Conditionally Risky Behavior vs Expected Value Maximization in Evolutionary Games

Patrick Roos[1,2] and Dana Nau[1,3,2]

[1] Department of Computer Science
[2] Institute for Advanced Computer Studies
[3] Institute for Systems Research
University of Maryland, College Park MD 20740, USA
{roos,nau}@cs.umd.edu

Abstract. Inspired by much empirical evidence of human decision-making under risk that does not coincide with expected value maximization, much effort has been invested into the development of descriptive theories of human decision-making involving risk (e.g. Prospect Theory, Regret Theory, SP/A Theory). An open question is how behavior corresponding to these descriptive models could have been learned or arisen evolutionarily, as the described behavior differs from expected value maximization. We believe that the answer to this question lies, at least in part, in the interplay between risk-taking and sequentiality of choice in evolutionary environments. We provide simulation results for evolutionary game environments where sequential decisions are made between risky and safe choices. Our results show there are evolutionary games in which agents that are sometimes risk-prone and sometimes risk-averse can outperform agents that make decisions solely based on the maximization of the local expected values of the outcomes.

1 Introduction

In most of the existing literature on models of decision making under risk, the construction of such models is approached primarily through the analysis of single or one-shot decisions among a set of choices, which generally are choices among lotteries with different types of payoff distributions and thus potentially different risks. Under the most traditional model of decision making, expected utility theory, a rational agent seeks to maximize the expected utility of choice outcomes and abides to the axioms of utility, which define preference relations on states or lotteries under a von Neumann-Morgenstern utility function [1].

Empirical evidence of human decision making under risk shows that humans are sometimes risk averse, sometimes risk seeking, and even behave in ways that systematically violate the axioms of expected utility [2]. Expected utility theory can account for different attitudes towards risk, such as risk-aversion or risk-seeking, through certain von Neumann-Morgenstern utility functions (e.g. [3]). Such risk propensities can differ greatly from simple expected value considerations on prospective outcomes. Researchers have invested much effort into constructing utility functions that appropriately model human decision making under risk under the expected-utility model (e.g. [3–5]). Researchers have also constructed alternative descriptive theories of decision making that claim to correspond more closely to how humans make decisions involving risk. Among the most popular of these alternative models are prospect theory [2, 6], regret theory [7], and SP/A (Security-Potential/Aspiration) theory [8–10]. One advantage of these models is that they more explicitly or perhaps more naturally model some of the mechanics involved in human decision making processes. For example, state-dependent attitudes toward risk are modeled in prospect theory by using a reference point with respect to which prospective outcomes can be interpreted as potential gains or losses, and are modeled in SP/A theory by including an aspiration level as an additional decision criterion in decisions involving risk.

A question that has received much less attention is how behaviors corresponding to the above decision-making models, or any other empirically documented risk-related behavior that differs from expected value maximization, could have arisen in human evolution or are learned in societies.

We believe that one part of the answer to this question is the interplay between risk-taking and sequentiality of choices; and in this paper we present simulation results to support this belief. In the spirit of numerous previous studies which have used evolutionary game simulations to explore and derive explanations for how the phenomenon of cooperation can arise in populations of individuals (e.g. [11–16]), we use an evolutionary game simulation approach to explore risk-related behavior.

In particular, we simulate a simple evolutionary game in which agents make sequential choices among lotteries that have equal expected value but different risks. Our experimental results demonstrate that depending on the game's reproduction mechanism, an agent that acts solely according to the local expected values of outcomes can be outperformed by an agent that varies its risk preference in ways suggested by descriptive models of human decision making.

2 Evolutionary Lottery Game

Our evolutionary game is based on a finite, homogeneous population model in which agents acquire payoffs dispensed by lotteries. In each generation, each agent must make a sequence of n choices, where each choice is between two lotteries with equal expected value but different risks. One lottery has a certain outcome of payoff 4 (with probability 1), we call this the *safe* lottery. The other lottery gives a payoff of 0 with probability 0.5 and a payoff of 8 with probability 0.5, we call this the *risky* lottery. Both lotteries have an expected value of 4, the only difference is the payoff distribution. Note an expected-utility player with a utility function that is the identity function on the values of the lottery would be indifferent between the two choices presented, as it would seek to maximize the expected value of its choice. We include all possible pure strategies in the environment in equal frequencies, as described in the following section. Through this setup, we don't model risk propensities of agents explicitly, but each of the agent types' strategies can be interpreted as different attitudes towards risk: some risk-averse, some risk-seeking, and some varying depending on previous choice outcomes.

2.1 Sequentiality and Strategies

We consider the case where agents make a single, one-shot choice among the two lotteries (i.e., $n = 1$), and the case where agents make two sequential choices (i.e., $n = 2$). If we only consider pure strategies, then for the case where $n = 1$ there are two possible strategies:

- Strategy S: always choose the safe lottery;
- Strategy R: always choose the risky lottery.

For the case where $n = 2$, there are six possible pure strategies:

Strategy	1st lottery	2nd lottery
SS	choose safe	choose safe
RR	choose risky	choose risky
SR	choose safe	choose risky
RS	choose risky	choose safe
R-WS	choose risky	choose safe if 1st lottery was won, risky otherwise
R-WR	choose risky	choose risky if 1st lottery was won, safe otherwise

2.2 Reproduction

Our evolutionary model uses non-overlapping populations: once all lottery choices have been made and payoffs have been dispensed, all agents reproduce into the next generation (a new population). Note that reproduction in evolutionary game theory does not necessarily need to be biological reproduction, but can be interpreted as a mechanism describing the process of learning [17] or the spread and adoption of cultural memes or behavioral traits [18], e.g. [19].

Replicator dynamics. In the literature on evolutionary game theory and social simulation, replicator dynamics are the most widely used reproduction mechanism. Under the replicator dynamic, the payoffs received by agents are considered to be a measure of the agent's fitness, and agent types reproduce proportional to these payoffs [20, 21]:

$$p^{new} = p^{curr} \frac{pay(agent_i)}{\overline{pay}} \tag{1}$$

where p^{curr} is the proportion of agents of type i in the current population, p^{new} is the corresponding proportion in the next generation, $pay(agent_i)$ is the summed payoff an agent of type i received from all games played, and $\overline{pay}$ is the average payoff received by all agents in the population. An agent's type is simply the strategy it employs to make choices among lotteries.

While the replicator dynamics enjoy much theoretical support (e.g., [20–22]), some of its characteristics, such as a potentially unlimited amount of reproduction (barring a population size limit) and guaranteed reproduction with some positive payoff, seem unrealistic in most real world environments. For example, organisms in real environments may die if they do not acquire a certain amount of resources, and there is certainly also a limit on the amount of offspring organisms can reproduce.

Tournament selection. Because of the above considerations, we felt that instead of focusing solely on replicator dynamics, it would be important to compare it with at least one other reproduction mechanism. For this purpose, we chose tournament selection, which is probably the second most widely used reproduction mechanism in the evolutionary game theory literature (e.g., [14, 23, 24]).

Under tournament selection, each agent in the population is matched up with a randomly drawn other agent in the population and the agent with the higher acquired payoff is reproduced into the next generation. If the payoffs of the matched agents is equal, one of the two agents is chosen at random to reproduce.

3 Preliminary Analysis

We now consider four different versions of our evolutionary lottery game, using all four combinations of the following parameters: the number of sequential choices ($n = 1$ or $n = 2$), and the reproduction mechanism (tournament selection or replicator dynamics).

We are interested in the dynamics of agent type (i.e. strategy) frequencies as our population of agents evolves over time. To anticipate the performance of different agents in our simulations, it is important to take a closer look at the nature of the reproduction mechanisms. Under replicator dynamics, agents' rate of change of population frequency is directly proportional to the payoff received relative to the average population payoff. Tournament selection on the other hand acts like a probabilistic, population dependent, threshold step-function, where the particular threshold an agent needs to achieve in order to produce one offspring is the payoff of the randomly drawn opponent from the population.

Recall that for $n = 1$ (i.e., the single choice game) there are only two pure strategies, S and R. S will always receive a payoff of 4, while R will have a 50% chance to receive a payoff of 8 and a 50% chance to receive 0. Hence in each case, the expected value is 4. Thus under replicator dynamics, by equation (1) we expect neither type of agent to have an advantage. Under tournament selection, an R agent will have a 50% chance to beat an S agent and a 50% to lose, thus we expect neither agent to have an advantage here either.

For $n = 2$ (i.e., a sequence of two lottery choices), there are six pure strategies. As before, they all have an expected value of 4 at each lottery choice, thus a total expected value of 8 for the sequence of two choices. For each of these strategies except *R-WS* and *R-WR*, the probability of being above the expected value equals the probability of being below the expected value; but for *R-WS* and *R-WR*, the two probabilities differ. *R-WS* has a 50% chance of receiving a payoff of 12, since it always chooses the safe option after the first lottery was one. It has a 25% chance of receiving the exact expected value 8, which occurs when only the second risky choice is won. Hence it has only a 25% chance of acquiring a payoff below the expected value. (Similarly, *R-WR* has only a 25% chance of acquiring a payoff above the expected value.) This lead us to hypothesize that there are circumstances in which *R-WS* will do better than the other strategies. The purpose of our simulation, described in the next section, is to test this hypothesis.

4 Simulation Results

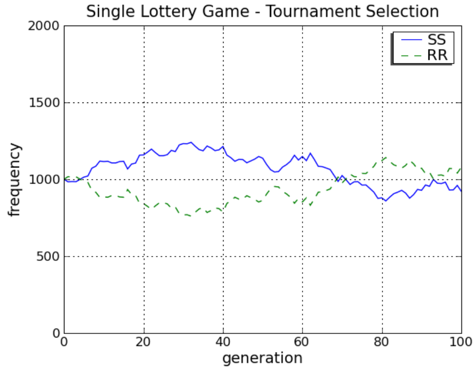
We have run simulations of our evolutionary lottery game using all four combinations of the following parameters: the number of sequential choices ($n = 1$ or $n = 2$), and the reproduction mechanism (tournament selection or replicator dynamics). The types of agents were the ones described in Section 2.1. All simulations started with an initial population of 1000 agents for each agent type and were run for 100 generations, which was sufficient for us to observe the essential population dynamics.

Figures 1(a,b) show the frequency for each type of agent when $n = 1$. As we had expected, both *S* and *R* performed equally well (modulo some stochastic noise) regardless of which reproduction mechanism we used.

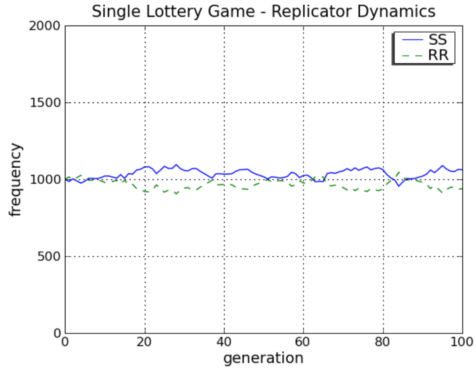
For $n = 2$ (Figures 1(c,d)), the results are more interesting and differ depending on the reproduction mechanism used. Under replicator dynamics, all of the strategies performed equally well and remained at their frequency in the original population. But under tournament selection, the conditional strategy *R-WS* outperformed the other strategies. *R-WS* rose in frequency relatively quickly to comprise the majority ($> 2/3$) of the population and remained at this high frequency throughout subsequent generations. One surprise, which we discuss in the next section, was that the two unconditional strategies *SR* and *RS* fell slightly in population but then remained, comprising the proportion of the population not taken over by *R-WS*.

Table 1. Performance of agent types for the single ($n = 1$) and sequential ($n = 2$) lottery game simulations under tournament selection and replicator dynamics.

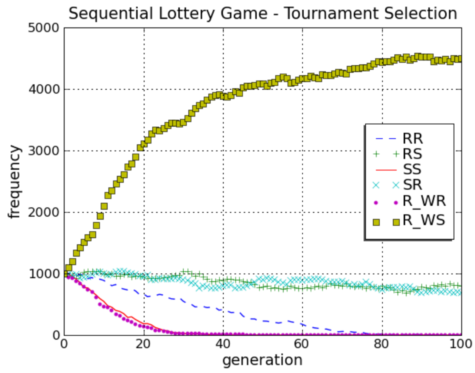
	Tournament	Replicator
Single	***S*, *R* same**	***S*, *R* same**
Sequential	***R-WS* best**	**all same**



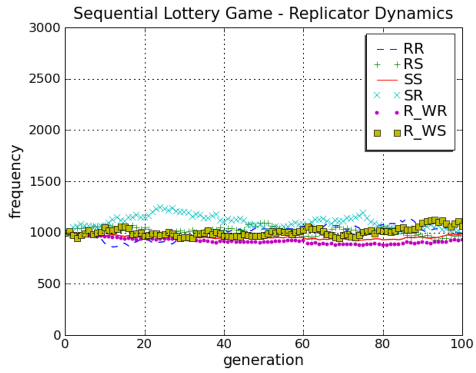
(a) $n = 1$, tournament selection



(b) $n = 1$, replicator dynamics



(c) $n = 2$, tournament selection



(d) $n = 2$, replicator dynamics

Fig. 1. Agent type frequencies for all four simulations over 100 generations.

5 Discussion

Table 1 summarizes the experimental results as they relate to the hypothesis we stated at the end of Section 3. The results confirm our hypothesis that there are circumstance under which the *R-WS* agent performs better than other agent types. In particular, it performed much better than all other agents in the sequential lottery using tournament selection simulation.

The following subsections discuss the impact of the reproduction mechanisms used, the population dynamics observed, and how the results relate to theoretical work concerning models of decision making under risk.

5.1 The Role of Reproduction Mechanisms

The reproduction mechanism played a central role in the results of our simulations. We now analyze the impact of the reproduction mechanisms by examining the expected payoff distributions of the agent types.

Table 2 lists these distributions of payoffs that agents are expected to receive from their choices in one generation of the sequential lottery game.

Table 2. Expected payoff distributions for all agent types in the sequential lottery game.

agent	*R-WS*			*R-WR*			*SR*		*RS*		*SS*	*RR*		
payoff	**12**	**8**	**0**	**16**	**8**	**4**	**12**	**4**	**12**	**4**	**8**	**16**	**8**	**0**
probability	.5	.25	.25	.25	.25	.5	.5	.5	.5	.5	1	.25	.5	.25

We can see that the *R-WS* agent had a 50% chance of acquiring a payoff of 12, a 25% chance of acquiring a payoff of 8, and a 25% chance of acquiring 0. Under tournament selection, *R-WS* had an advantage over the other agents because it had an increased probability of achieving a payoff at or above a certain reproduction threshold. This threshold is the payoff of a randomly drawn opponent, which has an expected value of 8 equal to the expected value of the lotteries. *R-WS* pays for this enlarged chance of being above the threshold through a small chance of doing much worse (payoff 0) than the summed expected utilities, which occurs when the first and the second risky choice is lost.

Since tournament selection only considers whether or not the agent's payoff is better than another agent's in order to decide whether the agent reproduces, the *extent* to which the agent is better is not significant. In contrast, the replicator dynamics define reproduction to be directly proportional to the amount by which the agent's payoff deviates from the population average. In this case the small chance of *R-WS* of being significantly below the expected value balances against the agent's larger chance of being slightly above it. Thus, under replicator dynamics, the *R-WS* agents had no advantage.

5.2 Further Population Dynamics

As noted in Section 4, the *SR* and *RS* agents did not go extinct in the sequential-lottery tournament-selection simulation. The reason *SR* and *RS* remained can be explained by considering all the agents' payoff distributions expected in a generation (Table 1). If we compare the payoff distribution of *SR* and *RS* with that of *R-WS*, we see that if these agents are matched up with each other in tournament selection there is an equal chance that either of the strategies reproduces. Thus once all other strategies are extinct, the population frequencies remain approximately unchanged. The reason *R-WS* rises in frequency so much faster early on is because *R-WS* has a significantly higher chance of beating an agent from the rest of the population. Against *SS* for example, *R-WS* has a 62.5% chance of winning: 50% of the time the payoff of 12 beats the sure payoff of 8 by *SS* and 1/2 of the time the two players are matched with equal payoff of 8 (25% chance), *R-WS* is favored. *SR* and *RS* on the other hand only have a 50% chance of winning against *SS*. Similar relations hold for *RR* and *R-WR*.

This shows an interesting dynamic of population-dependent success of agents:

- In an environment that contains *SR*, *RR*, and *R-WS* and no other strategies, all three will do equally well.
- In an environment that contains *SR*, *RS*, *SS* and *RR* and no other strategies, all four will do equally well.
- In an environment that contains *SR*, *RS*, *SS*, *RR*, and *R-WS*, *R-WS* will increase and the other strategies will decrease until *SS* and *RR* become extinct, at which point *SR* and *RS* will stop decreasing.

5.3 Relations to Alternative Decision Making Models

The manner in which the *R-WS* strategy deviates from expected value maximization in our lottery game can be characterized as risk-averse (preferring the safe choice) when doing well in terms of payoff and risk-prone (preferring the risky choice) otherwise. Similar risk behavior is suggested by models such as prospect theory [2, 6] and SP/A theory. In prospect theory, people are risk-seeking in the domain of losses and risk-averse in the domain of gains relative to a reference point. In SP/A theory [10], a theory from mathematical-psychology, aspiration levels are included as an additional criterion in the decision process to explain empirically documented deviations in decision-making from expected value maximization.

One explanation for the existence of decision-making behavior as described by such models is that the described behavioral mechanisms are hardwired in decision makers due to past environments in which the behaviors provided an evolutionary advantage [25]. Another interpretation, not necessarily unrelated, is that the utility maximized by decision makers is not the payoffs at hand, but a different perhaps not obvious utility function. Along these lines, [26] proposes a model of decision making that includes probabilities of success and failure relative to an aspiration level into an expected utility representation with a discontinuous (at the aspiration level) utility function. Empirical evidence and analysis provided in [27] provide clear support for the use of probability of success in a model of human decision making. All these descriptive theories provide for agents to be sometimes risk-prone and sometimes risk-averse, depending on their current state or past outcomes, such as the *R-WS* in our simulations.

The sequentiality of choices in our game simulations allow for such state-dependent risk behavior to be explicitly modeled. One could theoretically model the sequential lottery game in normal form, i.e. reduce the choices to a single choice between the payoff distributions listed in Table 2. Doing so would provide essentially equivalent results except that the asymmetry in the payoff distribution of lotteries would be the determining factor of agent successes. In such a representation however, the analysis of risky and safe choices, and agents' preferences among them becomes blurred. In fact, we believe that a tendency towards modeling games in normal form often leads people to overlook the impact of sequentiality on risk-related behavior.

We believe our results show that tournament selection models an important mechanism that can lead to the emergence of risk-taking behavior with similar characteristics to that captured in alternative, empirical evidence-based models of decision making like the ones discussed above. Whenever reproductive success is not directly proportional to payoff (i.e., a reproduction mechanism other than the pure replicator dynamics),[4] risk propensities that differ from expected value maximization have the opportunity to be more successful than agents that solely consider expected value in their local choices. This suggests that there are many other reproduction mechanisms for which expected-value agents can be outperformed by agents that vary their propensities toward risk-taking and risk-averseness.

6 Conclusion

Our simulation results in several evolutionary lottery games demonstrate how sequentiality and reproduction can affect decision making under risk. Our simulations show that a strategy other than expected-value maximization can do well in an evolutionary environment having the following characteristics:

- At each generation, the agents must make a sequence of choices among alternatives that have differing amounts of risk.
- An agent's reproductive success is not directly proportional to the payoffs produced by those choices. We specifically considered tournament selection; but as pointed out in Section 5.3, we could have gotten similar results with many other reproduction mechanisms.

The most successful strategy in our simulations, namely the *R-WS* strategy, exhibited behavior that was sometimes risk-prone and sometimes risk-averse depending on its success or failure in the previous lottery. This kind of behavioral characteristic is provided for in descriptive theories of human decision making based on empirical evidence. It is not far-fetched to suppose that when human subjects have exhibited non-expected-value preferences in empirical studies, they may have been acting as if their decisions were part of a greater game of sequential decisions in which the success of strategies is not directly proportional to the payoff earned. Apart from a purely biological interpretation, in which certain behavioral traits are hardwired in decision-makers due to past environments, perhaps such empirical studies capture the effects of the subjects' learned habit of making decisions as part of a sequence of events in their daily lives.

Our results also demonstrate (see Section 5.2) that the population makeup can have unexpected effects on the spread and hindrance of certain risk propensities. This may be an important point to consider, for example, when examining decision-making across different cultures or societies.

[4] We say "pure" here because replicator dynamics can be modified to make reproductive success not directly proportional to payoff. For example, if a death rate (e.g. [12]) is implemented as a payoff-dependent threshold function, we might expect risk propensities to differ depending on whether an agent is above or below that threshold, similar to an aspiration level in SP/A theory.

In conclusion, our simple lottery game simulations are a first step in exploring evolutionary mechanisms which can induce behavioral traits resembling those described in popular descriptive models of decision making. A specific related topic to explore is how the prospect-theoretic notion of setting a reference point may relate to evolutionary simulations with sequential lottery decisions. In general, there is much more opportunity for future work to use simulation for the purpose of exploring or discovering the mechanisms which induce, possibly in a much more elaborate and precise manner, the risk-related behavior characteristics described by prospect theory or other popular descriptive decision making models based on aspiration levels.

Acknowledgements

We would like to thank the reviewers for their helpful comments and suggestions. This work was supported in part by AFOSR grant FA95500610405, NAVAIR contract N6133906C0149, DARPA's Transfer Learning Program, DARPA IPTO grant FA8650-06-C-7606, and NSF grant IIS0412812. The opinions in this paper are those of the authors and do not necessarily reflect the opinions of the funders.

References

1. Neumann, J.V., Morgenstern, O.: Theories of games and economic behavior. Princeton University Press Princeton, NJ (1944)
2. Kahneman, D., Tversky, A.: Prospect theory: An analysis of decision under risk. Econometrica **47** (1979) 263–291
3. Friedman, M., Savage, L.J.: The utility analysis of choices involving risk. The Journal of Political Economy **56**(4) (August 1948) 279–304
4. Arrow, K.J.: Essays in the theory of risk-bearing. Markham, Chicago (1971)
5. Rabin, M.: Risk aversion and Expected-Utility theory: A calibration theorem. Econometrica **68**(5) (September 2000) 1281–1292
6. Tversky, A., Kahneman, D.: Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty **5** (1992) 297–323
7. Loomes, G., Sugden, R.: Regret theory: An alternative theory of rational choice under uncertainty. The Economic Journal **92** (December 1982) 805–824
8. Lopes, L.L.: Between hope and fear: The psychology of risk. Advances in Experimental Social Psychology **20** (1987) 255–295
9. Lopes, L.L.: Re-modeling risk aversion. In von Furstenberg, G.M., ed.: Acting under uncertainty: Multidisciplinary conceptions. Boston: Kluwer (1990) 267–299
10. Lopes, L.L., Oden, G.C.: The role of aspiration level in risky choice: A comparison of cumulative prospect theory and sp/a theory. Journal of Mathematical Psychology **43** (1999) 286–313
11. Axelrod, R., Hamilton, W.D.: The evolution of cooperation. Science **211**(4489) (March 1981) 1390–1396
12. Nowak, M.A., Sigmund, K.: A strategy of win-stay, lose-shift that outperforms tit for tat in the prisoner's dilemma game. Nature **364** (1993) 56–58
13. Nowak, M.A., Sigmund, K.: Tit for tat in heterogeneous populations. Nature **355** 250–253

14. Hales, D.: Evolving specialisation, altruism, and group-level optimisation using tags. In Sichman, J.S., Bousquet, F., Davidsson, P., eds.: MABS. Volume 2581 of Lecture Notes in Computer Science., Springer (2002) 26–35
15. Bowles, S., Gintis, H.: The evolution of strong reciprocity: cooperation in heterogeneous populations. Theoretical Population Biology **65**(1) (February 2004) 17–28
16. Nowak, M.: Five rules for the evolution of cooperation. Science **314** (2006) 1560–1563
17. Harley, C.B.: Learning the evolutionarily stable strategy. Journal of Theoretical Biology **89** (1981) 611–633
18. Dawkins, R.: The Selfish Gene. New York: Oxford University Press (1976)
19. Hales, D.: An open mind is not an empty mind: Experiments in the meta-noosphere. Journal of Artificial Societies and Social Simulation **1**(4) (1998)
20. Hofbauer, J., Sigmund, K.: Evolutionary game dynamics. Bulletin of the American Mathematical Society **40**(4) (July 2003) 479–519
21. Gintis, H.: Game Theory Evolving: A Problem-centered Introduction to Modeling Strategic Behavior. Princeton University Press (2000)
22. Imhof, L.A., Fudenberg, D., Nowak, M.A.: Evolutionary cycles of cooperation and defection. Proc Natl Acad Sci U S A **102**(31) (August 2005) 10797–10800
23. Riolo, R.L., Cohen, M.D., Axelrod, R.: Evolution of cooperation without reciprocity. Nature **411** (2001) 441–443
24. Hales, D.: Searching for a soulmate - searching for tag-similar partners evolves and supports specialization in groups. In Lindemann, G., Moldt, D., Paolucci, M., eds.: RASTA. Volume 2934 of Lecture Notes in Computer Science., Springer (2002) 228–239
25. A L Houston, J M McNamara, M.D.S.: Do we expect natural selection to produce rational behaviour? Philosophical Transactions of the Royal Society B **362** (2007) 1531–1543
26. Diecidue, E., Ven, J.V.D.: Aspiration level, probability of success and failure, and expected utility. International Economic Review **49**(2) (May 2008) 683–700
27. Payne, J.W.: It is whether you win or lose: The importance of the overall probabilities of winning or losing in risky choice. Journal of Risk and Uncertainty **30** (2005) 5–19