STAIRS 2012
K. Kersting and M. Toussaint (Eds.)

doi:10.3233/978-1-61499-096-3-113

Adopting a Risk-Aware Utility Model for Repeated Games of Chance

Nathaniel GEMELLI [a,b], Jeffrey HUDACK [a,b] and Jae OH [a]
[a] *Computer Science Department, Syracuse University, Syracuse, NY, USA*
[b] *Information Directorate, Air Force Research Laboratory, Rome, NY, USA*

Abstract. We present a risk-aware utility model for action selection in 2-player, non-cooperative, repeated games of chance. Our model augments expected utility calculations and adapts to changes in accumulated utility, demonstrating rational game play. Motivated by risk aversion and the utility of wealth, our model is parameterized by an agent's wealth, the payoffs of a game, and the probability associated with gain and loss. Using expected utility combined with our model, we can impose desired behavior onto an agent that mimics more closely the types of behaviors we see in social and economic situations where risk is involved. We define our model theoretically and present empirical results showing the effectiveness of a risk-aware utility model against a Nash equilibrium mixed strategy in a repeated game of chance.

Keywords. game theory, risk aversion, utility theory, expected utility

1. Introduction

The application of expected utility for game-theoretic analysis is an effective method for an agent to determine preference of one action over another. In the simplest case, expected utility can be computed as a function of the payouts assigned by the game definition. However, an agent may have additional constraints or motivations that will influence its expected utility of the actions available. In utility theory, risk attitudes with respect to investment are a way of personalizing the utility for an individual agent [1]. A *risk averse* agent will prefer actions that are more likely to provide a safe payoff while a *risk seeking* agent will prefer riskier actions typically associated with higher payoffs. In general, a *risk neutral* approach is taken as a mathematical convenience to avoid making assumptions about an agent's internal representation of value with respect to the payoffs specified by the game. We consider the effects of risk attitude on the expected utility of what we refer to as a *risk-aware* agent and demonstrate greater performance in the play of a repeated game than a Nash equilibrium mixed strategy.

The *utility of wealth* [2] specifies that static payouts prescribed by a game have less value to an agent when that agent has more wealth. This is an intuitive concept that seems to capture the motivation behind a wide range of human behavior in sociology and economics. For example, consider two risk-aware agents;

one who is relatively wealthy as compared to another who has only \$20. Given a high chance of winning, the wealthy agent is not as interested in making a bet for \$10 as an agent with only \$20 might be. For the wealthy agent, the amount is insignificant, while the other agent considers it an opportunity to increase their wealth significantly. It is not the case that the payout has a different intrinsic value, but each agent has a unique internal utility function that values the payout differently. We model risk aversion in an agent using wealth-based utility.

In games where agents are competing to accumulate wealth, and must play the game repeatedly, *relative wealth* becomes important. Rather than focusing singularly on choosing actions that increase wealth, an agent must also consider its own wealth with respect to the wealth of other agent(s) that it is competing against. For an agent that values having more wealth than another agent, its risk attitude is concerned with maintaining a lead and only taking risks when the lead is sufficient. Alternatively, an agent may be competing with a large number of agents and values maintaining a position in a top percentile of the population for accumulated wealth. In these cases, a risk-aware model must change dynamically as the wealth of both the agent and the competing population changes.

The remainder of the paper is presented as follows. Section 2 discusses risk behavior and utility functions, as well as related work. In Section 3, we define a 2-player repeated investment game. In Section 4, we introduce a model that incorporates risk-aversion into expected utility for action selection. To measure the performance of our approach we simulate a repeated investment game in which a risk-aware agent with varying risk attitudes plays against an agent that is not risk-aware, as described in Section 5. Our results in Section 5 demonstrate two methods of using the risk-aware model to maximize the chance of winning, or to minimize the time spent playing while still maintaining an advantage. In Section 6, we summarize our findings and discuss future work.

2. Background

In the following sections we will often use the terms “wealth” or “investment.” This provides an intuitive way of explaining the concepts that we have borrowed from utility theory, but should not be construed as a limitation on the applicability of our work to only monetary investment options or wagers. The results of this work would apply to any game with transferable utility, as defined in [3]. In this section we will discuss the properties of utility functions, how we will use those properties to produce desired agent behavior, and related work.

2.1. Risk Behavior

To study risk behavior, we will define a utility function, $U(\cdot)$, that is twice differentiable and demonstrates two unique properties: *non-satiation* and *risk attitude*. The non-satiation property is determined by the first derivative of $U(\cdot)$ and is interpreted as ‘more wealth is always better than less or none’. For a utility function to be non-satiated its first derivative must be greater than zero, $U'(\cdot) > 0$. Risk attitude is the rate of change, or rate of gratitude for further accumulation of money. Risk attitude can take on the characteristic of being risk averse, risk seeking, or risk neutral, all of which have distinct second derivatives.

$$U''(\cdot) = \begin{cases} \textbf{risk averse} \text{ if } U''(\cdot) < 0; \\ \textbf{risk neutral} \text{ if } U''(\cdot) = 0; \\ \textbf{risk seeking} \text{ if } U''(\cdot) > 0. \end{cases} \quad (1)$$

Risk averse behavior is an individual's reluctance to invest in a risky option and preference for a more certain or even guaranteed result. Risk averse behavior, with regards to utility, is a concave function. In contrast, risk seeking behavior is a convex function that reflects an individual's preference for risk when presented with investment options. Typically, individuals are risk averse, but risk seeking behavior has been studied and observed in the negative domain of value functions in work such as Prospect Theory [4]. Risk neutral behavior gives no preference to a choice between two equally expected value options.

Non-satiation and risk aversion is analog to the *Diminishing Marginal Utility* (DMU), first introduced by Bernoulli [5]. DMU is defined by two laws: (1) "The marginal utility of each (homogeneous) unit decreases as the supply of units increases (and vice versa)" and (2) "The marginal utility of a larger-sized unit is greater than the marginal utility of a smaller-sized unit (and vice versa)" [6]. We are most interested in law 1; as the total amount of wealth increases (decreases), the marginal utility of wealth decreases (increases).

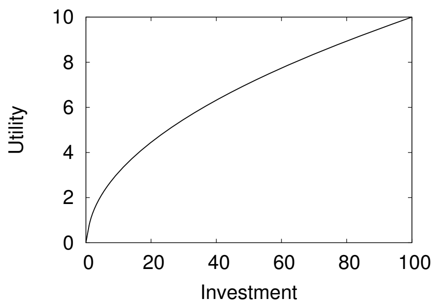
In this work, we consider risk aversion as a means of determining the appropriate amount of wealth or resource to invest in a particular game against a particular opponent. In Section 4 we will discuss how risk averse behavior is tailored to individual game scenarios.

2.2. Risk-Averse Expected Utility

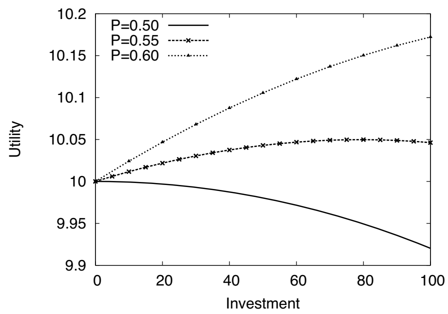
Risk averse agents will seek to limit the amount of risk they are willing to take in an investment. Figure 1a depicts a utility function with the characteristic concave down curve of a risk averse agent with non-satiation. The utility function is the square root function shown in Eq. (2). As the amount of wealth increases, the marginal utility of that wealth decreases. This equation exhibits the DMU of wealth and evaluates as risk averse ($U'(w) = 0.5w^{-0.5} > 0$) and non-satiated ($U''(w) = -0.25w^{-1.5} < 0$).

$$U(w) = \sqrt{w} = w^{0.5} \quad (2)$$

We use this utility function to evaluate a set of actions, $A = \{a_0, a_1, ..., a_n\}$. These actions are a set of monotonically increasing investment options with respect to n. For any action, $a_i \in A$, there is a probabilistic potential for gain, $G(a_i)$, and a probabilistic potential for loss, $L(a_i)$. We assume that both $G(a_i)$ and $L(a_i)$ return a real-valued number and that an agent has a certain amount of wealth, w, that is affected positively and negatively by $G(a_i)$ and $L(a_i)$, respectively. Given some probability P of winning, we calculate the *expected utility* of action a_i, $EU(a_i)$, with the general form:



(a)



(b)

Figure 1. (a) The square root utility function that is risk averse and non-satiated. (b) Utility curves for the square root expected utility function with maximums at 0, 79, and 100 for $P =$ 0.50, 0.55, and 0.60, respectively.

$$EU(a_i) = [P \times U(w + G(a_i))] + [(1 - P) \times U(w - L(a_i))] \tag{3}$$

As an example, assume we have \$100 to invest in a "risky" investment game. The set of actions available to us is to invest $\$0 \leq n \leq \$100$ of our money in the game. This investment has a potential positive gain $G(n) = (x \times n)$, and a potential negative loss $L(n) = (y \times n)$ where x and y are the rate of return on investment n. Our goal is to determine the appropriate amount, n, to place in this investment that will maximize $EU(n)$, where

$$EU(n) = [P \times U(100 + G(n))] + [(1 - P) \times U(100 - L(n))] \tag{4}$$

The graph in Figure 1b shows Eq. (4) with $x = 25\%$ and $y = 25\%$. Notice that an agent that is risk averse and given an equally probable positive and negative return of equal value will reject the opportunity to invest. However, the same risk averse agent, given a higher probability of positive return ($P = 0.55$) will venture into the investment. At a certain point, P increases high enough that an agent will invest all of his money (such as $P = 0.60$), even though there may still be a risk of a loss occurring.

What we are actually observing in Figure 1b is a set of *indifference* or *utility curves* [7], each with its own unique maximum value. For example, the maximum of the $P = 0.55$ curve is \$79. An investment option that maximizes the utility value for any given utility curve can be interpreted as the "optimal" investment option for that agent with that given expected utility function. This is important as it gives us a way of calculating a single option, or action, that the agent should take given an investment's parameters. In Sections 4 and 5 we will be using our generalized form of the expected utility function defined in Eq. (3) as a decision mechanism for 2-player non-cooperative game settings where wealth is accumulated. The return rates x and y used in the example above, and being used to define the utility curves in Figure 1b, are weighting factors. In instances where the values of x and y are not equivalent we find similar results. When $x > y$, the maximum of the utility curve shifts right. Conversely, when $x < y$, the maximum of the utility curve shifts left.

2.3. Related Work

Wealth-based utility functions for games of chance (risk) are not new. In economics there has been much investigation into choices involving risk and risk averse behavior. From the earliest work of Friedman and Savage [8], Pratt [9] and Kahneman and Tversky's Prospect Theory [4] to more recent advances of Holt [10] and Lybbert [11], there has been no shortage of attempts to explain why humans make the investment decisions they do. Work focused on human investment behavior motivates our interest in defining how computer agent interactions will occur when faced with risk-based investment decisions that are not under the agent's complete control. In [12], the authors look at varying behavioral game play in the a Cournot duopoly game based on an exponent in the utility function. They mitigate risk based on different player types that result from observing opponent behavior. However, there is no dynamic update function that takes proportional wealth into consideration. The shift in behavior is based on the type of player persona that has been adopted at the current moment.

Instead of adjusting risk from a utility function perspective, there has been work in adjusting the actual payouts of the game being played. In [13], the authors look at heterogeneous payoffs from an evolutionary game theory perspective. Based on average interaction within groups of agents, an agent can use an aging function to update how long it wishes to continue participating in a social game, specifically the spatial iterated prisoner's dilemma. They also look at dynamic payoff values that are based on experience in the simulation where the agents' payoff matrix is changing. Rather than change the game structure, our model maximizes an investment based on the fixed normal form. [14] and [15] also promote using evolutionary game theory to study dynamic payoffs in the spatial iterated prisoner's dilemma game. In [16], the authors suggest changing the payout structure of the normal form game. This could represent a shift in the utility preferences a player has for a particular result but it does not represent the actual game being played. The games being studied are stochastic games as opposed to repeated stage games.

3. Game Description

The *investment game* is a 2-player, zero-sum, game of chance that is characterized by a probability of winning and a maximum investment. The probability of winning the game (and conversely an opponent losing the game) is determined by Nature and is therefore out of the agents' control. The maximum investment is fixed and known to all agents and the payout resulting from both players' action is determined by the minimum investment. For example, if one agent invests 10 and the other agent invests 2, then the maximum payout (or loss) for both agents is 2. For this work we consider a game with two agents that have the same maximum investment value, represented as an integer. This could easily be extended to more players with varying maximum investment values and real-valued investment options.

Given a maximum investment $I_M \in \mathbb{Z}_{>0}$, a set of actions $A = \{a_1, a_2, ..., a_n\}$, with action a_i representing an investment of value i, and a probability $P \in [0,1]$, we can define the payoff for both agents. Let N be the action taken by Nature, sampled from a continuous uniform distribution $\mathcal{U}(0,1)$. Let p_1 and p_2 be the payoff for the row and column player, respectively, such that $p_1, p_2 \in \mathbb{Z}$. For all actions $a \in A$ we define the outcome function $O : a \times a \rightarrow (p_1, p_2)$ that maps a pair of actions to the payoffs for both agents.

$$O(a_1, a_2) = \begin{cases} (min(a_1, a_2), -min(a_1, a_2)) & \text{if } N < P \\ (-min(a_1, a_2), min(a_1, a_2)) & \text{otherwise} \end{cases} \tag{5}$$

The full extensive form with maximum investment I_M and probability of Player 1 winning P is shown in Figure 2. The information sets in this game are the result of the random selection of Nature's move and the move played by the other agent. Because the maximum investment is variable, the actual size of the game may vary.

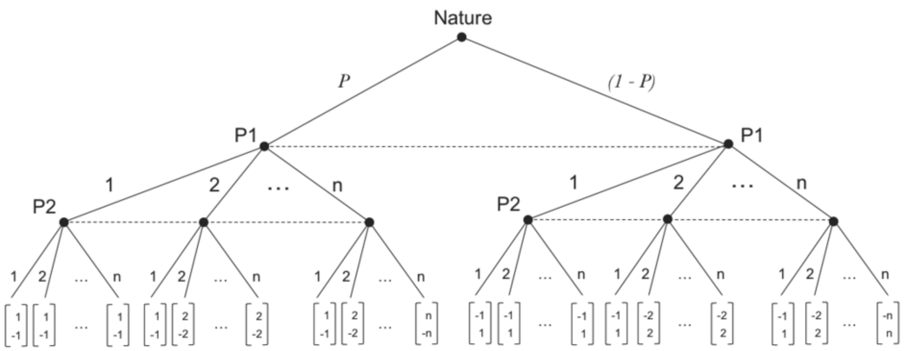


Figure 2. Extended form representation of the investment game.

3.1. Utility Function for the Investment game

In the remainder of the paper, we will take the perspective of the row player, and it is implicitly assumed that i represents the row player and j represents the column player. We will formulate our utility function in the same general form as that of Eq. (3). Assuming we have a current wealth value of w and Nature's move is determined by the random variable P, we define the expected utility for the investment game as,

$$EU(a) = [P \times U(w + \overline{O}_i(a))] + [(1 - P) \times U(w - \overline{O}_i(a))] \tag{6}$$

and the payoff of a particular investment to the row player is,

$$\overline{O}_i(a) = \sum_{j=1}^{n} [s_{opp}(a_j) \times O(a, a_j)] \tag{7}$$

Where n is the maximum number of actions available to both players and $s_{opp}(a_j)$ is a probability density function representing the support the opponent player gives to action a_j. We adopt the definition of support from a game theory perspective; the probability of selecting a particular action from the set of all available actions, as in a mixed strategy [1]. $\overline{O}_i(a)$ is a single value, real number, that represents the weighted average of an investment option against an opponent with some assumed mixed strategy.

4. Risk-Aware Utility Model

Incorporating the concepts of risk aversion and non-satiation to achieve risk-aware behavior will require us to parameterize the utility function given in Eq. (2). We will define a new utility function we call the *Risk-Aware Utility* (*RAU*) model, and parameterize it with the variable λ, which defines an agents' diminishing marginal utility on w,

$$RAU(w, \lambda) = w^{\lambda} \tag{8}$$

where w is the resulting level of wealth from the investment, such that $w = w_p + G(a_i)$ or $w = w_p - L(a_i)$, and w_p is the level of wealth prior to the investment result. The RAU model can alter the behavior we observed in the original utility function to become more or less risk averse based on the value of λ. When λ is bounded, $0 < \lambda < 1$, the behavior for the RAU is guaranteed to behave risk aversely. Allowing λ to vary provides an agent with the capability to shift behaviors between aggressive and non-aggressive game play in scenarios that dictate such behavior.

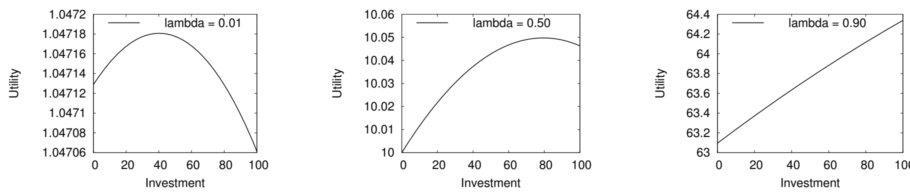
4.1. Effects of λ

Consider again the example from Section 2.2. When $P = 0.55$, the agent employing the risk averse utility function, $U(w) = w^{0.5}$, would invest \$79 in the investment. In the context of utility of wealth, not everyone would feel comfortable investing nearly 4/5's of their money in this investment. The only mechanism an agent has in the RAU model to control the utility curve resulting from expected utility calculations is to vary λ. The agent has no control over P. Figure 3 shows the resulting utility curves for different λ settings with $P = 0.55$. Given $P \neq 1$, in the $\lim_{\lambda \to 0}$, the utility function demonstrates increasing aversion to risk (Figure 3a) and will give the lowest possible investment the highest utility. Conversely, in the $\lim_{\lambda \to 1}$, the utility function demonstrates decreasing aversion to risk, nearing in on risk neutral behavior (Figure 3c).

We can now incorporate the RAU model into the expected utility function.

$$EU(a) = [P \times RAU(w_p + \overline{O}_i(a), \lambda)] + [(1 - P) \times RAU(w_p - \overline{O}_i(a), \lambda)] \tag{9}$$

Depending on the game being played and the utility function being augmented, the function for assigning a value to λ, or other parameterized variables within



(a) $\lambda = 0.01$ (b) $\lambda = 0.50$ (c) $\lambda = 0.90$

Figure 3. λ values of 0.01, 0.50, and 0.90 with utility curve maximums at 10, 79, and 100, respectively. $P = 0.55$, the x-axis is the amount of money for investment, and the y-axis is the relative utility between investment options.

the utility function, will vary. For our particular case of a square root function, any function that generates a real-valued number, $0 < \lambda < 1$, can be used. In Section 5 we introduce two definitions of λ with different agent goals in mind.

5. Experimental Design and Results

We simulate an environment where two agents play repeated instances of the investment game described in Section 3. Each agent starts with an equal amount of initial wealth W_I and the current wealth of each agent carries over between each stage game. The maximum investment I_M is also shared by the agents. The probability of row player winning, P, is drawn from a Normal Distribution with $\mu = 0.5$ and standard deviation $\sigma \in [0, 0.2]$. The simulation ends when one agent has exhausted all of its wealth or the maximum simulation time (100,000 steps) is reached. For all simulations we have an agent using the RAU model with a specified λ function against an agent playing a uniform random investment strategy. This scenario represents the RAU model agent playing against an opponent using a Nash equilibrium mixed strategy.

Because the mean probability is 50% the game is inherently fair in repeated play. However, we aim to show that the agent employing the RAU model can take advantage of subtle shifts in probability during single stage instances of the game that are imperceptible to the opponent. For that reason, the RAU agent is given the exact value of P at each episode of game play. Our intent is to demonstrate the importance of risk aversion in expected utility calculations as opposed to using the strict value of the payoffs. Since there is an inherent connection between P and λ, allowing the RAU agent to percieve P allows us to place our focus on studying the effects of λ. The advantage of perceiving P can be interpreted as an experienced player with expert domain knowledge playing a game with a less experienced palyer. Admittedly, more sophisticated techniques from opponent modeling [17] or Bayesian belief networks [18] would be a better way of representing this knowledge of P.

The λ function is the key to adjusting the risk attitude of the agent and should be designed with the goals of the agent in mind. We provide two functions for λ, each with different goals, but we acknowledge that they may not be the optimal approach for this environment. Our intention is to show that the behavior of the agent varies significantly when the λ function is changed to reflect different priorities. The λ function shown in Eq. (10) is intended to maximize wins in a repeated game against a single opponent. The λ function in Eq. (11) aims to balance winning with reducing the number of time steps the game is played. As λ is bounded below by 0, we set $\epsilon = 0.0001$. W_{rau} and W_{opp} are the current wealth levels of the RAU model agent (row) and opponent agent (column), respectively.

$$\lambda_1 = \begin{cases} 1 - W_{opp}/W_{rau} & \text{if } W_{opp} < W_{rau} \\ \epsilon & \text{otherwise} \end{cases} \tag{10}$$

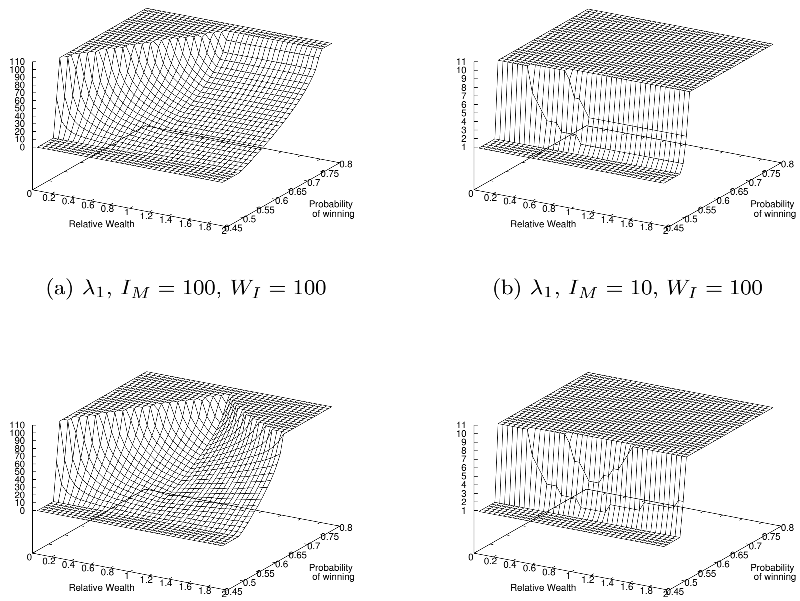
$$\lambda_2 = \begin{cases} 1 - W_{opp}/W_{rau} & \text{if } W_{opp} < W_{rau} \\ 1 - W_{rau}/W_{opp} & \text{if } W_{opp} > W_{rau} \\ \epsilon & \text{otherwise} \end{cases} \tag{11}$$

The graphs in Figures 4a and 4b show the investment that an agent will make using λ_1 with respect to the relative wealth and probability of winning. In this case, the agent will incur more risk when it has the lead (left side), but will be more risk averse when it has relative wealth below the opponent (right side of graph). The same investment values for an agent using λ_2 function are shown in Figures 4c and 4d and show that the agent now incurs more risk when it is falling behind. For both λ functions in Figure 4, as the maximum investment is scaled from 100 to 10, the relative value of the bet as scaled against current wealth changes and the maximum investment is quickly reached with even a small probabilistic advantage. In all cases when the agent has a 50% probability or less to win it will play the minimum investment to minimize expected loss.

We show the performance of the λ_1 RAU model and the effect of varying the standard deviation of the sampled probability used by Nature in Figure 5a. The simulations were executed using $W_I = \{100, 200, 400, 800, 1600\}$ and $I_M = 100$. The RAU agent is able to quickly take advantage of small variations in the probability and improves as the margin of the standard deviation increases. Furthermore, as the starting wealth for both agents is increased, the RAU model is able to better take advantage of the opportunities for exploitation and safeguard against low probability outcomes. Conversely, when the starting wealth is lower, a string of bad luck on investments with a 55% probability of winning can result in running out of wealth early in the simulation. This is especially true when the maximum investment is equal to current wealth and there is the potential to lose the game immediately, even with odds in the agent's favor.

The average time per simulation, which ends when one agent has 0 wealth, is shown in Figure 5b. In all cases the time per simulation decreases as the standard deviation increases. This is due to the increased opportunities presented by probabilities above 50%. Our algorithm will take greater risk as the probability of winning increases, ending the simulation faster. Additionally, as the initial wealth increases simulations will inevitably take longer to exhaust the wealth of the losing agent.

We show similar results for λ_2 model seen in Figures 5c and 5d. Unlike the λ_1 model this approach plays more aggressively when it is losing by a significant



(a) λ_1, $I_M = 100$, $W_I = 100$

(b) λ_1, $I_M = 10$, $W_I = 100$

(c) λ_2, $I_M = 100$, $W_I = 100$

(d) λ_2, $I_M = 10$, $W_I = 100$

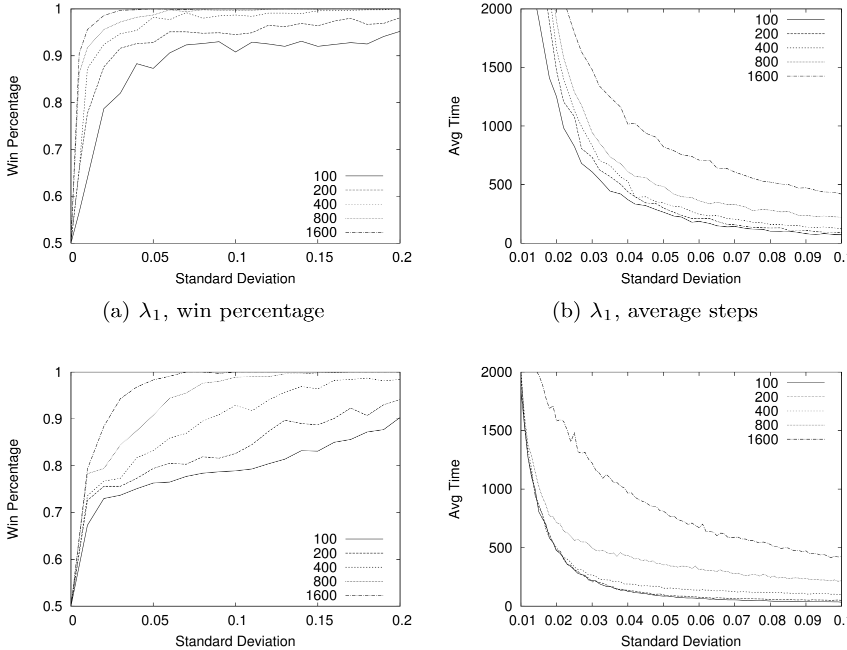
Figure 4. Investment (z-axis) based on the adjusted utility of an agent using RAU with the λ_1 and λ_2 functions. Relative wealth is a proportion of the RAU model agent wealth (W_{rau}) to the opponents wealth (W_{opp}) and is computed as W_{opp}/W_{rau} with 1 being equal wealth at the center of the axis. The λ_1 function prescribes being risk averse when the agent is losing to the opponent, while the λ_2 function is less risk averse when either player has an advantage.

margin. This is motivated by the Prospect Theory work of Kahneman and Tversky [4] where they showed that individuals typically take more risk when below a reference point. The reference point in our work is the point at which both agents' wealth is equal. This type of play results in a lower percentage of wins but also finishes games in less time while still maintaining an advantage above the reference point. This trade-off between performance and time can be utilized in situation where time may incur an additional cost that exceeds the benefit of winning more often. Ideally, a complete λ function would incorporate this cost directly.

6. Discussion

We have presented a method for dynamically adjusting an agent's strategy by adjusting risk attitude in repeated game play. This is achieved by using the agent's current wealth, the potential gain and loss, and the probability of gain to compute the expected utility of each action. The resulting utility function displays the desired non-satiation and risk-aversion properties as defined in utility theory.

We have introduced a 2-player, zero-sum, investment game that embodies a competitive environment in which agents must choose the level of risk they are



(a) λ_1, win percentage

(b) λ_1, average steps

(c) λ_2, win percentage

(d) λ_2, average steps

Figure 5. Win percentage and average time steps for both λ_1 and λ_2 functions.

willing to incur. This risk level is modeled as an investment that is used to scale the gain and loss that each agent is willing to accept in order to maximize expected utility. The λ function is used to adjust the risk-aversion of the agent with respect to changes in the environment. We use repeated play of the investment game as an environment for competition with wealth being maintained between stages of the repeated game.

Finally, we presented two versions of the λ function that are intended to exhibit different strategies for play in the repeated investment game. By using the relative accumulated wealth between two players we provide agents that take actions geared towards meeting specific objectives for the repeated game. Experimentation and results using this model have been shown to be successful for this repeated investment game, but there are many possible extensions of this work.

6.1. *Future Work*

In our simulation the probability of winning was provided to the agent. In most scenarios this is infeasible. We would like to investigate ways which an agent can estimate the probability of success. In situations where an agent can fully observe the outcome of a previous state-action transition, we could use inferencing or machine learning techniques. Inferencing could be driven by knowledge of the environment or via Bayesian methods [19] based on past perception. This approach would likely motivate very cautious play until a sufficient level of confidence is reached regarding the probability of success in a given state.

Because the level of investment is controlled by both agents there is potential for opponent modeling techniques, such as those found in [17], that will allow for taking advantage of states in which the other agent is estimating the probability or managing risk incorrectly. An opponent modeling technique would be even more effective in situations where the game is being played against multiple opponents with different behavior and/or belief systems.

References

[1] Y. Shoham and K. Leyton-Brown. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge Univ Pr, 2009.

[2] H. Markowitz. The utility of wealth. *The Journal of Political Economy*, 60(2):151–158, 1952.

[3] R.B. Myerson. *Game theory: analysis of conflict*. Harvard Univ Pr, 1997.

[4] D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, pages 263–291, 1979.

[5] D. Bernoulli. Exposition of a new theory on the measurement of risk. *Econometrica: Journal of the Econometric Society*, pages 23–36, 1954.

[6] Thorsten Polleit. What can the law of diminishing marginal utility teach us? http://mises.org/daily/5014/What-Can-the-Law-of-Diminishing-Marginal-Utility-Teach-Us, February 2011.

[7] J. Dutka. On the st. petersburg paradox. *Archive for History of Exact Sciences*, 39(1):13–39, 1988.

[8] M. Friedman and L.J. Savage. The utility analysis of choices involving risk. *The Journal of Political Economy*, 56(4):279–304, 1948.

[9] J.W. Pratt. Risk aversion in the small and in the large. *Econometrica: Journal of the Econometric Society*, pages 122–136, 1964.

[10] C.A. Holt and S.K. Laury. Risk aversion and incentive effects. *American economic review*, 92(5):1644–1655, 2002.

[11] T.J. Lybbert and C.B. Barrett. Risk-taking behavior in the presence of nonconvex asset dynamics. *Economic Inquiry*, 49(4):982–988, 2011.

[12] J.H. Knott, J. Verkuilen, R. Pahre, B. Peyton, and T. Hammond. A Computer Tournament of Adaptive Decision Making: Two-Person Iterated Games of Limited Information. 2003.

[13] G. Rezaei and M. Kirley. Heterogeneous Payoffs and Social Diversity in the Spatial Prisoner's Dilemma game. *Simulated Evolution and Learning*, pages 585–594, 2008.

[14] M. Tomochi and M. Kono. Spatial prisoner's dilemma games with dynamic payoff matrices. *Physical Review E*, 65(2):026112, 2002.

[15] M. Perc and A. Szolnoki. Social diversity and promotion of cooperation in the spatial prisoner's dilemma game. *Physical Review E*, 77(1):011904, 2008.

[16] R. Joosten, H. Peters, and F. Thuijsman. 10. Games with Changing Payoffs. *The Economics of growth and technical change: technologies, nations, agents*, page 244, 1994.

[17] S. Ganzfried and T. Sandholm. Game theory-based opponent modeling in large imperfect-information games. In *Proceedings of the Tenth International Conference on Autonomous Agents and Mutli-Agent Systems (AAMAS)*, 2011.

[18] J. Pearl. Fusion, propagation, and structuring in belief networks. *Artificial intelligence*, 29(3):241–288, 1986.

[19] G.E.P. Box, G.C. Tiao, and WISCONSIN UNIV MADISON DEPT OF STATISTICS. *Bayesian inference in statistical analysis*. Wiley Online Library, 1992.