Subjective Expectations in Learning Classifier Systems

Part 1: Risk Neutrality in The Bucket Brigade

Justin T.H. Smith
University of New Mexico
Department of Economics
May, 2008

Abstract

Both economics and biology have come to agree that successful behavior in a stochastic environment responds to the variance of potential outcomes. Unfortunately, when biological and economic paradigms are mated together in a learning classifier system (LCS), decision-making agents called classifiers are generally indifferent to risk. LCS are an adaptive rule-based production system for machine learning relying on both an evolutionary algorithm, the Genetic Algorithm (GA), to shape the pool of behavioral rules over the long run and an economic algorithm, the Bucket Brigade (BuB), that allocates credit to individual rules in the short run. Since a fundamental problem of learning in a stochastic environment is risk management, LCS have not always performed as well as theoretically predicted. In this three-part series, I develop the first comprehensive theoretical and mathematical framework for understanding learning in LCS through the economic and reproductive behavior of classifiers, then rebuild LCS from the bottom-up with risk-sensitivity as a fundamental axis of decision-making. The unifying principle is simple: **fitness is the relationship between economic behavior and reproductive behavior**.

Part 1 develops the basic short-run model of risk-neutral behavior in a traditional BuB market. Before changes to LCS design can be credibly proposed, I first demonstrate the applicability of the basic model to the typical design and reexamine two basic issues where traditional LCS performance fails to meet expectations: 1) default hierarchies, where a general classifier that activates in multiple states is prevented from making mistakes in certain states by another classifier possessing more specific information, and 2) long chains of coupled classifiers, where gains are distributed down a chain of BuB transactions to reward rules that made the payoff at the top possible through previous action. Risk neutrality creates dynamic instability in both areas, while homogeneity of risk preferences results in extreme attenuation of changing price signals down classifier chains. Despite these limitations and the simplicity of even risk-neutral classifiers, long-term learning in is LCS more flexible than Bayesian updating while capable of emulating such behavior when adaptive.

1. Introduction

Economists have been pondering the problem of risky decision-making for centuries. Daniel Bernoulli's (1738) hypotheses that 1) individuals maximize not the expected value of their wealth but the personal "happiness" or "utility" received from it and 2) individuals are risk averse, weighing losses more than equivalent-sized gains, have been at the heart of the economic approach since von Neumann and Morgenstern's (1944) original axiomatic derivation and proof. The role of risk in evolutionary and ecological systems wasn't explored in depth until more recent decades, primarily in the study of foraging and predation behavior in animals. The general result from this work is that not only is stochasticity an undeniable feature of biological environments, but also that risk-sensitive behavior is often adaptive and favored by natural selection. Both economics and biology have come to agree that successful behavior responds to the variance of potential outcomes. Unfortunately, when biological and economic paradigms are mated together in a learning classifier system (LCS), agents making decisions are generally indifferent to risk. LCS are an adaptive rule-based production system for machine learning, developed by John H. Holland in the 1970's (Holland and Reitmanm 1978), that rely on both an evolutionary algorithm (the *Genetic Algorithm*, or GA) to shape the pool of behavioral rules available over the long run and an economic algorithm (the *Bucket Brigade*, or BuB) that allocates credit to individual rules in the short run[1]. Rules are represented as independent agents called classifiers, with more economically successful classifiers more likely to be selected by the GA to reproduce. Reproduction is often imperfect, such that new variants of successful classifiers are introduced and, over time, individual lineages and the entire system as a whole evolve and learn in response to reinforcement from the environment.

LCS have not always performed as well as theoretically predicted. Specifically, LCS suffer from severe weaknesses in their ability to develop and support local assemblies of classifiers that cooperatively solve problems, most notably 1) default hierarchies, where a general rule that activates in multiple states is prevented from making mistakes in certain states by another classifier possessing more specific information, and 2) long chains of coupled classifiers, where gains are distributed down a chain of BuB transactions to reward rules that made the payoff at the top possible through previous action. The literature devoted to these problems and the associated battery of tested modifications is extensive. A full review is outside the scope of this paper, but see Wilson and Goldberg (1989) for starters. A solid understanding of these problems has been difficult as LCS are not well-defined mathematically. The goal of this paper is to develop a mathematical model of classifier behavior focused initially on the Bucket Brigade market that can shed new light on the default hierarchy and coupled chain problems. This will only provide some of the basic framework for a treatment on the potential of risk-sensitivity in LCS. Uniting the economic and Genetic Algorithms of LCS under the same theoretic and mathematic framework will require risk-sensitive behavior on the part of individual classifiers. Since risk-sensitivity will have dramatic consequences for both the economic and reproductive behavior of classifiers, requiring a fundamental reinterpretation of the GA and demanding new market mechanisms to realize the full potential of such behavior, the development of a unifying model will follow separately. Before changes to LCS design can be credibly proposed, I will first demonstrate the applicability of the basic model to the typical design. Parts 2 and 3 will follow in additional installments.

2. Subjective Expectations

The model takes the perspective that a classifier's behavior is not generated by objective knowledge of the true state of the world, but by the equivalent of an internal model of the individual agent's local environment, shaped by its personal experiences. The foundational model for this approach is the subjective expected utility (SEU) model of

[1] This paper focuses on Holland- or Michigan-style LCS where the GA acts on individual rules, as opposed to Pittsburg-style LCS where the GA acts on sets of rules.

decision-making first developed by Savage in his revolutionary 1954 *The Foundations of Statistics*. I will first provide a brief overview of the SEU model's development and operation, and then adapt the theory and methodology to LCS.

2.1 Utility

A classifier's economic behavior consists of competitively bidding for input messages in auctions with other classifiers. Winning the bid competition and "investing" in an input message with a share of its accumulated wealth allows a classifier to then output its own message for sale via auction. The environment itself acts as a classifier by selling messages that characterize the external state and by purchasing certain output messages, serving as the ultimate source of value for the system by rewarding a set of desired behaviors. Thus the value a classifier places on an input message somehow reflects the value it expects from its own output message. The goal is to understand and describe how classifiers make such valuations in a stochastic environment.

The basic idea comes from Daniel Bernoulli's (1738) solution to the famous St. Petersburg posed by his cousin Nicholas Bernoulli in 1713, which demonstrated that people do not maximize the expected value of any real currency or wealth (w). Instead, Daniel focused on *utility*, a sort of "satisfaction" or "happiness" from wealth which is not necessarily linearly related to wealth but may vary nonlinearly in a manner that captures an individual's preferences for risk. Instead of maximizing the expected value of wealth, individuals make choices to maximize the expected value of their utility. Bernoulli's expected utility model became the dominant paradigm in economics when given a mathematical basis by von Neumann and Morgenstern's (1944) proof that certain properties of the basic structure of an agent's preferences imply and are implied by behavior that maximizes some continuous function that exhibits the *expected utility property*. Consider a stochastic event as a lottery L with n prizes, $[y_1, y_2, ..., y_n]$ received with probabilities $[p(y_1), p(y_2), ..., p(y_n)]$. The expected utility property says that the utility of such a lottery is simply the expectation of the utility from its individual prizes, such that utility is additively separable over outcomes and linear in the probabilities:

(1) $$EU(L) = \sum_{i=1}^{n} p(y_i) v(y_i), \ \sum_{i=1}^{n} p(y_i) = 1,$$

where $v(y_i)$ is the utility of each outcome level of income. Expected value maximization is a special case of expected utility maximization where the utility function is linear, for example: $EV(L) = \sum_{i=1}^{n} p_i y_i$, for convenience letting $p_i = p(y_i)$.

The von Neumann and Morgenstern theorem relies on full objective knowledge of both the likelihood and value of risky outcomes. Savage took this one step further by demonstrating a consistent theory of choice whereby these expectations are formed from an agent's "personal" probabilities as "degrees of belief." An individual's subjective probabilities for lottery L are given by $[\pi_1, \pi_2, ..., \pi_n]$, such that the subjective expected utility is:

(2) $$SEU(L) = \sum_{i=1}^{n} \pi_i v(y_i), \ \sum_{i=1}^{n} \pi_i = 1.$$

Given a choice between lotteries L and K, a subjective expected utility maximizer will choose whichever it *believes* has the greater expected utility based on its subjective probabilities for the outcomes of each lottery. How are these subjective probabilities related to any underlying objective probabilities? This subjectivist approach can be traced back to Thomas Bayes (1763), and these subjective probabilities as used by Savage are Bayesian probabilities, subject to Bayes' law. Learning in the SEU model thus occurs via Bayesian updating such that $\pi \to p$, as information causes subjective probabilities to adjust to match underlying objective probabilities.

This subjective expected utility function fully captures an agent's preferences if certain axioms are satisfied[2]. As a behavior model, the focus is thus on *revealed* preferences through observed choices. Even when systematic biases that conflict with preference axioms are observed in individual choice behavior, such as ambiguity aversion as seen in Ellsberg's paradox (1961), Savage's approach is still often seen as a normative model for rational choice—how rational individuals *ought* to behave. We have it a bit easier with LCS because classifier behavior is driven by a well-behaved bidding function[3] satisfying typical axioms such as transitivity and completeness. We needn't worry about reconstructing preferences from choice behavior since this underlying process of introspection is open to the system's designers. The SEU model by itself, however, does not provide an adequate descriptive model of classifier behavior.

2.2 Fitness

While Savage extended the expected utility model with subjective probabilities, his model and most of its derivatives still assume that the values of outcomes in a lottery are fully known, learning a matter only of measuring likelihood. Classifiers, however, are tasked with learning not only how likely their output messages are to sell, but also the price they can expect to receive if and when they do find a customer. This will require one small step past Savage, considering simultaneously subjective beliefs about both the size of outcomes and their probabilities. Interaction between multiple beliefs makes for flexibility, supporting behaviors outside the predictions of traditional expected utility theory. Fortunately, this is an easy extension to implement, though the implications are not as straightforward. As will be shown below, this allows learning that is fundamentally *non*-Bayesian. Non-Bayesian learning is important because a classifier's beliefs don't only affect its bidding behavior in the BuB, but also its *reproductive behavior*. In the presence of the GA, reproductive success is, ultimately, the whole point of risking and earning wealth in the BuB. Behaviors persist and spread when they are adaptive, not necessarily when they are based on accurate beliefs. Once you consider a classifier's strength as a currency that can be spent on reproduction, the label of "utility" doesn't do justice to the trade-offs being optimized. What we're talking about is *fitness* in the implicit sense, as expressed through behavior. This is the ultimate benefit of the axiomatic/mechanistic approach—a set of simple conditions under which we can fully represent behavior that is implicitly fitness-maximizing with an explicit fitness function, suitable fodder for a Genetic Algorithm. For this initial paper, I'll assume a GA consistent with Grefenstette (1991): the growth rate (expected number of offspring) for classifier x is given by a composition of three functions:

(3) $$gr(x) \equiv select\left(u\left(f\left(x\right)\right)\right) \equiv select\left(v\left(w_x\right)\right).$$

Working outward, the objective function $f(x)$ is simply the classifier's wealth, w. Next, $u(w)$ is the fitness function, which, defined over wealth, represents implicit fitness such that $v(w_x) = u(f(x))$ just as in a von Neumann-Morgenstern utility function, thus assumed to be strictly monotonic. Last is the selection algorithm, here also strictly monotonic by assumption such that the GA is *strict* by Grefenstette's definition. In a following paper, Part 3, I will make stronger claims on the selection algorithm, extending the role of subjective expectations, but for now it's enough that classifiers can expect to produce more offspring than those with a lower implicit fitness, and vice versa.

Both interacting subjective beliefs and the pressure of selection separately imply that, unlike utility functions, implicit fitness functions are unique, with even a linear transformation bringing behavioral changes, as will be seen below with the scale constant k in the bid function. Together, these features of LCS and similar systems lead to a model that is more than just a simple adaptation of SEU theory, so I will refer to it as the *subjective implicit fitness* model. While SEU is a theory of consistent choice under uncertainty, SIF is a model of adaptive choice. Fitness is implicitly derived from first

---

[2] Following Savage, a number of others developed alternative axiomizations and variations on the subjective expected utility model with different axiomatic foundations. For an early, but thorough review, see Fishburn (1981).

[3] Classifier systems such as Goldberg's (1989, 1990) which inject random noise into bids effectively obscure the expressed preferences of classifiers, so that we might observe "preference reversals" where lottery L is preferred to K in one case and K to L in another even if beliefs are invariant. This kind of "irrationality" is intended to make auctions non-deterministic, but since the interest of this paper is the deterministic behavior of classifiers it will be assumed that bids are free from such noise.

principles, and subjective beliefs and preferences drive behavior through subjective expectations. In the next section, I develop the model up through the implicit fitness function and show how subjective preferences and beliefs interact through fitness-maximizing bidding behavior.

3. Classifiers

3.1 Subjective Beliefs and Preferences

A classifier's preferences, expressed through the function $v(w)$, indicate the subjective tradeoff between risky rewards from producing outputs and the certain losses due to purchasing inputs. Preferences are important even in the case of risk neutrality because they implicitly determine how much a classifier values the boundaries of its own knowledge—the *tolerance* for risk and uncertainty. The willingness, or unwillingness, to risk all of one's wealth even in the face of a favorable gamble depends on preferences. This implicit savings behavior has a large impact on fitness, as will be seen when we get into the particulars of how these preferences and beliefs are formed and expressed in classifiers.

The environment as subjectively experienced by an individual classifier is at best a crude approximation of reality, captured in just a few key variables. Each economic decision a classifier makes plays out across two transactions: purchasing an input and selling an output, modeled as two sequential, independent[4] lotteries. In the first lottery, the bid competition (L^B), the prize for making the winning bid (B) on an input is essentially a lottery ticket for the latter sale auction (L^S) with a chance to win an actual "monetary" prize. A traditional classifier has what I'll call *simple beliefs* about these lotteries. In L^B, a simple classifier is one that has no means of predicting the bids of its competitors; more on the implications of this in the next section. In L^S, a simple classifier's beliefs consist of just two components: a belief about the price received for an output if sold (R) and a belief about the likelihood of making the sale (π^+), such that the predicted average reward is simply $\pi^+ \cdot R$. More generally, beliefs about selling outputs could be modeled and constructed with a continuous probability distribution over the set of all nonnegative real numbers, more apt to accurately describe the return in a stochastic environment, but the simple form has its benefits and adequately describes most classifiers in the literature.

The compound of sequential lotteries can be reduced into an equivalent simple lottery[5] that in this case adds the state of losing the bid competition L^B to the number of outcomes considered in the sale auction L^S. Thus a simpler classifier's subjective market environment can be described in just three dimensions: the no-output state (w^0) indicating no production, and the states where its output either sells in the next time step (w^+) or fails to sell (w^-). If we denote the level of wealth (strength) accumulated by a classifier at step index τ by w_τ, assuming for now no taxes or other exogenous adjustments to the classifier's strength between time steps, the expected amount of strength achieved in each state at the next step index $\tau+1$ is given by:

(4) $$w^0 = w_\tau,\ w^+ = w_\tau - B + R, \text{ and } w^- = w_\tau - B,$$

and the likelihood of these different states occurring can be described with some probabilities, in this case q^0, q^+, and q^-, respectively, such that $q^+ = (1-q^0)\pi^+$ and $q^- = (1-q^0)(1-\pi^+)$. The expected fitness of this three-state lottery is:

(5) $$SIF(w^0, w^+, w^-) = q^0 v(w^0) + q^+ v(w^+) + q^- v(w^-),\ q^0 + q^+ + q^- = 1.$$

[4] The probability a classifier has of selling its output and the reward received upon completing a sale are assumed to be independent of the price paid for an input.

[5] This Reduction of Compound Lotteries is an explicit axiom or derivative in most expected utility models, but doesn't always hold up in decision-makers as complex as human-beings (Budescu and Fischer, 2001).

3.2 Fitness for Auction

Even the three-state formulation with simple beliefs is a bit cumbersome, but the problem can be slimmed down further if classifiers' bids reveal their full demand for the sale auction. Usually, the structure of the auction determines whether the dominant strategy is to bid one's full demand. In most traditional LCS, the winner in each bid competition is chosen probabilistically, the chance of purchasing an input proportional to the relative size of each classifier's bid, $\partial q^0 / \partial B \leq 0$. Such probabilistic auctions don't fit into the Vickrey's (1961) classification scheme that economists use to describe real-world auctions, where the highest bidder prevails; real markets are more concerned with *efficiency* while machine learning seems to demand experimentation[6]. This trade-off in auction design will be discussed in detail in Part 2 to follow, but for now it's enough to observe that classifiers may not face structural incentives to reveal their full demands. However, simple classifiers lack the machinery to estimate the q probabilities of the bid competition, so the risk of not making the winning bid is best managed by doing the best one can independent of the odds of being out-competed. Auctions are implicitly demand-revealing when bidders are too myopic to anticipate their competition[7]. For traditional risk-neutral classifiers, this implies the familiar steady-state bid equal to the expected average return. More generally, a demand-revealing classifier chooses its bid such that it is indifferent to winning or losing in the bid competition. The demand-revealing bid satisfies the following indifference condition:

(6) $$SIF(w^0) = SIF(L^S),$$

for classifiers with simple beliefs, the optimal bid thus satisfies:

(7) $$v\left(w^0\right) = \pi^+ v\left(w^+\right) + \pi^- v\left(w^-\right), \ \pi^+ + \pi^- = 1,$$

where π^+ is again the classifier's subjective probability of selling an output conditional on it having won the bid competition. Substituting this indifference condition into Equation (5) and simplifying, an equivalent two-state problem is reached that allows the maximization of subjective implicit fitness independent of the odds in the bid competition:

(8) $$SIF(w^0, w^+, w^-) = \pi^+ v\left(w^+\right) + \pi^- v\left(w^-\right).$$

The two-dimensional subjective fitness maximization problem in Equation (8) can be illustrated as in Figure 1, with the sell outcome along the horizontal axis and the state of no sale along the vertical. The key geographical feature of this environment is the $45°$ certainly line, along which every outcome has the same value whether or not the classifier's output would sell ($w^+ = w^-$). The classifier begins at such a point of certainty, with a given amount of wealth w_t, the assumption of no taxes to penalize non-bidding allows the convenient representation of both w_τ and w^0 by a single point along this certainly line. Beliefs about the sale price are represented as running parallel to the certainty line in a 'reward set,' $w^+ = w^- + R$ shown in orange, containing all possible outcomes believed to be in w^+ / w^- space with the expected reward R. For a classifier with more complex beliefs about the size of the reward, the reward set would be a probability cloud, the pdf stretched along the right side of the certainty line. Preferences can be represented given the indifference condition in Equation (7) by an indifference curve connecting the point w^0 on the certainty line with all points that are judged to be equally fit. The classifier's preferences determine the shape of the indifference curve via the functional form of $v\left(w\right)$, such that a risk-neutral implicit fitness function yields a linear form while risk-averse and risk-loving preferences are respectively convex and concave in w^+ / w^- space. The subjectively optimal, fitness-maximizing bid occurs at the intersection of preferences and beliefs.

---

[6] See De Groot (1970) and Goldberg (1989). In probabilistic auctions, the average return a classifier receives selling to its peers is lower than would be received if it could connect directly to the highest bidder. Classifiers pay for the chance to experiment.

[7] More complex classifiers able to monitor and attempt to predict the bidding behavior of competitors may not make for smarter bidders. As Vickrey showed, demand-revealing behavior can be the optimal strategy even when bidders can fully observe the bids of rivals, as in the English or progressive "open" auctions, so there's little justification for additional complexity here.

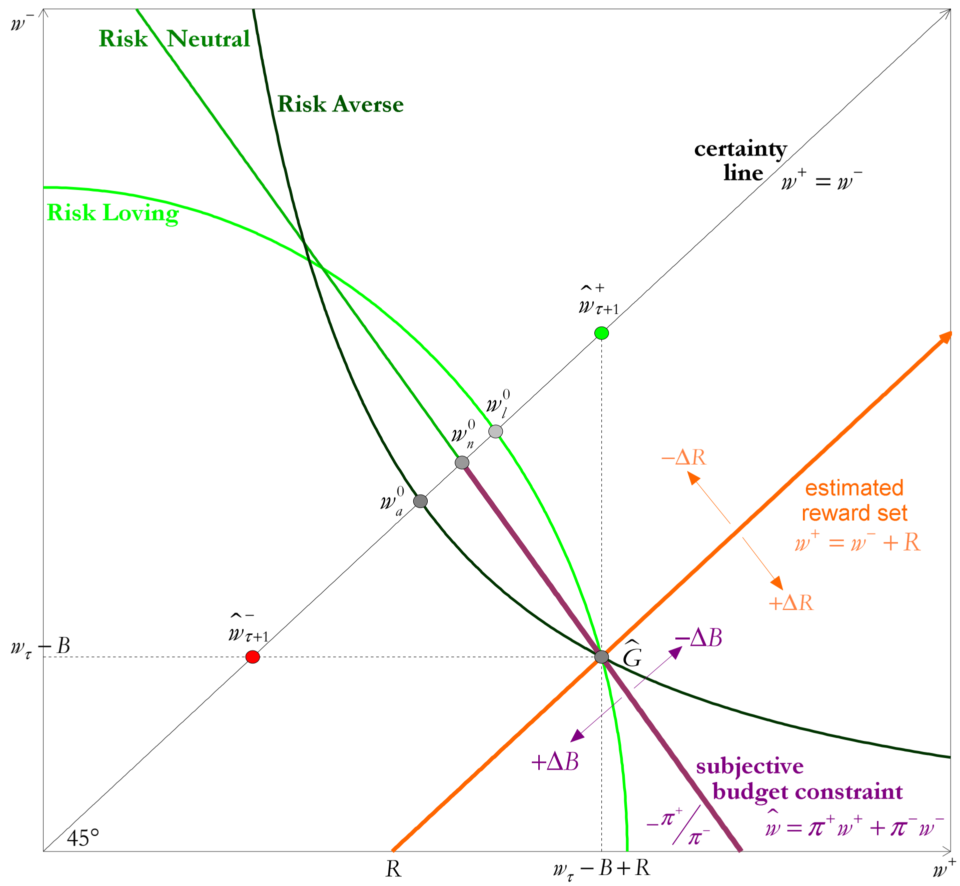

Figure 1 –Preferences (green) and beliefs (orange and purple) come together to subjectively determine a classifier's optimal bid. Where preferences and beliefs overlap, preferences are shown above the certainty line, beliefs below.

The classifier's wealth and the size of its bid determine where in the reward set the optimal subjective gamble $\hat{G} = \left(w_\tau - B + R,\ w_\tau - B\right)$ will be located. By winning the bid competition lottery, the classifier trades its certain position w^0 for a risky but, given beliefs and preferences, subjectively equally fit position $\hat{G}$. Passing through this chosen gamble is a purple line indicating the 'subjective budget constraint[8]' which contains all points in $w^{+}\ /\ w^{-}$ space believed to have an expected value of wealth as the sale auction: $\hat{w} = E\left(w^{+},\ w^{-}\right) = \pi^{+}w^{+} + \pi^{-}w^{-}$. This budget constraint intersects the certainty line where $\hat{w} = w^{+} = w^{-}$, it's position relative to w^0 indicating the average profit or loss expected while its slope reveals the subjective odds of payoff, $-\pi^{+}/\pi^{-}$. Notice how, *ceteris paribus*, changes in the perceived size of the simple reward shift the position of the reward set without changing its slope from $45°$, while changes in the classifier's bid shift the budget constraint while also preserving its slope. Traditional risk-neutral classifiers, with a linear bid function, have preferences that overlap this subjective budget constraint, so choose a bid such that $w_n^0 = \hat{w}$, the average "steady-state" condition. Risk-loving classifiers are willing to loose wealth on average ($w_l^0 > \hat{w}$), while risk-averse classifiers expect to be compensated for the risk of outputs not selling by only engaging in favorable gambles ($w_a^0 < \hat{w}$). Risk preferences present new opportunities but also new problems, so the remainder of this first paper will focus on risk-neutral classifiers with simple beliefs.

---

[8] 'Constraint' is a term inherited from the SEU and other economic models, but here the budget really isn't constrained in the traditional sense, dependent on the classifier's choice of bid.

3.3 Optimal Risk-Neutral Behavior in the Bucket Brigade

Traditional risk-neutral classifiers with linear bid functions offer an easy starting point for analysis, since their preferences coincide with their beliefs. If the variance in returns is irrelevant, only the subjective mean of the reward matters; for example, the risk neutral classifier would bid as much for a 1% chance for a reward of 1000 units as it would for a 100% chance of a 10 unit reward, the expected value the same in each case. Not only are the risk-neutral classifier's preferences revealed in its bidding behavior, so too are its beliefs. Beginning with the indifference condition in Equation (7) and substituting in the values of the outcome states from Equation (4) and a linear functional form, $w = \pi^+ \cdot (w - B + \mathrm{R}) + \pi^- \cdot (w - B)$, which simplifies to:

(9) $$B(\pi^+, \mathrm{R}) = \pi^+ \cdot \mathrm{R},$$

The risk-neutral classifier only cares about the weighted average return and believes its return to be normally distributed. Variance is irrelevant as the magnitude of losses is valued equally to that of equivalent-sized gains. This can be verified from Figure 1, where the slope of the subjective budget constraint is determined by $\pi^+ / \pi^- = \frac{B}{B - \mathrm{R}}$ such that $\pi^+ = \frac{B}{\mathrm{R}}$.

This theoretically predicted bidding behavior can be tied to the bidding function used by simple classifiers:

(10) $$B(\sigma, w) = \sigma \cdot k \cdot w,$$

where the bid is determined by the product of the specificity of the classifier's match to the input tag (σ), the classifier's wealth at the time (w), and a bid constant (k). The domain of these parameters predicts the following decomposition:

(11) $$\pi^+(\sigma) = \sigma \in [0,1] \text{ and } \mathrm{R}(w) = k \cdot w \geq 0.$$

Specificity is measured by the degree to which a classifier's genetically-determined input tag matches the input message up for bid. This translates directly into the classifier's degree of belief that its output will sell given the information about the state of the world carried in the input message. We can also get a little more specific about R, the expected reward from selling an output. While the classifier's bid B reveals its demand for participation in the sale auction, R can be thought of as the classifier's demand for rewards from selling an output independent of the likelihood of finding a buyer. The implicit fitness function is normalized so that rewards are the numéraire:

(12) $$v(w, p_{\mathrm{R}}) = w / p_{\mathrm{R}},$$

where p_{R} is the implicit price in wealth of one unit of expected reward. Using Roy's Identity this demand can be derived directly from the normalized implicit fitness function as a function of wealth and the price of rewards:

(13) $$\mathrm{R} = -\frac{\partial v(w, p_{\mathrm{R}})}{\partial p_{\mathrm{R}}} \Big/ \frac{\partial v(w, p_{\mathrm{R}})}{\partial w}.$$

Taking the partial derivatives of Equation (12) and plugging into Equation (13) results that $\mathrm{R} = w / p_{\mathrm{R}}$, such that p_{R} reflects the classifier's subjective value of its wealth relative to its reward. Let the bid constant $k = 1/p_{\mathrm{R}}$ such that it determines a classifier's *tolerance* for risk, higher values of k inducing a preference for riskier bids farther from the certainty line. This parameter determines the rate at which classifiers convert their wealth into an estimate of the reward for their output, the distance between the certainty line and the estimated reward set. Simple risk-neutral classifiers have the convenient result that $v(w, p_{\mathrm{R}}) = \mathrm{R}$, considering themselves as fit as this expected value of their output. Summing up, the risk-neutral bid function is:

(14) $$B(\sigma, w) = B(\pi^+, \mathrm{R}) = \pi^+ \cdot k \cdot w.$$

4. Inaccurate Beliefs

4.1 Adaptation and Error

The key insight, made first by Arrow (1971) and reinforced in an evolutionary context by Real (1987), is not that subjective beliefs be accurate, but that they be "adaptive." Accuracy-based reward schemes in LCS as sought by Wilson (1995) and others entirely miss this point. The principle can be more generally illustrated in Figure 2, which combines the subjective and objective environments to examine the interplay between subjective preferences and beliefs and an objective reality. Let's assume there's some known, fixed selling price, P, represented by an objective reward set parallel to the certainty line colored red in the figure, and some objective probability of selling an output, p^+ such that the slope of the objective budget constraint, in blue, is $-p^+/(1-p^+)$.

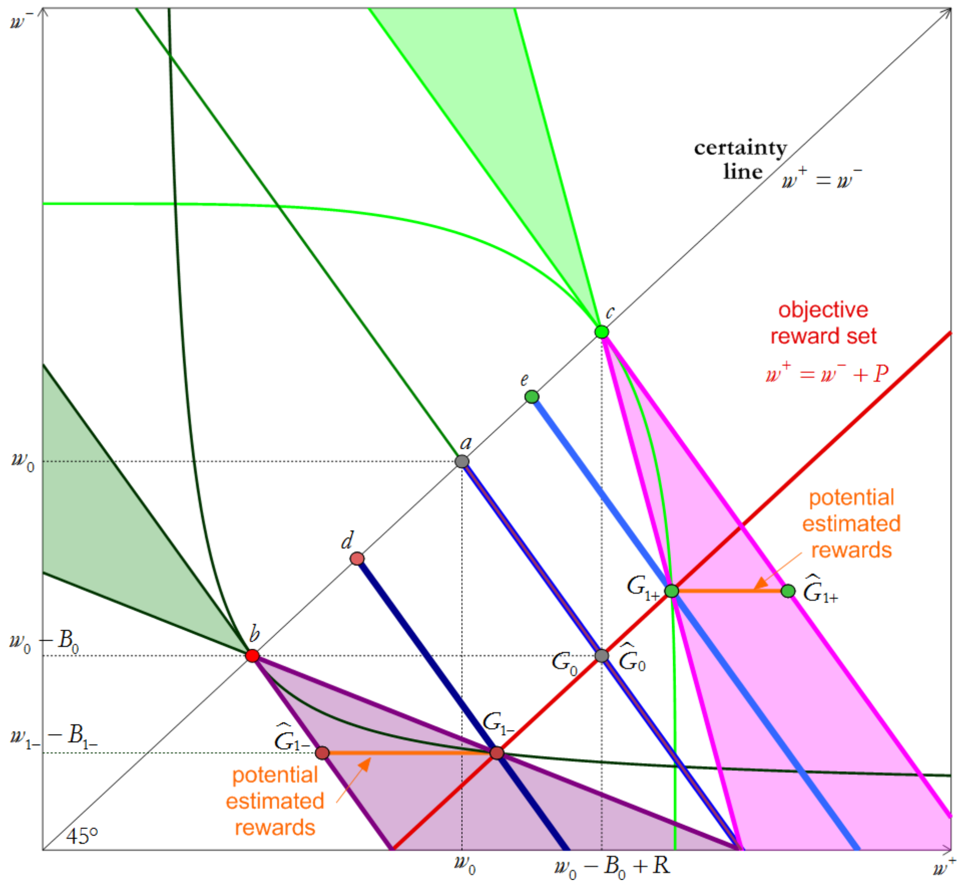

Figure 2 - The interplay of subjective (green, orange, purple) and objective (red, blue) realities.

Suppose a risk-neutral classifier begins at point a with wealth w_0 and completely accurate beliefs, $R = P$ and $\pi^+ = p^+$, such that the optimal objective gamble G_0 agrees with the optimal subjective gamble $\widehat{G}_0$. This is the traditional conception of the "steady state." Assuming for now that the price and objective probability of selling are fixed, what happens the next time the classifier encounters an input message? If the classifier wins in the bid competition, it will have ended up either at point b if it failed to sell its output or point c if it found a buyer. It's clear that, except for the special case where $p^+ = 1$, "steady state" is a misnomer. The picture here is exaggerated for visual clarity, the variance around point a decreasing with lower values of k, but the spread is not distributed normally except for the special case $p^+ = p^- = 1/2$. Not only are risk-neutral classifiers unable to sustain their supposed "steady state" strength in a stochastic environment, the distribution around the mean outcome is skewed from the normal. While classifiers faced with odds $p^+ > p^-$ are most likely to be above average, the median outcome for a low-probability seller is below the average. The frequency and magnitude of over- and under-bids is thus proportional to the relative probabilities of gaining and losing. This has clear implications for fitness and survivability. As this graphical demonstration shows, if a gain is more likely than a loss, the probability distribution around the average strength is skewed with a longer tail on the low end. Below-average outcomes are less likely, but they are spread farther from the mean than the more common outcomes above. On the other hand, riskier environments will lead to subpopulations dominated by a majority of below-average outcomes offset by much rarer high-strength outcomes. These distributional effects of repeated play merely hint at the consequences to classifiers who ignore variance, but even risk-neutral classifiers have some tricks up their sleeves.

Continuing with the graphical example, in the no-sale case, if the classifier had the same perfect knowledge, it would make the same bet this second time, but the penalty for another round of bad luck in this particular example would be death as $w_0 - 2B_0 = 0$. But classifiers are myopic, and the initial bout of bad luck causes the classifier to revise down its belief about the size reward and bid for the gamble $\widehat{G}_{1-}$ at the genetically-constrained specificity π^+, unlike a Bayesian learner who would have revised its subjective probability of payoff to bid for the gamble G_{1-} with a more shallow-sloped subjective budget constraint. The critical insight is that, due to the behavioral flexibility of even the simple belief structure, even though their beliefs differ, both the SIF and Bayesian gamblers make the same bid expecting the same average outcome at point b! The orange line indicates the set of subjective rewards $R \leq P$ that result in the same bid given the appropriate set of beliefs about the probability of payoff in the purple shaded region and appropriate risk-neutral preferences in the green shaded region. But bidding behavior need not be limited to the risk neutral—the same bid would also be made by a classifier with accurate beliefs about the size and probability of the reward who is sufficiently risk averse, given the convex indifference curve, that it demands to be compensated by the average outcome of point d for taking the gamble G_{1-}. After a run of bad luck when a classifier's wealth is below the steady state, inaccurate beliefs in the form of depressed expectations about the price of its output lead the simple risk-neutral classifier to behave as if it were a risk-averse Bayesian learner. Furthermore, this pseudo-risk-averse behavior is *adaptive* in the sense that it leads on average to an increase in wealth and fitness. On the other hand, selling an output moves the classifier's wealth to a point above the steady state, resulting in apparent risk-loving behavior with an average decrease in wealth and fitness that's less-arguably adaptive, but it is *necessary* to maintaining a level of wealth in the neighborhood of the steady state. Risk-neutral classifiers are designed to make mistakes.

This dynamic instability deserves a closer look, as this adjustment process is important in understanding how learning occurs in the BuB. The average error in a risk neutral classifier's bid is:

(15) $$\widehat{\varepsilon_\tau} = p_\tau^+ \cdot \varepsilon_\tau^+ + p_\tau^- \cdot \varepsilon_\tau^-\text{, where:}$$

$\varepsilon_\tau^+ = P_\tau - \pi_\tau^+ \cdot \mathrm{R}_\tau$ is the risk neutral error conditional on selling an output and

$\varepsilon_\tau^- = -\pi_\tau^+ \cdot k \cdot w_\tau$ is the risk neutral error when an output fails to sell.

$$\widehat{\varepsilon_\tau} = p_\tau^+ \cdot P_\tau - \pi_\tau^+ \cdot \mathrm{R}_\tau.$$

The error is the change in a classifier's wealth after paying for an input and attempting to sell an output. When the classifier has fully adjusted to its environment, the error is zero *on average*, but if $p_\tau^+ < 1$, every bid is subject to some nonzero error whether its output sells (ε_τ^+) or not (ε_τ^-). Except in the special case of known, certain rewards, each attempt at selling an output message subjects the classifier to the shock of a return either above or below average, sufficient to push it away from any steady state. It'll be useful to keep phrasing things in terms of averages for a basis of comparison, easily calculating the effect of a positive or negative random error by resolving $p_\tau^+ = 1$ or $p_\tau^+ = 0$, respectively. What's really important is the relative magnitude of shocks, either from random error or real changes in a classifier's environment. The *average error rate* is the average error relative to the classifier's bid:

(16)	$\widehat{e}_\tau = \dfrac{\widehat{\varepsilon}_\tau}{B_\tau} = \dfrac{p_\tau^+}{\pi_\tau^+} \cdot \dfrac{P_\tau}{R_\tau} - 1$.

Error depends on the degree to which subjective beliefs deviate from reality. How does this error affect a classifier? The output-induced *average adjustment rate* in an active risk neutral classifier's wealth is given by:

(17)	$\widehat{\omega}_\tau = \dfrac{\widehat{\dot{w}}_\tau}{w_\tau} = \dfrac{\widehat{\varepsilon}_\tau}{w_\tau} = \left(p_\tau^+ \cdot {P_\tau}/{w_\tau}\right) - \left(\pi_\tau^+ \cdot k\right)$.

Since bid error only causes changes in the classifier's wealth when actually producing an output, we can express the overall average growth rate at a given level of strength by accounting for the frequency with which a classifier can win the bid competition. Assuming no taxes or other exogenous adjustments, the *average growth rate* is:

(18)	$\widehat{\Omega} = \left(1-q^0\right) \cdot \widehat{\omega}_\tau = \left(1-q^0\right)\left[\left(p_\tau^+ \cdot {P_\tau}/{w_\tau}\right) - \left(\pi_\tau^+ \cdot k\right)\right]$.

In any case, positive adjustment depends on objective reality: the true likelihood of selling an output at a given point in time and the relative magnitude of the actual selling price. Positive adjustments are probabilistic. The rate of negative adjustment, however, is a certainty, and *solely* due to subjective parameters: the tolerance for risk and personal beliefs about the likelihood of payoff. A stable steady state exists

(19)	$\omega_\tau = 0$ if and only if $\pi_\tau^+ = p_\tau^+ = 1$.

The stochastic steady state bid $B^* = p^+ \cdot P$ occurs where:

(20)	$\widehat{\omega}_\tau = 0 \Rightarrow \left(\pi_\tau^+ \cdot k\right) = \left(p_\tau^+ \cdot {P_\tau}/{w_\tau}\right)$.

Again, notice that this condition doesn't require that subjective beliefs be accurate, only that they combine with subjective preferences to produce a certain behavior (the steady-state bid). The true problem with random error is easy to see in this formulation—how can the classifier separate random changes in w , one variable in the right-hand side of Equation (20), from important changes in its environment via changes in p^+ and P ? This problem will be examined in more detail below when discussing classifier chains. The stochastic steady-state wealth,

(21)	$w_\tau^* = \dfrac{p_\tau^+}{\pi_\tau^+} \cdot \dfrac{P_\tau}{k}$,

can vary as $\sigma_\tau <> p_\tau^+$. Fitness and thus the expected reward for risk-neutral classifiers bidding at the stochastic steady state depend positively on the on average real payoff but are inversely proportional to subjective beliefs about the likelihood of reward:

(22)	$v\left(w_\tau^*\right) = \dfrac{p_\tau^+}{\pi_\tau^+} \cdot P_\tau = R^*$,

A risk-neutral classifier's steady-state fitness is only equal to the true value of its output if its beliefs are accurate. Steady-state fitness and expected reward are independent of the risk-neutral classifier's tolerance for risk, k .

Learning in the Bucket Brigade boils down to a classifier changing its bid as a response to environmental changes, either in the magnitude or likelihood of the reward. Define the adjustment rate in a classifier's bid as:

(23)	$b_\tau = \dfrac{\dot{B}_\tau}{B_\tau} = \dfrac{B_{\tau+1} - B_\tau}{B_\tau}$.

Risk neutrality is defined by the one to one response to changes in wealth, so with risk neutral classifiers we can use the adjustment rate in wealth as a proxy for learning:

(24)	$b_\tau = \omega_\tau$.

In order to understand how a risk-neutral classifier's strength and thus bid respond to the environment, consider the *elasticity*, the growth rate of one variable induced by a one percent increase in another, as a measure of sensitivity. The elasticity of wealth and bids to variations in rewards, for a risk-neutral classifier, is simply the relative size of its bid:

(25) $$\mathrm{E}_{B_\tau,\varepsilon_\tau} = \mathrm{E}_{w_\tau,\varepsilon_\tau} = \frac{\widehat{\omega_\tau}}{\widehat{e_\tau}} = \frac{\widehat{\varepsilon_\tau}}{w_\tau} \cdot \frac{B_\tau}{\widehat{\varepsilon_\tau}} = \frac{B_\tau}{w_\tau} = \pi_\tau^+ \cdot k .$$

In the elasticity formulation, all stochasticity cancels out. Response happens only when deviations have been experienced, for better or for worse. Furthermore, deviations in actual return from expectations due to random error in a stochastic environment have the same effect on a risk-neutral classifier's strength and bid as real changes in the environment. The most important result is that while a classifier's environment determines the magnitude of deviations from expectations when attempting to sell an output, the degree of response to this information is solely dependent on personal, *subjective* factors. The sensitivity to deviations between reality and the risk-neutral classifier's expectations is directly proportional to its tolerance for risk, k, such that classifiers with a higher tolerance for risk are subject to greater behavioral perturbations due to bid error, i.e. learn "faster" from environmental changes real and imagined. A risk-neutral classifier tempers its response to environmental fluctuations in direct proportion to the subjective belief about the amount of stochasticity in its reward, neutrally compensating for the perceived risk that any shocks to its selling price are random error rather than meaningful information. A risk-neutral classifier discounts its response to shocks the same way it discounts its strength when predicting the price of its output.

As an example, suppose a classifier with $k = 0.2$ faces a stochastic environment with known odds of selling $\pi_\tau^+ = p_\tau^+ = 0.5 \,\forall \tau$ at the price $P_0 = 20$. Suppose that the classifier starts at the steady state, making the steady state bid of $B_0 = B^* = 10$ at the steady-state wealth of $w_0 = w^* = 100$, expecting a mean return of 10. The wealth elasticity of errors is $\mathrm{E}_{w_0,\varepsilon_0} = 0.1$; a one percent deviation in selling price from expectations results in a 0.1% change in wealth. If the classifier wins the bid competition and then sells its output at a price with a 100% upward deviation from the expected average ($\frac{20-10}{10}$), the classifier responds to the positive shock of profiting with a 10% increase in wealth, $w_1 = 110$, and a corresponding 10% increase in its next bid, $B_1 = 11$. If, on the other hand, the output failed to sell, the zero price a 100% downward deviation from the expected mean, a 10% decrease in strength and the next bid would occur. The classifier's initial bid accurately predicted the average return, but once this transaction is completed its next bid will be off by 10% in one direction or the other, relying on mistakes in the other direction to pull the classifier back towards the steady state. What's absurd is that the classifier *expects* the correct price it received $(P_0 = 20) = (\mathrm{R}_0 = k \cdot w_0)$ probabilistically, it just doesn't *care*. The essence of risk neutrality is ignoring variance and only focusing on expected mean outcomes. This dynamic instability is a product of willful ignorance on the part of risk-neutral classifiers.

4.2 Default Hierarchies

Under this perspective, the SIF model offers a new explanation for the role of default hierarchies and can help predict the economic and evolutionary environments in which these symbiotic arrangements can evolve due to self-interested adaptive behavior leading to cooperation between different "species" of classifiers. At the most basic level, default hierarchies arise because of Equation (22)—a classifier can raise its fitness by being pessimistic about the likelihood of being rewarded for an output. More specifically, a classifier can maintain a bid for a given input if it concurrently lowers its subjective probability of payoff ($\pi_2^+ < \pi_1^+$) and holds an inversely biased estimate of the reward ($R_2 > R_1$) such that it can reach the stochastic steady state in Equation (20), where $\pi_2^+ \cdot \mathrm{R}_2^* = \pi_1^+ \cdot \mathrm{R}_1^*$. Though both classifiers make the same steady-state bid, they are treated differently by the Genetic Algorithm since $\mathrm{R}_2^* = v_2^* > v_1^* = \mathrm{R}_1^*$. A classifier can raise its fitness by simultaneously holding pessimistic beliefs about the odds of selling an output and optimistic beliefs about the price it earns when selling, and it rises to this new fitness by pseudo risk-averse bidding behavior. Risk-averse behavior of this sort only has an evolutionary advantage over more accurate beliefs if the gains to wealth from being pessimistic about probabilities aren't entirely lost when the classifier is exposed to other input states. This is exactly what may happen in LCS, where $\pi^+ < p^+$ implies generality, a classifier ignoring potentially useful information with don't care symbols that may lead it to make the same bid for other inputs that don't lead to the same expected reward. The inability to distinguish between states can only persist if another classifier, which does make use of the ignored information, intercedes by outbidding it for the misleading input, "covering" for it by restricting the set of states in which it's exposed to feedback. Thus a default hierarchy is at best an example of commensalism but not mutualism, since it only benefits the general classifier; the specific classifier's behavior is indifferent to the presence of the profiting generalist[9]. Still, successfully exploiting Equation (22) requires the development of complex parallelism and coordination. Unfortunately, the dynamic instability of the "steady state" and the need for risk-neutral classifiers to overbid is a serious problem for the creation and maintenance of default hierarchies as demonstrated empirically by Riolo (1987b) among others, even with the positive bias over-general classifiers receive from the GA.

The generalist's differences with its more specific ancestor aren't limited to an increase in steady-state fitness. Since its steady state wealth is higher, the generalist has to accomplish more growth to reach this target. Via Equation (17), the generalist has a faster average adjustment rate than a specialist at the same level of wealth, but because it bids lower has a smaller chance of winning the bid competition. On average, as in Equation (18), either one may grow faster depending on competition for input messages, implying a tradeoff between the quantity and magnitude of learning experiences. Either one may need more successful sales of output messages to reach its full stochastic steady-state strength and fitness, which has an impact on their comparative performance under the GA. Default hierarchies are one of many areas where this tradeoff between growth and reproduction is an important behavioral consideration.

4.3 Classifier Chains

Complex information processing requires more than just a population of classifiers individually responding to environmental information. Default hierarchies aren't the only means of cooperation and specialization open to learning and evolving classifiers. Classifiers that purchase information from the environment are called *receptors*, while those that sell to the environment indicating the system's response—some outcome state or action to take—are called *effectors*. By forming chains, receptors at one end, effectors on the other, information can be processed and correlated over many steps in a chain of transactions in the BuB market. In the traditional design, information flow down classifier chains is often problematic. One issue is attenuation, market signals in the form of changes in environmental rewards getting weaker as they flow down chains. Attenuation can dramatically affect the rate at which a classifier chain can learn, with classifiers far from the reward needing to accumulate many more BuB transactions to learn an ultimate change in their value imputed by changes in environmental rewards (Wilson, 1986).

In a classifier chain, one classifier's bid is another's reward, allowing us to propagate adjustment rates down a link in the chain:

(26) $e_{j,\tau} = b_{j-1,\tau-1} = \omega_{j-1,\tau-1}$, where decreasing j moves down the chain from effector at $j = 0$.

Notationally, the series of transactions down a chain all occurs within one step index τ, a useful simplification here since it takes a full activation of the entire chain to transmit the effect of an environmental shock down one link in the chain. To measure how deviations in rewards are amplified down a chain, again consider elasticity as a measure of sensitivity of price changes between levels of a chain. Via Equations (25) and (26):

[9] An argument could be made that general classifiers in default hierarchies are actually parasites to the degree that the higher fitness of the generalist crowds out the covering classifier in the GA, potentially reducing the covering classifier's replicative rate.

(27) $$\mathrm{E}_{\varepsilon_{j,\tau},\varepsilon_{j-1,\tau-1}} = \frac{\hat{e}_{j,\tau}}{\hat{e}_{j-1,\tau-1}} = \frac{\hat{\omega}_{j-1,\tau-1}}{\hat{e}_{j-1,\tau-1}} = \pi^{+}_{j-1,\tau} \cdot k_{j-1} .$$

The rate of adjustment going down a level is attenuated by a fraction determined *solely* by the subjective preferences and beliefs of the buying classifier. Attenuation can only be avoided in the special case where a classifier takes upon itself the maximum amount of risk, $\sigma_\tau = k = 1$. In general, $\mathrm{E}_{\varepsilon_{j,\tau},\varepsilon_{j-1,\tau}} \in [0,1)$ and attenuation ensues. Down a chain of multiple links, attenuation is multiplicative, such that the elasticity of shocks to the selling price of a classifier's output at level j induced by a deviation in returns from the external environment at the head of the chain, j activations before, is given by the *attenuation elasticity*:

(28) $$\mathrm{E}_{\varepsilon_{j,\tau},\varepsilon_{0,0}} = \prod_{i=0}^{j-1} \frac{\hat{e}_{i+1,\tau+1}}{\hat{e}_{i,\tau}} = \prod_{i=0}^{j-1} \pi^{+}_{i,\tau} \cdot k_i .$$

Shocks, both price signals to the whole system and random errors, are dampened down classifier chains. Receptors at the bottom of chains are insulated from shocks to effectors at the top and therefore learn less in one transaction from changes in the environment.

What is to be done? Complex mechanisms such as Holland's (1985) "bridging classifiers" jury-rig connections between top and bottom-level classifiers to transfer additional strength from effectors to receptors. The attenuation elasticity provides a more direct handle. Riolo (1987a) demonstrated that, as Equation (28) predicts, attenuation is reduced by raising the tolerance for risk, k_i. Trouble is, in traditional implementations, the tolerance for risk is fixed across the population of classifiers and thus across members of a chain, $k_i = \bar{k} \, \forall i$, who are forced to respond the same way to risk. But the risks faced by effectors at the tops of chains are attenuated down the chain along with the magnitude of shocks. Classifiers at different levels face different risk environments. Receptors at the bottom might be paid by another classifier whether or not an effector at the top of a chain gets paid by the environment, and are insulated from much of the effectors' risks. It makes sense to allow the risk tolerance preference to evolve freely rather than be fixed across all classifiers. If a risk-neutral classifier's risk tolerance k_j was inherited along with subjective beliefs σ_j and similarly subject to mutational changes, elements in a chain could adapt their responsiveness to change to the actual risk in their individual niches. For example, only effectors might have low risk tolerances given their interactions with a stochastic environment, whereas receptors in much more stable environments might find a higher tolerance for risk and sensitivity to changes propagating down from above *adaptive*. The level of attenuation across chains could adapt, balancing increases in responsiveness and learning against risks of staking more on outcomes. And given Equation (22), freeing the risk tolerance preference does not affect the steady state fitness of classifiers, just the growth pattern taken toward the steady state. Equations (25) and (28) aptly demonstrate the important tradeoff between risk and learning, and there's no reason to expect the existence of some globally optimal solution.

While attenuation dampens random bid errors to some extent, potentially ameliorating some of the impact of risk-neutral dynamic instability on a classifier's suppliers, the probabilistic nature of auctions used to resolve bid competition means that the selling price is chosen randomly out of the entire pool of potential buyers, with the result that the price that clears the BuB market is much more variable than the first- or second-price bidder's random bid error. Baum and Durdanovic (2000) report that the inefficiency caused by this violation of property rights can severely harm the system's ability to form long chains of classifiers, another potentially large source of dynamic instability impairing chain formation.

5. Towards Risk Aversion

At this stage, SIF is largely a descriptive model of adaptive learning in LCS in the face of the risk and uncertainty of a stochastic environment. Long-term learning in LCS is more flexible than Bayesian updating, capable of emulating such when such behavior is adaptive. But SIF isn't limited to describing classifier behavior, offering great potential for not only understanding but also improving the learning and performance of LCS. It seems logical for example that the tolerance for risk might be better determined endogenously, adaptively, by relinquishing the fixed risk preferences imposed on all members of the population and adding the bid constant k to the genome and allowing different response strategies to emerge, but this is really a minor alteration. Large-scale assemblies of classifiers such as default hierarchies and classifier chains suffer unduly under the dynamic instability introduced by probabilistic auctions and risk-neutral preferences ignoring the distinction between random error and real fluctuations in the environment. In contrast, risk aversion lends a motive to insure oneself against random noise. For example, classifiers with access to a more complete set of markets would be willing to make the risk-neutral bid if they could trade any stochastic profits, should their return fall above expectations against any potential shortfalls from expected returns, and thus remain stably at the steady state unless the price of this insurance changed. In a following paper, Part 2, I will show how risk-averse classifiers can pool information and trade risk through decentralized insurance cooperatives and betting markets, what I'll call a *Betting Brigade* (BeB). Risk aversion also faces some important challenges. The optimal bid condition given in Equation (7) is not solvable analytically except in the risk-neutral case of a linear implicit fitness function, $v(w)$. Therefore Part 2 will take a closer look at the structure of auctions used in the various markets to solicit bids, also examining the dynamic instability of probabilistic auctions and the trade-off between efficiency and experimentation.

The ultimate potential for risk aversion lies in a radical rethinking of reproduction and evolution in LCS. **Fitness is the relationship between economic behavior and reproductive behavior.** A classifier's fitness represents its subjective valuation of itself, and a classifier can make decisions about reproduction based on the expected market performance of products of its own reproduction. For example, a classifier can easily compare the subjective fitness implicit in having all its strength locked up in one self risking some amount in betting on inputs in w^{+}/w^{-} space versus the subjective implicit fitness of dividing its strength into two daughter classifiers in order to make two independent but smaller bets. In the case of asexual reproduction, where daughter cells are approximately $100\,\%$ related to the parent aside from some small chance of mutation, it's easy to see that a classifier would subjectively gain fitness from dividing its strength w^{0} by some fraction φ if and only if:

(29) $$v\left(\varphi\cdot w^{0}\right)+v\left(\left(1-\varphi\right)\cdot w^{0}\right)>v\left(w^{0}\right),\ \varphi\in\left(0,1\right).$$

Risk-proclivic classifiers would never want to reproduce, preferring the risk of having all strength locked up in one risk-seeking bid at a time. Risk neutral classifiers have a different problem; they are indifferent to reproduction, don't care how the return on their strength is divided up among themselves and any number of descendants. Risk neutral classifiers are content to go along with the subjectively arbitrary reproductive decisions of the GA. In a following paper, Part 3, I will go in a new direction. Risk-averse classifiers subjectively value the ability to hedge against the risks of a stochastic environment by using reproduction to spread their strength across offspring, and can make reproductive decisions for themselves. The axiomization of implicit fitness provides the conditions under which endogenous fitness preferences can be represented by an explicit real-valued fitness function, making efforts to endogenize fitness through discrete resource models such as Booker (2000) and Holland's ECHO (1995) superfluous. Indeed, it's easy to calculate a classifier's willingness to pay to reproduce, but we can go further. SIF provides a framework upon which to develop a "life history" of classifiers and derive from first principles a whole suite of behaviors tying the rate of growth and learning to questions about the optimal size and level of parental investment in offspring, the role of sex, and other issues in the context of subjective implicit fitness maximization. In Part 3, I will make the genetic operations provided by the GA subservient to the reproductive decisions of classifiers as expressed in a market for reproductive access, what I'll call a *Baby Brigade* (BaB).

LCS have much untapped potential as vehicles for studying learning in and by markets. Hooking a market up to some interface to directly test its ability to adapt through the distributed learning and discovery of individual agents is an approach worth furthering. In addition, LCS may be most provocative as a model of cognition—diverse 'subconscious' elements processing information and competing for access to the 'conscious'/'effective' mind. SIF offers a testable model of a mind composed of discrete elements shaped by both evolutionary adaptation and dynamic learning. Holland and Miller (1991) have advocated LCS as an adaptive agent suitable for use in economic modeling. Subjecting a 'learned' LCS to some of same preference tests devised by economists and psychologists such as Khaneman and Tversky (1979) might reveal if and when LCS are subject to some of the same biases found in people and animals and what sorts of environments make these biases adaptive, or maladaptive. Even if individual classifiers are fully 'well-behaved' in the classical sense, the entire system may not be. It's an open question as to whether individual rationality or lack thereof reflects simple biases at lower levels or whether irrationality can emerge even when information is transmitted and processed rationally. For now, Learning Classifier Systems need a lot of work to become both plausible and functional stand-ins for human beings in computational models.

6. References

Arrow, Kenneth J. (1971). *Essays in the Theory of Risk Bearing*. North-Holland.

Bayes, Thomas. (1763). An Essay Toward Solving a Problem in the Doctrine of Chances. *Philosophical Transactions of the Royal Society*, 53, p. 370-418.

Baum, Eric B. and Igor Durdanovic (2000). Evolution of Cooperative Problem Solving in an Artificial Economy. *Neural Computation*, 12, p. 2743-2775.

Bernoulli, Danuel. (1738). Exposition of a New Theory on the Measurement of Risk. Translated in 1954 in *Econometrica*, 22:1, p. 23-36.

Booker, Lashon B. (2000). Do We Really Need to Estimate Rule Utilities in Classifier Systems? In *Lecture Notes in Artificial Intelligence 1813*, Pier Luca Lanzi, Wolfgang Stolzmann, and Stewart W. Wilson, (Eds). Springer.

Budescu, David V. and Ilan Fischer. (2001). The Same but Different: An Empirical Investigation of the Reducibility Principle. Journal of Behavioral Decision-Making, 14, p. 187-206.

De Groot, Morris H. (1970). *Optimal Statistical Decisions*. McGraw-Hill.

Ellsberg, Daniel. (1961). Risk, Ambiguity, and the Savage Axioms. *Quarterly Journal of Economics*. Vol 75, p. 643-669.

Fishburn, Peter C. (1981). Subjective Expected Utility: A Review of Normative Theories. *Theory and Decision*, 13:2, p. 139-199.

Goldberg, David E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley.

Grefenstette, John J. (1991). Conditions for Implicit Parallelism. In *Foundations of Genetic Algorithms*, Gregory J.E. Rawlins, ed. Morgan Kaufmann Publishers.

Holland, John H. (1985). Properties of the Bucket Brigade Algorithm. In *Proceedings from the First International Conference on Genetic Algorithms*, Lawrence Erlbaum Associates, p. 1-7.

Holland, John H. (1995). *Hidden Order: How Adaptation Builds Complexity*. Addison-Wesley.

Holland, John H. and John H. Miller. (1991). Artificial Adaptive Agents in Economic Theory. The American Economic Review, 81:2, p. 365-370.

Holland, John H. and J.S. Reitman. (1978). Cognitive Systems Based on Adaptive Algorithms. In *Pattern Directed Inference Systems*, D.A. Waterman and F. Hayes-Roth, eds. Academic Press.

Holland, John H., Lashon B. Booker, Marco Colombetti, Marco Dorigo, David E. Goldberg, Stephanie Forrest, Rick L. Riolo, Robert E Smith, Pier Luca Lanzi, Wolfgang Stolzmann, and Stewart W. Wilson. (1998). What Is a Learning Classifier System? In *Lecture Notes in Artificial Intelligence 1813*, Pier Luca Lanzi, Wolfgang Stolzmann, and Stewart W. Wilson, (Eds), p. 3-32). Springer.

Kahneman, Daniel and Amos Tversky. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47:2, p. 263-292.

Krebs, J.R. (1973). Behavioral Aspects of Predation. In *Perspectives in Ethology*, P. Bateson and P. Klopfer, eds, p. 73-111.

McClure, Samuel M., David I. Laibson, George Loewenstein, and Jonathan D. Cohen. (2004). Separate Neural Systems Value Immediate and Delayed Monetary Rewards. *Science*, 306, p. 503-507.

Ramsey, Frank P. (1926). Truth and Probability. In Ramsey, 1931, *The Foundations of Mathematics and Other Logical Essays*, Ch. VII, p. 156-198, edited by R.B. Braithwaite. Harcourt, Brace and Company. 1999 electronic edition.

Real, Leslie A. (1980). Fitness, Uncertainty, and the Role of Diversification in Evolution and Behavior. *The American Naturalist*, 11:5, p. 623-638.

Real, Leslie and Thomas Caraco. (1986). Risk and Foraging in Stochastic Environments. Annual *Review of Ecology and Systematics*, 17, p. 371-390.

Real, Leslie A. (1987). Objective Benefit Versus Subjective Perception in the Theory of Risk-Sensitive Foraging. *The American Naturalist*, 130:3, p. 399-411.

Riolo, Rick L. (1987a). Bucket Brigade Performance: I. Long Sequences of Classifiers. In *Proceedings from the Second International Conference on Genetic Algorithms*, John J. Grefenstette, ed, p. 184-195. Lawrence Erlbaum Associates.

Riolo, Rick L. (1987b). Bucket Brigade Performance: II. Default Hierarchies. In *Proceedings from the Second International Conference on Genetic Algorithms*, John J. Grefenstette, ed, p. 196-201. Lawrence Erlbaum Associates.

Savage, Leonard J. (1954). *The Foundations of Statistics*. Wiley.

Schoener, Thomas W. (1971). Theory of Feeding Strategies. *Annual Review of Ecology and Systematics*, 2, p. 369-404.

Stearns, Stephen C. (2000). Daniel Bernoulli (1738): Evolution and Economics under Risk. *Journal of Biosciences*, 25:3, p. 221-228.

Vickrey, William. (1961). "Counterspeculation, Auctions, and Competitive Sealed Tenders." Journal of Finance. 16:1, p. 8-37.

Wilson, Stewart W. (1986). Hierarchical Credit Allocation in a Classifier System. *Research Memo RIS No. 37r*. The Rowland Institute of Science.

Wilson, Stewart W. (1995). Classifier Fitness Based on Accuracy. *Evolutionary Computation*, 3:2, p. 149-175.

Wilson, Stewart W. and David E. Goldberg. (1989). A Critical Review of Classifier Systems. In *Proceedings from the Third International Conference on Genetic Algorithms*, J. David Schaffer, ed., p. 244-255. Morgan Kaufmann.

Wright, Sewall. (1937). The Distribution of Gene Differences in Populations. *Proceedings of the. National Academy of Sciences*. 23:6, p. 307-320.