A FUZZY RESOLUTION OF THE PRISONER'S DILEMMA

Kalyan Raman
Professor, School of Management, University of Michigan Flint &
Associate Faculty, CSCS, University of Michigan Ann Arbor

January 2002

The author is indebted to Professor John Hauser, Sloan School of Management, M.I.T., Cambridge, Dr. Rick Riolo, CSCS, University of Michigan Ann Arbor, and to participants of the 2001 Marketing Science Conference in Wiesbaden, Germany for their helpful comments and insightful suggestions on this research. Part of this work was completed at Katholieke Universiteit Leuven, Belgium during Summer and Fall of 2001. The author gratefully acknowledges the resources generously made available by Katholieke Universiteit Leuven, Belgium. Partial support for this project came from a FDAC grant awarded by the University of Michigan Flint.

Abstract

The prisoner's dilemma is a powerful metaphor for the difficulty of obtaining a cooperative outcome from the interactions among a group of self-interested individuals. Proposed ways of attaining the cooperative outcome include infinitely repeated interaction among the players or finitely repeated interaction with an uncertain date for the last play of the game, fictitious play, tit-for-tat strategy, credible threats, and coalitions. However, each of these solutions has well-known limitations, thus motivating a need for fresh approaches.

This analysis uses the concepts of bounded rationality, heuristics and fuzzy control theory to bring a new perspective on the use of threats as a means of seeking cooperation. Strictly binary classification of a signal as either a threat or not a threat makes cooperation difficult to achieve in a prisoner's dilemma. I show that a fuzzy control analysis permits the emergence of various modes of cooperative behavior. The fuzzy view makes both conceptual and pragmatic sense, and includes the conventional binary prisoner's dilemma as a special case. Classical analysis of the prisoner's dilemma separates concepts and their opposites into rigidly dichotomous, diametrically opposed and mutually exclusive categories of meaning. Fuzzy analysis views concepts and their opposites as mutually overlapping linguistic categories, such that the meaning of each concept gradually shades off into the meaning of its opposite, and the fuzzy perspective leads to the emergence of many partially cooperative outcomes including the extreme outcomes of total cooperation and total non-cooperation.

KEYWORDS

Competition, Cooperation, Prisoner's Dilemma, Marketing Strategy, Fuzzy Control Theory

INTRODUCTION

The prisoner's dilemma (Gibbons 1992) is a powerful metaphor for the difficulty of obtaining a cooperative outcome from the interactions among a group of self-interested individuals. Proposed ways of moving towards the cooperative outcome include infinitely repeated interaction among the players or finitely repeated interaction with an uncertain date for the last play of the game, fictitious play (Fudenberg and Tirole 1991), Axelrod's tit-for-tat strategy (Axelrod 1984, Axelrod and Dion 1988) and the formation of coalitions (Fader and Hauser 1988). There is a vast literature on the use of credible threats to induce cooperation. Each of these solutions has well-known limitations; infinite plays guarantee cooperation but an infinite horizon game lacks face validity for some applications, uncertain ending dates do not guarantee cooperation, fictitious games provide only asymptotic assurance of cooperation, the tit-for-tat strategy can lead to long non-cooperative cycles when signals are misinterpreted to be threats, and the work on coalitions is concerned with the evolution of cooperation as a result of the influence of the third player on the other two in a three-player, repeated game.

The importance of finding ways to resolve the prisoner's dilemma, hereafter referred to as PD, can hardly be overstated, given its conceptual power and far-ranging applications in fields spanning economics, marketing, political science, production management, cross-functional team management, biology and engineering. Rigorous microeconomic resolutions of the PD have been based on fully rational behavior and strict binary logic, but none of the proposed solutions has proven to be entirely satisfactory. The assumption of binary logic implies that players classify signals in a dichotomous manner. Interpreted purely in terms of binary or Aristotelian logic, a signal is either seen as a threat or it is not. However, binary classification is unduly restrictive and somewhat artificial, for in the real world, both empirical observations and common sense inform us that signals are perceived as threats to varying degrees. Fuzzy logic accommodates such a perspective whereas classical binary logic does not. An example will illustrate the difference between these two views. In the fuzzy framework, when a company invests in additional capacity, the degree to which a competitor perceives it as a threat is an increasing function of the amount of investment. In the binary framework, a competitor either perceives the investment as a threat or does not. Surely, the fuzzy view makes both conceptual and pragmatic sense. I develop one possible resolution of the PD through fuzzy analysis, thereby bringing a new perspective to the use of threats as a means of seeking cooperation. Classical microeconomic analysis of the PD is grounded in the twin assumptions of binary logic and full rationality. Fuzzy methodology discards these restrictions, and is distinguished by its ability to incorporate bounded rationality (Simon 1957) and accommodate a rich classification of signals along a continuum.

DESCRIPTION OF A FUZZY APPROACH TO THE PD

Since descriptions of the PD abound elsewhere (Axelrod 1984, Gibbons 1992), I will move quickly to the details of my analysis. I assume a perfectly symmetric game so that all players face the same penalties, rewards and choices. I assume two players in this analysis, but the fuzzy controller logic does not require any distinction between the two-player and multiple player case. Therefore, in principle, the fuzzy controller accommodates multiple players. In the traditional PD analysis, a signal is either perceived as a threat or it is not. If a signal of magnitude x is perceived as a threat, then it is classified in the set of threatening signals T. This binary logic implies a discontinuous membership function $T_{bin}(x)$ for the set of threatening signals defined as follows $T_{bin}(x) = 1$, if x is perceived as a threat, and 0 if it is not. In the fuzzy perspective, the set of threatening signals is defined by a membership function $T_{fuzz}(x)$, where $T_{fuzz}(x)$ is a continuous function ranging from 0 to 1 to admit different grades of membership. $T_{fuzz}(x) = 0$ indicates that x is perceived to be non-threatening, $T_{fuzz}(x) = 1$ indicates that x is perceived to be completely threatening, and intermediate values of $T_{fuzz}(x)$ indicate that x is perceived to be partially threatening. In the language of fuzzy logic, intermediate values of $T_{fuzz}(x)$ define the degree to which the signal is perceived to be a threat. Clearly, the conventional binary case is subsumed within this framework. Following standard terminology, we say that the set T is crisp (Kosko 1997) in the binary case and fuzzy otherwise. Without loss of generality, I assume that the signal space is the unit interval [0, 1] and that the signal's threat content increases with x.

The usual binary analysis of the PD restricts players to two possible actions: Cooperate or Defect. In the fuzzy case, the action set can be represented more generally as a fuzzy set C. This assumption basically implies that players choose from a continuous action space. Fader and Hauser (1988) also postulate a continuous action space for the players but do not use fuzzy analysis. Without loss of generality, I assume that the cooperation space is the unit interval [0, 1] and that the outcome's cooperative nature increases with x. Thus, 0 indicates total non-cooperation and 1 indicates total cooperation. Full membership in C ($C(x) = 1$) indicates cooperative behavior, zero membership in C ($C(x) = 0$) indicates non-cooperative behavior, and intermediate values of $C(x)$ indicate varying degrees of cooperative behavior. While intermediate values of $C(x)$ are not interpretable within a literal interpretation of the PD (each prisoner either cooperates or does not), they make sense in many real-world applications of the PD. For example, in leader-follower games prevalent in oligopolies, firms generally follow price reductions or advertising increments of the leader to varying degrees. Even when intermediate values of $C(x)$ are not interpretable as a prescription for a specific action, they are interpretable as the frequency of cooperative behavior.

What are the conceptual implications of treating the level of cooperation as a continuous variable ranging over [0,1]? In the binary treatment of the PD, the outcome is considered to be completely non-cooperative below a threshold level, and above this threshold, the outcome is considered to be completely cooperative. In the fuzzy treatment of the PD, the outcome is considered to be increasingly non-cooperative as we go lower and lower below a threshold level, and the outcome is considered to be increasingly cooperative as we go higher and higher above this threshold. Thus, in the fuzzy framework, cooperation is a matter of degree rather than an "All or Nothing" outcome.

Thus, by postulating a continuum of cooperation levels (any number in the unit interval [0,1]), the *fuzzy controller places no restriction on the possible nature of the cooperative outcomes* and does not impose an artificial binary choice (Cooperate or Defect) on the players. The decision space is the entire continuum [0,1]. *We will however see that when the players classify threats dichotomously, they will always generate a binary pattern of cooperation/defection, but that when they view threats as a matter of degree, then they will generate a variety of cooperation patterns ranging from complete non-cooperation to complete cooperation with infinitely many intermediate degrees of cooperation.* We may diagrammatically represent the logic of the binary and fuzzy games as in Figure 1.

OPTIMALITY ISSUES

The classical economic paradigm of humans as fully rational utility-maximizing beings has been challenged by the Nobel-prize winning economist Herbert Simon, who proposed that humans rely upon an alternative process that he termed "bounded rationality" in making their choices (1957). In Simon's view, evolution did not give rise to optimal agents, but to agents who are locally optimal at best, or locally satisfactory in some sense. Indeed, even if agents were capable of solving complicated optimization problems that a mathematician might find daunting, they simply would not have all the information necessary to be globally optimal. Thus, in Simon's opinion, a theory based upon optimal behaviors is tenuous at best. In the "bounded rationality" paradigm, agents are assumed to be rational within limits and to satisfice rather than optimize. An immediate dilemma posed by the concept of satisficing is that there may be many different ways to satisfice and many possible outcomes. The ideal optimizer *Homo economicus* has a well-defined objective and an unambiguous course of action, but the boundedly rational being must choose the manner in which she will bound her rationality. Recent work in psychology and consumer behavior suggests a plausible way of bounding rationality. This approach is spawned by the literature on heuristics or cognitive shortcuts, a research stream rooted in extensive studies of decision-making by Kahneman (1973).

Seminal research by Kahneman and Tversky (1987, 1984) shows that subjective judgments generally do not obey the basic normative principles of decision theory, and that the axioms of rational choice are often violated consistently by sophisticated as well as naive respondents. Human judgments appear to follow certain principles that often lead to reasonable answers. Kahneman and Tversky's work shows that people often use logical shortcuts called heuristics. Lindsay and Norman (1972) define heuristics as "rules of thumb" or general action plans that people use in problem-solving. Such heuristics are generally operationalized as production systems featuring "If-Then-Else" rules. Heuristics are simple rules that are easy to apply, and they exist because they reduce cognitive effort by simplifying decision-making in the face of cognitive complexity. Appealing to this literature, the cognitive demands placed by rational behavior are bounded by heuristics in this analysis. The use of heuristics is a widely-accepted account of consumer decision-making in modern theories of information processing.

In the context of the PD, players are assumed to use the following heuristics in guiding their choice of strategy. The player observes the signal, classifies the extent to which it is a threat and makes a decision according to the following rules:

"If the threat is low, then do not cooperate"

"If the threat is high, then cooperate"

In the next level of sophistication, players judge both the level of the threat and its credibility and use that information to guide their action. This leads to the following heuristics:

"If the threat is low and its credibility is high, then do not cooperate"

"If the threat is low and its credibility is low, then do not cooperate"

"If the threat is high and its credibility is low, then do not cooperate"

"If the threat is high and its credibility is high, then cooperate"

FUZZY ANALYSIS OF THE PD

I will assume that the signal is received without error. I will farther assume that the signal is nonnegative and finite-valued; without loss of generality, I may then assume that the signal strength x lies between 0 and 1. A signal x is received and classified as a threat to degree T(x). Thus the player's input is the fuzzy set T. A crisp prescription for action is obtained by mapping this fuzzy input into a fuzzy output set and then defuzzifying it. The output fuzzy set is the action set, taken to be the unit interval. The fuzzy controller maps the input fuzzy set to the output fuzzy set by using the rules of fuzzy logic to implement each of the "If-Then" rules that prescribe behavior as a function of observed threat. Figure 2 shows an overview of the fuzzy control system and Figure 3 shows its architecture.

FUZZY CONTROLLER LOGIC AND OUTPUT

The fuzzy controller uses fuzzy inference rules of the following form:

If T = □, Then C = Δ

In the above proposition, T denotes the level of threat, C the level of cooperation, and the rectangular box and triangular box are filled in with fuzzy linguistic qualifiers such as "Low," "High," "Moderate," or any appropriate descriptors that are meaningful for the problem. For example, we may successively substitute the words Low or High in both the boxes together. The states of T (Low or High) are called *antecedents* and the states of C (Low or High) are called *consequents*. Only a rule for which the compatibility of the measured value of the threat with the antecedent is *positive* will participate in determining the value of the controlled variable (advertising spending). Thus, for example, suppose that levels of threat measured to be greater than 0.25 are considered high and those below 0.25 are considered low threat situations. If the threat level is measured to be 0.40, then this value of the state has zero degree of membership in the Low Threat fuzzy set and hence zero compatibility with the fuzzy inference rule "If C = Low, Then S = Low"; hence this particular rule will not participate in the control process. On the other hand, the measured value of the state 0.40 does have positive membership degree in the High Threat fuzzy set and hence the fuzzy inference rule "If C = High, Then S = High" will become instrumental in determining the value of the controlled variable. The fuzzy inference rules that participate in the control mechanism are called the *rules that fire*.

Thus a specific measured value of the input variable is received by the controller, its compatibility with the antecedent of each fuzzy inference rule is assessed through the membership functions for the fuzzy sets measuring threat, and the rules that fire are thereby determined. An inference is made by each rule that fires. The conclusion obtained by each rule that fires is a fuzzy set. Given the conclusions obtained by the individual fuzzy inference rules, the overall conclusion is obtained by taking the fuzzy union of all the individual conclusions, resulting in the final fuzzy set for prescribing the controller's action. This last step requires a process called defuzzification the purpose of which is to convert the fuzzy set representing the overall conclusion into a real number in [0,1] that in some sense, best represents the fuzzy set. This real number determines the cooperation level. Defuzzification is accomplished in a number of ways, the most popular of which is the Center of Gravity technique (Kosko (1991, 1997), Klir 1997) abbreviated COG by some authors. Using COG defuzzication, the player takes an action "c" determined as follows: if H(x) is the membership function of the overall conclusion fuzzy set, then the center of gravity of the fuzzy set H(x), denoted COG[H(x)], is given by:

$$\mathrm{COG}[H(x)] = \frac{\int_{-\infty}^{\infty} x\, H(x)\, dx}{\int_{-\infty}^{\infty} H(x)\, dx}$$

With these concepts established, the fuzzy control algorithm works as follows.

1. Compute the centroid c_j and volume V_j of each output fuzzy set. Denoting the membership function of the j-th output fuzzy set by $b_j(x)$, these are defined as:

$$V_j = \int_{-\infty}^{\infty} b_j(y)\, dy$$

$$c_j = \frac{\int_{-\infty}^{\infty} y\, b_j(y)\, dy}{\int_{-\infty}^{\infty} b_j(y)\, dy}$$

The integrals are replaced by p-dimensional integrals if the output is a vector in R^p. In that case the integrals are of the form $\int_{R^p} dy_1 dy_2 ... dy_p$.

2. Applying the Standard Additive Model (SAM) and invoking Kosko's (1997) SAM Theorem, we compute the input-output map $x \rightarrow F(x)$ as follows:

$$F(x) = \frac{\sum_{j=1}^{j=m} w_j\, a_j(x)\, V_j\, c_j}{\sum_{j=1}^{j=m} w_j\, a_j(x)\, V_j}$$

The weights w_j are chosen to reflect the importance of each rule or heuristic stored in the controller. Since we have no conceptual or empirical reason to discriminate between the rules, we assume equal weights. Setting them all equal yields the following mapping

$$F(x) = \frac{\sum_{j=1}^{j=m} a_j(x)\, V_j\, c_j}{\sum_{j=1}^{j=m} a_j(x)\, V_j}$$

The above fuzzy controller logic essentially implements the heuristic reasoning of each player. Therefore, in this sense, it makes no difference whether two or more players play the PD game.

Theorem 1

When players classify threats in a strict binary manner, the traditional PD outcomes are obtained.

Proof

Assume that the membership functions for the low and high threat fuzzy sets are respectively indexed by j = 1 and 2. When players classify threats in a strict binary manner, only threshold level information about a signal is considered and all magnitude information is discarded. In other words, if the signal exceeds a critical level τ, it is considered a threat; otherwise, it is not. This implies that the membership functions $a_H(x)$ for the high threat set and $a_L(x)$ for low threat set are representable in terms of the unit step function U(x), where U(x) = 1 for x > 0 and U(x) = 0 otherwise. Using a shifted unit step function, we can write:

$a_L(x) = a_1(x)\, U(x - \tau)$

$a_H(x) = a_2(x)\, (1 - U(x - \tau))$

This specification of the membership functions for low and high threat fuzzy sets renders those sets crisp since the boundary between low and high threats is sharp and unambiguous. Hence $a_L(x) = 0$ if and only if $a_H(x) \neq 0$. When $x \leq \tau$, the signal is not considered to be a threat at all and when $x > \tau$, the signal is considered to be completely a threat. Substitution of these discontinuous membership functions and straightforward algebra reduces the mapping $x \rightarrow F(x)$ to the following expression.

$$F(x) = \begin{cases} c_1 & \text{for } x \leq \tau \\ c_2 & \text{for } x > \tau \end{cases}$$

In the above mapping, c_1 is the centroid of the non-cooperation set and c_2 is the centroid of the cooperation set. In fuzzy control theory, the centroid of a fuzzy set represents that set. Hence the above map translates into the following prescription for action.

Do not Cooperate if the threat is low

Cooperate if the threat is high

These are exactly the two extreme outcomes obtained in the classic PD.

Q.E.D.

Theorem 2

When players view signals as threats to varying degrees rather than in a strict binary manner, an infinite number of partially cooperative outcomes are obtained. Furthermore, the degree of cooperation increases with the degree of threat under the following conditions, where $a_1(x)$ and $a_2(x)$ are respectively the membership functions for the low threat and high threat fuzzy sets:

$$\operatorname*{Limit}_{x\to 1}[a_1(x)] = 0 \text{ and } \operatorname*{Limit}_{x\to 0}[a_2(x)] = 0$$

$$\operatorname*{Limit}_{x\to 0}[a_1(x)] \text{ and } \operatorname*{Limit}_{x\to 1}[a_2(x)] \text{ exist and are finite-valued}$$

The traditional PD outcomes are obtained only in the extreme cases when a signal is considered to be completely threatening or completely non-threatening.

Proof

Assume that the membership functions for the low and high threat fuzzy sets are respectively indexed by j = 1 and 2. The condition $\underset{x->1}{\text{Limit}}[a_1(x)] = 0$ is the very reasonable requirement that as the threat grows larger and larger in magnitude and approaches its maximum value, its membership in the low threat fuzzy set must go to zero. The condition $\underset{x->0}{\text{Limit}}[a_2(x)] = 0$ says that as the threat grows smaller and smaller in magnitude and approaches its minimum value, its membership in the high threat fuzzy set must go to zero. When players view signals as threats to varying degrees rather than in a strict binary manner, all signals are considered to be threatening to some degree and non-threatening to some degree. It follows that the membership functions $a_1(x)$ for the low threat set and $a_2(x)$ for the high threat set are not representable in terms of unit step functions U(x), since there is no sharp boundary demarcating a threatening signal from one that is non-threatening. In this case, the mapping $x \to F(x)$ is:

$$F(x) = \frac{\sum_{j=1}^{j=m} a_j(x) V_j c_j}{\sum_{j=1}^{j=m} a_j(x) V_j}$$

Next I establish that F(x) is a monotonically increasing function of x. The first derivative F'(x) of the mapping F(x) is given by:

$$F'(x) = \frac{(a_2(x)\ a_1'(x) - a_1(x)\ a_2'(x))\ (c_1 - c_2)\ V_1\ V_2}{(V_1\ a_1(x) + V_2\ a_2(x))^2}$$

The quantities V_1 and V_2 are always positive because they are volumes. The denominator is always positive since it is a squared quantity. Since $(c_1 - c_2) < 0$, it follows that $F'(x) > 0$ if $(a_2(x)\ a_1'(x) - a_1(x)\ a_2'(x)) < 0$. Both $a_1(x)$ and $a_2(x) > 0$ since they are membership functions. Since $a_1(x)$ is the membership function for the low threat set, it decreases as x increases, hence $a_1'(x) < 0$. Since $a_2(x)$ is the membership function for the high threat set, it increases as x increases, hence $a_2'(x)) > 0$. Putting all this together, it follows that $F'(x) > 0$, and therefore F(x) is an increasing function of x.

Finally, the conditions that $\underset{x->0}{\text{Limit}}[a_1(x)]$ and $\underset{x->1}{\text{Limit}}[a_2(x)]$ exist and are finite-valued ensure that $\underset{x->0}{\text{Limit}}[F(x)] = c_1$ and $\underset{x->1}{\text{Limit}}[F(x)] = c_2$, where c_1 is the lowest level of cooperation (non-cooperation) and c_2 is the highest possible level of cooperation. Thus, as the signal becomes completely non-threatening, we obtain the non-cooperative outcome, and as the signal becomes completely threatening, we obtain the fully cooperative outcome. In between these two extremes, intermediate degrees of threat elicit intermediate degrees of cooperative outcomes, and there are infinitely many such outcomes since F(x) is real-valued.

Q.E.D.

DISCUSSION OF RESULTS

Theorem 1 establishes that binary classification of signals into a threat/non-threat dichotomy always induces total cooperation or total non-cooperation as the only possible outcomes. This is true even though the players have an infinite continuum of choices ranging between total cooperation and total non-cooperation. Figure 4 illustrates the theorem in the case where all levels of the signal below 0.25 are considered non-threatening and all signal levels above this threshold are considered threatening. The totally cooperative outcome is always obtained whenever the signal level exceeds 0.25 and the totally non-cooperative outcome is obtained whenever the signal level is below 0.25. In Figure 4, the membership functions are constant step functions. Note however that Theorem 1 has been proven for the more general case of non-constant membership functions. Thus, the membership function $a_1(x)$ for the low threat fuzzy set in Figure 4 could be an arbitrarily decreasing function of x and the membership function $a_2(x)$ for the high threat fuzzy set in Figure 4 could be an arbitrarily increasing function of x and we would still obtain total cooperation and total non-cooperation as the only possible outcomes. The key point of Theorem 1 is that it holds whenever players make a binary classification of threat signals such that a signal is considered "Totally Threatening" or "Totally Non-threatening" according to whether or not it exceeds a threshold.

Figures 5 through 9 illustrate the content of Theorem 2. In Figure 5, the membership functions are linear and have non-zero overlap. The conceptual implications of the overlapping membership functions in Figure 5 are: (a) when the signal level is close to zero, it is totally non-threatening, but becomes less and less non-threatening as its level increases, (b) as the signal level exceeds the horizontal axis-intercept of the high threat membership function, the signal is to some degree *both* threatening and non-threatening, and (c) as the signal level exceeds the horizontal axis-intercept of the low threat membership function, it loses its non-threatening character completely and becomes increasingly threatening as its level rises, becoming an unambiguous threat as it approaches its maximum value of unity. In this case, we see that a range of partially cooperative outcomes exist between total cooperation and total non-cooperation.

Figure 6 illustrates a case where signals share threatening and non-threatening characteristics at every possible level ranging from zero to infinity. Unlike the linear case of Figure 5, the membership functions never become zero but only approach it asymptotically. Thus a signal, no matter how small its value, is always threatening to some small degree; and conversely, a signal, no matter how big its value, is always non-threatening to some small degree. In this case, we obtain a continuous range of outcomes ranging from asymptotic non-cooperation to asymptotic cooperation.

Figures 7 through 9 illustrate these ideas for the more general case when players consider the credibility of the signal in addition to its actual level. As the literature suggests, a threat can only be effective if it is credible. Consistent with the fuzzy analysis for the one-dimensional case incorporating a continuous range of variation for threats, credibility level is also permitted to vary along a continuum. As in the uni-dimensional case, so in the multi-dimensional case: membership functions for threat and credibility that have no overlap imply strictly dichotomous classification of signals into mutually exclusive threat/non-threat and credible/non-credible categories. Figure 7 shows that strictly dichotomous classification engenders only two possible extreme outcomes: total cooperation or total non-cooperation. Infinitely many (threat, credibility) pairs generate a non-cooperation sheet at the bottom of Figure 7 and infinitely many other (threat, credibility) pairs generate a cooperation sheet at the top of Figure 7. With binary classification, no possibility of partially cooperative outcomes exists, as seen by the discontinuous jump from the bottom sheet to the top sheet.

However, as with threat, so with credibility: in real life, most signals have some degree of credibility rather than being completely with or completely without credibility. This feature is reflected in membership functions with overlap in Figures 8 and 9. Those figures show that with overlapping classification, infinitely many partially cooperative outcomes exist, as seen by the continuous transition from the bottom sheet to the top sheet.

CONCLUSIONS

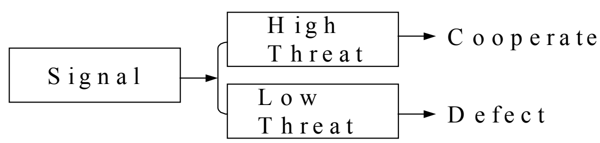
In many practical applications, cooperation and non-cooperation are not matters of 'All or Nothing." In the classical version of the PD, it is indeed the case that the prisoners can only cooperate or defect. But decision makers do not always face a stark choice between total cooperation and total non-cooperation. Furthermore, the bleak dichotomy between total cooperation and total non-cooperation is not the best interpretation of the alternatives available to players in all applications of the PD. Consider for example, the PD games that firms in oligopolies often unwittingly play when they engage in sub-optimally high advertising or promotions. Advertising and promotions are continuous variables with an uncountably infinite number of values, and a realistic application of the PD to such situations of rivalry ought to reflect the continua of choices available to firms that compete on the basis of advertising and promotions. Indeed, most real-world market structures are oligopolies and marketing strategy in oligopolies is heavily dependent on both advertising and promotions because these are the *critical* marketing variables in oligopolistic competition (Berkowitz, Kerin, Hartley, and Rudelius 2000). Thus PD games with continua of choices are the rule rather than the exception in real-world markets.

This analysis shows that fuzzy control enjoys a number of advantages over classical crisp analysis in resolving the PD. Action spaces are often continuous in the real world, human rationality is bounded and people rely upon heuristics to simplify cognitive complexity and ease decision-making. Fuzzy control theory accommodates all these phenomena in a natural way. Fuzzy control makes it evident that the theoretical difficulty of realizing cooperation is closely related to the binary classification of concepts such as signals into threats/non-threats and the binary classification of outcomes into a cooperative/non-cooperative dichotomy. This rigid "All or Nothing" dichotomy makes it difficult to link the PD to observed empirical behavior because its theoretical predictions of non-cooperative behavior are hard to reconcile with the many observed instances of partial cooperation in both the business world and in nature. Fuzzy control shows that the inherent bounded rationality and heuristic nature of human decision-making lead in a natural way to an abundance of partially cooperative outcomes, provided that dichotomous classifications are eliminated. In the fuzzy analysis of the PD, the content of each concept (threat) is partially contained in the content of its polar opposite (non-threat). Thus, concepts (threat/non-threat) and outcomes (cooperation/non-cooperation) share overlap of meaning in the fuzzy *Weltanschaung*, and it is precisely from the shared overlap of meaning between linguistic categories that partial cooperation is emergent.

Much work remains to be done on the PD, and given its extensive cross-disciplinary relevance, it is important to explore a variety of strategies that might elicit cooperation. My goal in this article was to encourage adoption of fuzzy control methodology as a powerful tool in the pursuit of cooperation in PD games. The fuzzy analysis of the PD presented here is meant to be suggestive and provocative rather than the final word on the topic. There are many problems for which the binary "All or Nothing" conceptualization is indeed the correct way to model the issues, and in those cases, the fuzzy approach is unnecessary because it essentially reduces to classical binary analysis. However, in problems where a continuous action space makes conceptual sense, the fuzzy control formulation shows a great deal of promise for generating partially cooperative outcomes.

Figure 1

Binary PD



Fuzzy PD

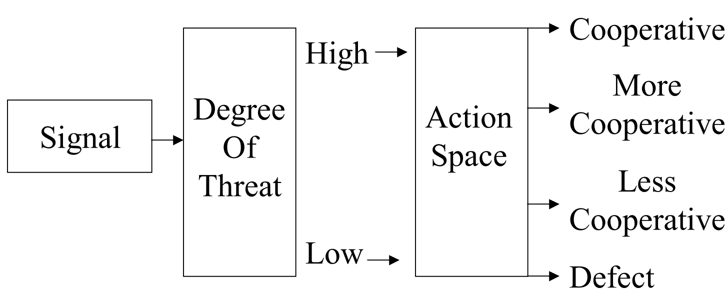


Figure 2

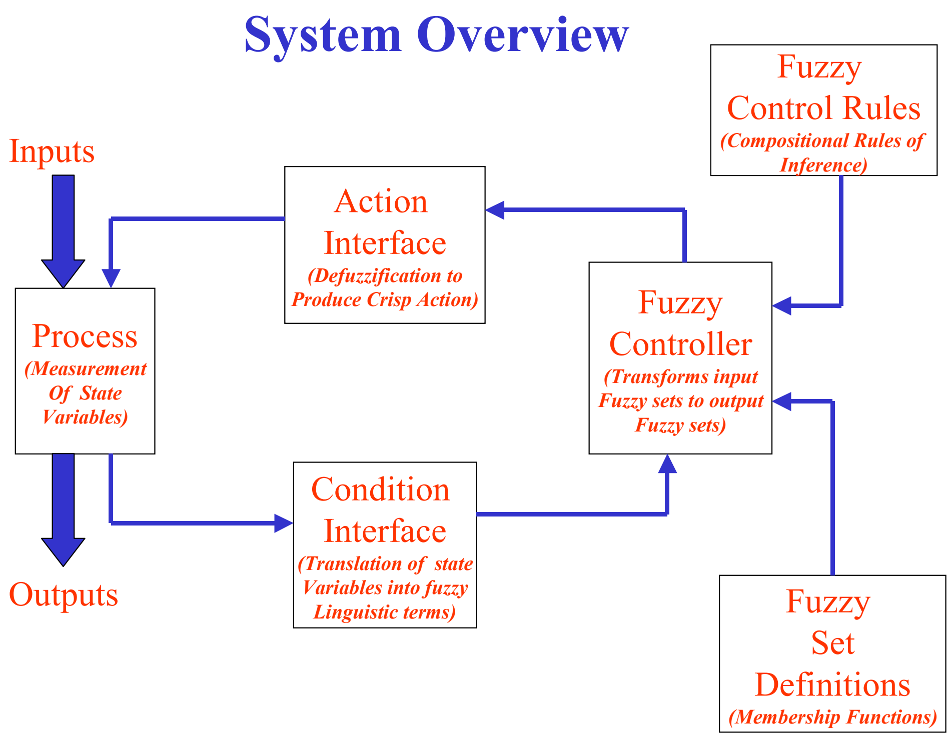

System Overview
Fuzzy
Control Rules
(Compositional Rules of
Inference)
Inputs
Action
Interface
(Defuzzification to
Produce Crisp Action)
Process
(Measurement
Of State
Variables)
Fuzzy
Controller
(Transforms input
Fuzzy sets to output
Fuzzy sets)
Condition
Interface
(Translation of state
Variables into fuzzy
Linguistic terms)
Outputs
Fuzzy
Set
Definitions
(Membership Functions)

Figure 3

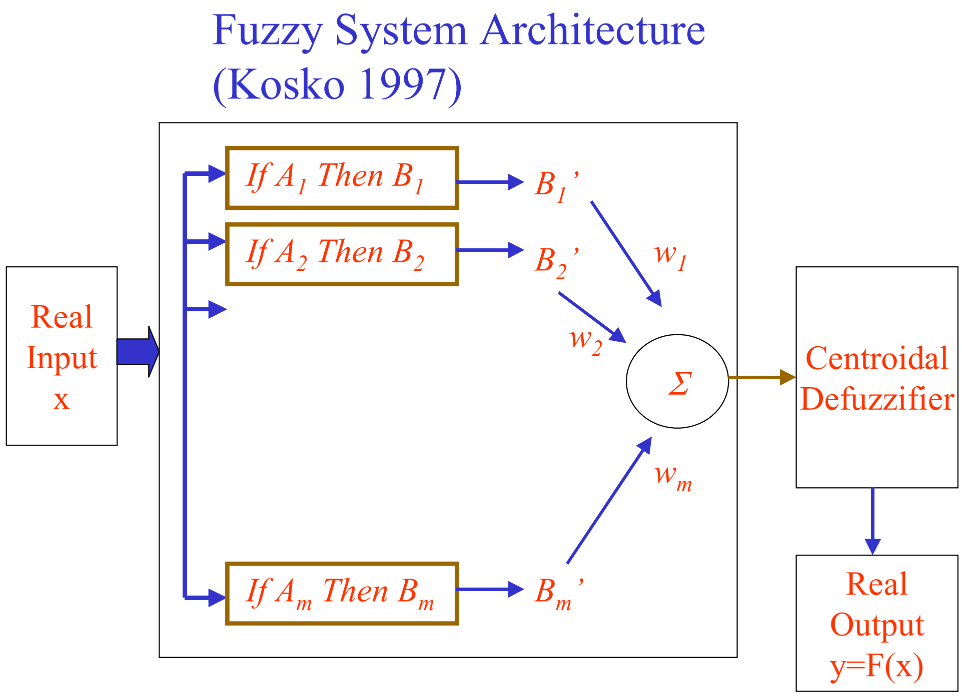

Fuzzy System Architecture
(Kosko 1997)
Real
Input
x
If A_1 Then B_1
If A_2 Then B_2
If A_m Then B_m
B_1'
B_2'
B_m'
w_1
w_2
w_m
Σ
Centroidal
Defuzzifier
Real
Output
y=F(x)

Figure 4

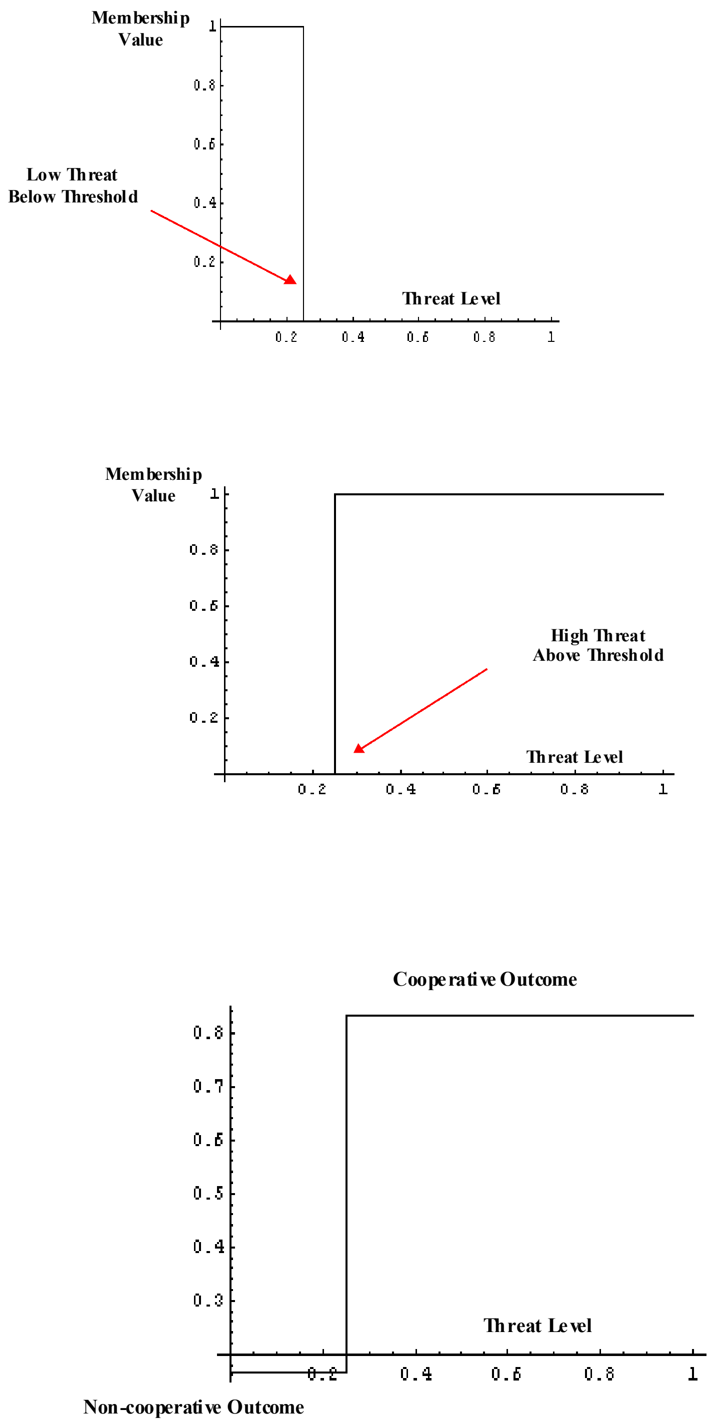
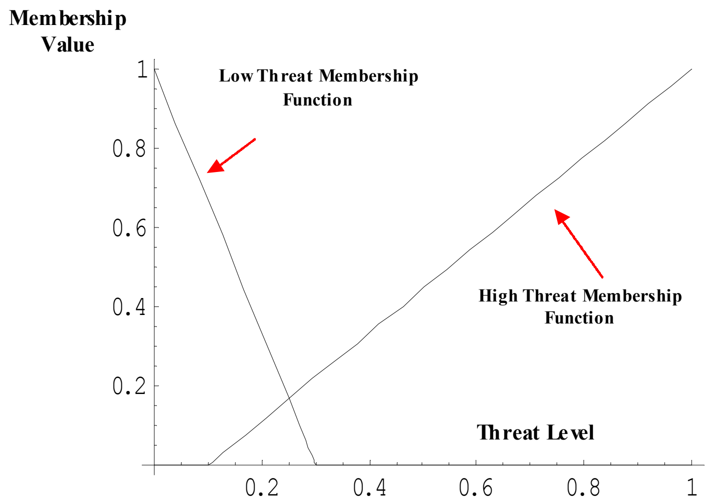
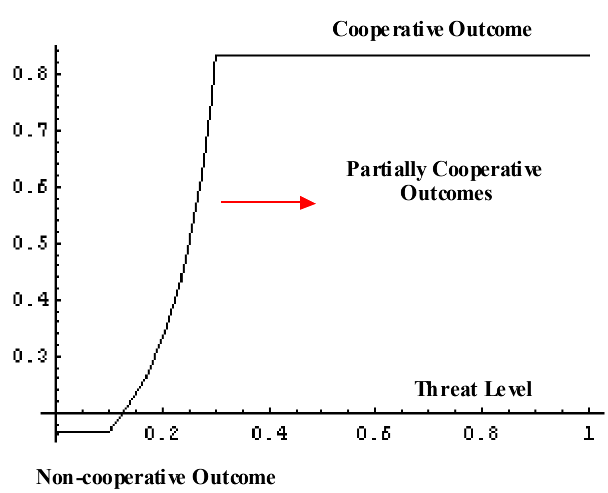

Membership Value
1
0.8
0.6
0.4
0.2
Low Threat Below Threshold
Threat Level
0.2
0.4
0.6
0.8
1
Membership Value
1
0.8
0.6
0.4
0.2
High Threat Above Threshold
Threat Level
0.2
0.4
0.6
0.8
1
Cooperative Outcome
0.8
0.7
0.6
0.5
0.4
0.3
Threat Level
0.2
0.4
0.6
0.8
1
Non-cooperative Outcome

Figure 5

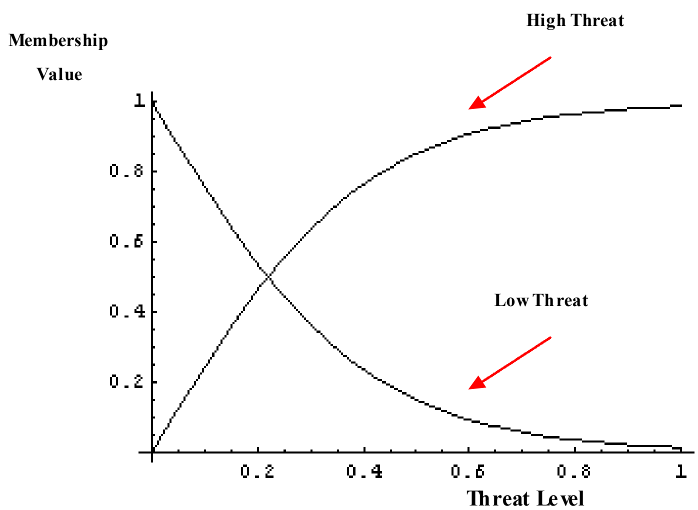


Membership
Value
1
0.8
0.6
0.4
0.2
Low Threat Membership
Function
High Threat Membership
Function
Threat Level
0.2
0.4
0.6
0.8
1

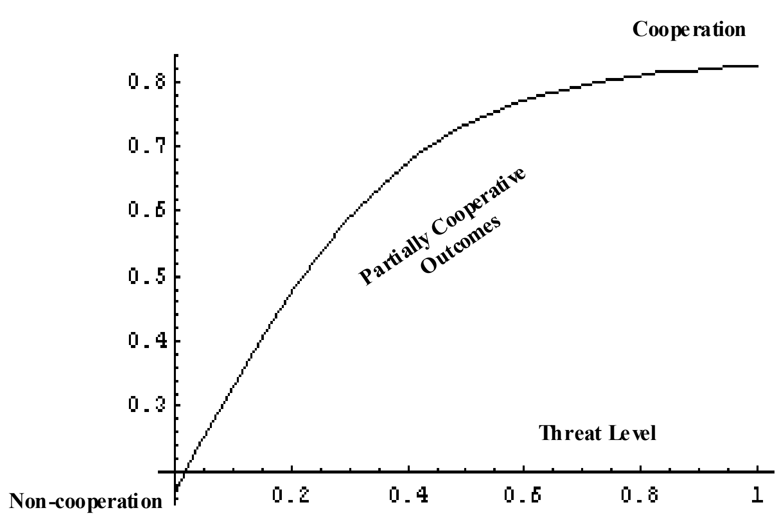


Cooperative Outcome
0.8
0.7
0.6
0.5
0.4
0.3
Partially Cooperative
Outcomes
Threat Level
0.2
0.4
0.6
0.8
1
Non-cooperative Outcome

Figure 6

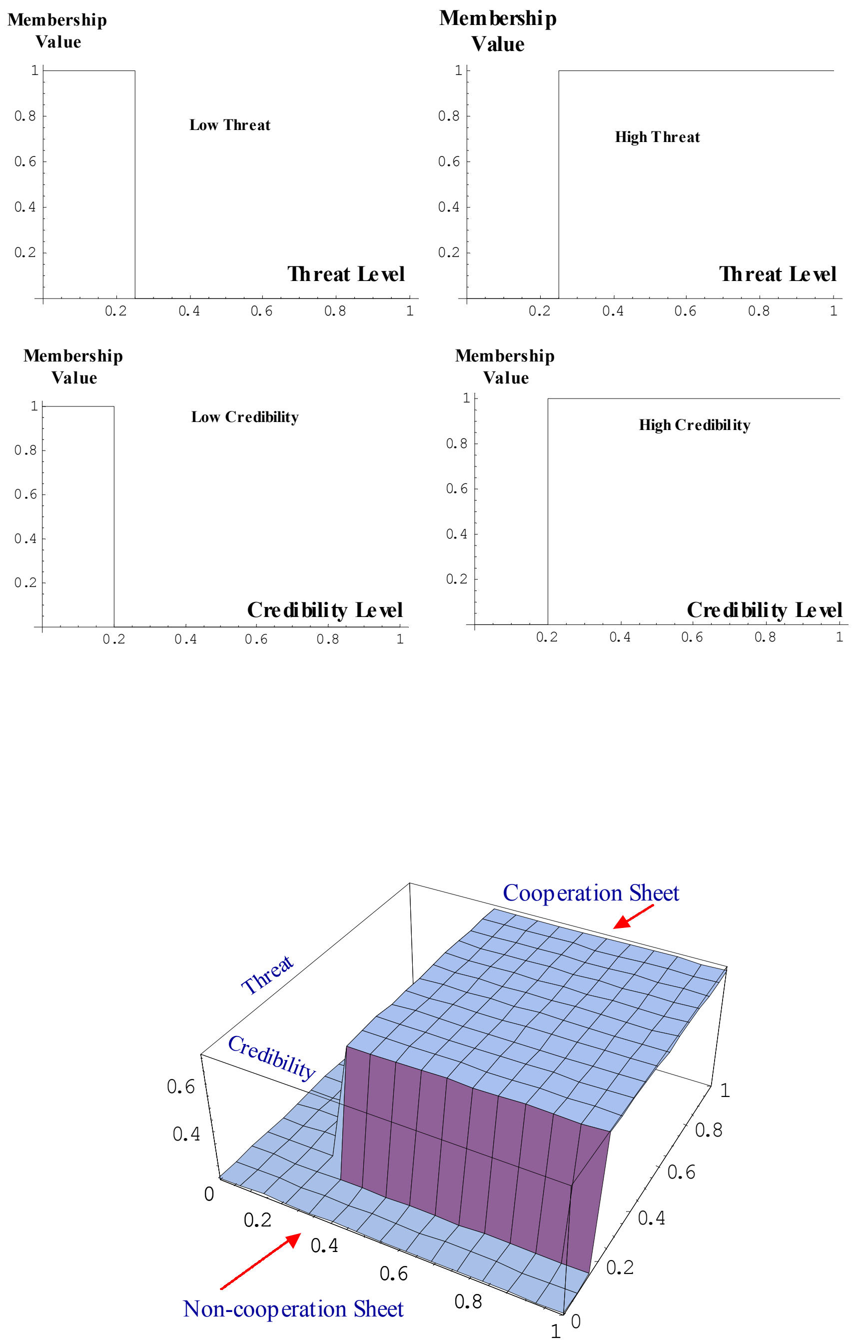


Membership
Value
High Threat
Low Threat
1
0.8
0.6
0.4
0.2
0.2
0.4
0.6
0.8
1
Threat Level



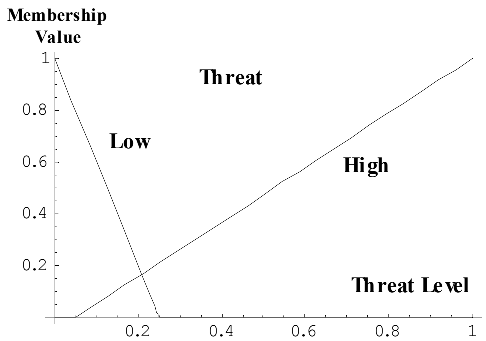
Cooperation
0.8
0.7
0.6
0.5
0.4
0.3
Partially Cooperative
Outcomes
Threat Level
Non-cooperation
0.2
0.4
0.6
0.8
1

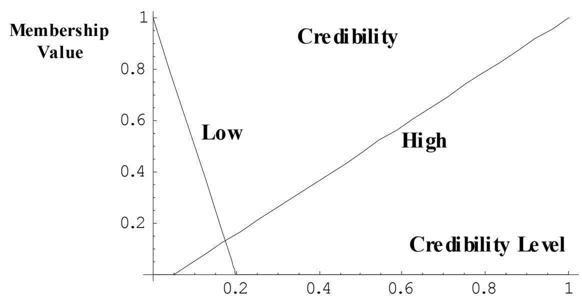
Figure 7



Membership Value
Low Threat
Threat Level
Membership Value
High Threat
Threat Level
Membership Value
Low Credibility
Credibility Level
Membership Value
High Credibility
Credibility Level
Cooperation Sheet
Threat
Credibility
Non-cooperation Sheet

Figure 8


Membership Value
Threat
Low
High
Threat Level
1
0.8
0.6
0.4
0.2
0.2
0.4
0.6
0.8
1


Membership Value
Credibility
Low
High
Credibility Level
1
0.8
0.6
0.4
0.2
0.2
0.4
0.6
0.8
1

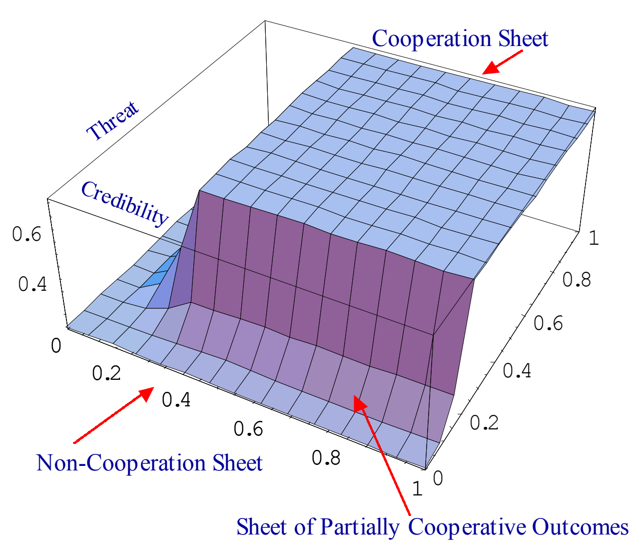
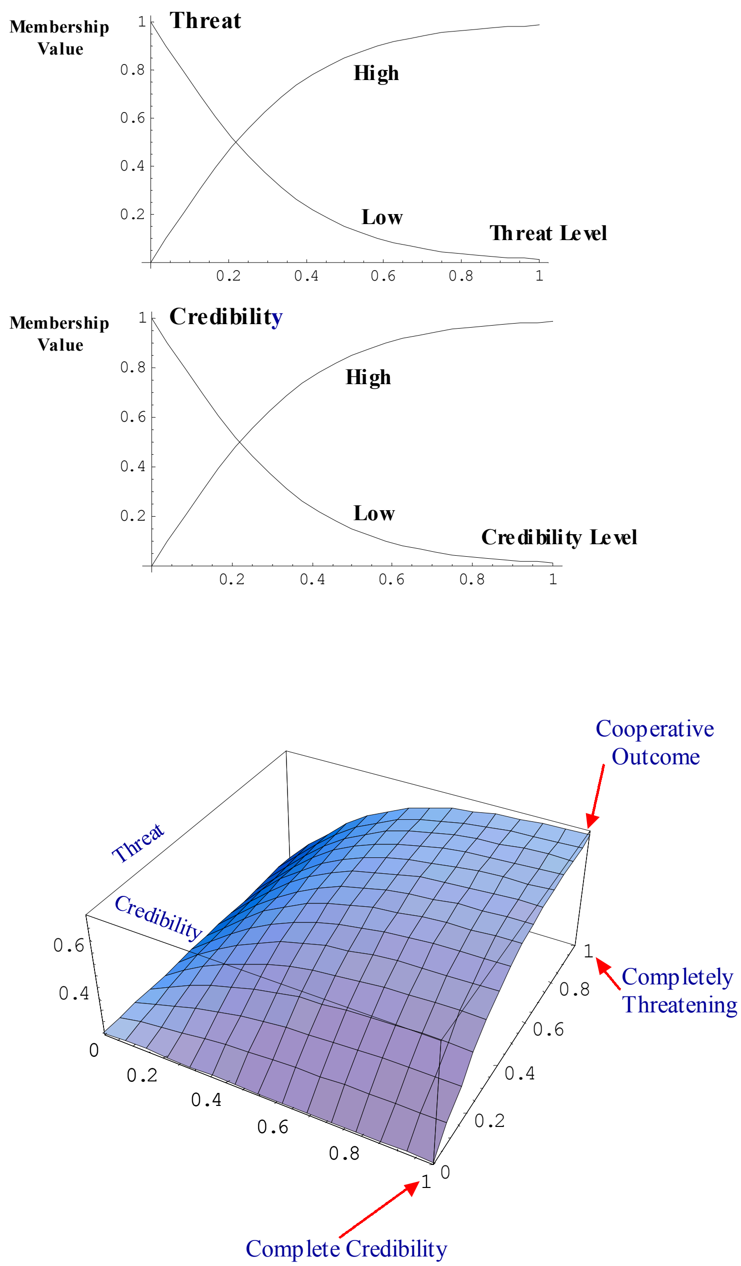

Cooperation Sheet
Threat
Credibility
0.6
0.4
0
0.2
0.4
0.6
0.8
1
0
0.2
0.4
0.6
0.8
1
Non-Cooperation Sheet
Sheet of Partially Cooperative Outcomes

Figure 9


Membership Value
Threat
High
Low
Threat Level
Membership Value
Credibility
High
Low
Credibility Level
Threat
Credibility
Cooperative Outcome
Completely Threatening
Complete Credibility

REFERENCES

Axelrod, R. (1984), *The Evolution of Cooperation*. New York: Basic Books.

Axelrod, R. and D. Dion (1988), *The Further Evolution of Cooperation*." Science, Vol. 242, (December), 1385-1390

Berkovitz, E., R.A. Kerin, S.W. Hartley and W. Rudelius (2000), *Marketing*, 6th Edition, Burr Ridge, IL:Irwin.

Fader, Peter and John R. Hauser (1988), "Implicit Coalitions in a Generalized Prisoner's Dilemma**,"** *The Journal of Conflict Resolution*, V. 32 (September) p. 553-82.

Fudenberg, D. and J. Tirole (1991), *Game Theory*, MIT Press.

Gibbons (1992), Game Theory for Applied Economists, Princeton, NJ: Princeton University Press.

Kahneman, D. (1973), *Attention and Effort*, Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D. and A. Tversky (1984), "Chocies, Values and Frames," *American Psychologist*, 39, 341-350.

Klir, G. (1997), *Fuzzy Set Theory: Foundations and Applications*, Upper Saddle River, NJ : Prentice Hall.

Kosko, Bart (1991), *Neural Networks and Fuzzy Systems*, Upper Saddle River, NJ: Prentice-Hall.

Kosko, Bart (1997), *Fuzzy Engineering*, Upper Saddle River, NJ: Prentice-Hall.

Lindsay, P.H. and D. A. Norman (1972), *Human Information Processing: An Introduction to Psychology*, New York: Academic Press.

Simon, H. (1957), *Models of Man*. Wiley, New York.

Simon, H. (1962), The Architecture of Complexity. *Proceedings of the American Philosophical Society*, 26, 467-482.

Tversky, A. and D. Kahneman (1987), *Judgment under Uncertainty: Heuristics and Biases*, chapter 1, pages 3--22. Cambridge University Press,.