Decision-Making in Policy Governed Human-Autonomous Systems Teams

Diane Felmlee[1], Emil Lupu[2], Cassie McMillan[1], Erisa Karafili[2] and Elisa Bertino[3]

Abstract—**Policies govern choices in the behavior of systems. They are applied to human behavior as well as to the behavior of autonomous systems but are defined differently in each case. Generally humans have the ability to interpret the intent behind the policies, to bring about their desired effects, even occasionally violating them when the need arises. In contrast, policies for automated systems fully define the prescribed behavior without ambiguity, conflicts or omissions. The increasing use of AI techniques and machine learning in autonomous systems such as drones promises to blur these boundaries and allows us to conceive in a similar way more flexible policies for the spectrum of human-autonomous systems collaborations. In coalition environments this spectrum extends across the boundaries of authority in pursuit of a common coalition goal and covers collaborations between human and autonomous systems alike.**

In social sciences, social exchange theory has been applied successfully to explain human behavior in a variety of contexts. It provides a framework linking the expected rewards, costs, satisfaction and commitment to explain and anticipate the choices that individuals make when confronted with various options. We discuss here how it can be used within coalition environments to explain joint decision making and to help formulate policies re-framing the concepts where appropriate. Social exchange theory is particularly attractive within this context as it provides a theory with "measurable" components that can be readily integrated in machine reasoning processes.

Keywords-**Decision making, autonomous systems, social-exchange theory.**

I. INTRODUCTION

At the most generic level "policies" have been characterized as statements "governing the choices in the behavior of a system" [24]. It is advantageous for the following discussion to remain deliberately vague on what we mean by "statements" and what we mean by "systems". Thus in different contexts statements may be defined as objectives, rules of behavior or utility [14] and "systems" may refer to humans, devices, or a mixture of both.

Traditional military coalitions exhibit multiple levels and contexts for policy decision making. At a national level military doctrine governs the way in which military forces engage in operations and the rules of engagement with both civilian and other military forces. At an operational level the principles govern the way in which decisions are made and the desired outcomes of actions. At the coalition level, military policy defines how the military forces from different nation states will cooperate to achieve joint aims and how decisions will be made by the coalitions themselves. At the joint coalition operation level, policy defines the rules of interaction and collaboration that govern tactical operations.

These policies are elaborated and implemented in part by humans and in part by the equipment associated with the military coalitions. When applied to equipment (and we focus here primarily on software systems, although the argument could easily be extended to hardware implementations), policy takes the form of declarative configuration rules that are precise (cover all possible sets of parameter values), deterministic (provide a single decision/outcome) and complete (no actions other than those dictated by the policy are taken) [4]. Changes to the behavior of the equipment require a change of policy to be made through human intervention. Examples of policy are numerous and range from network configuration information [5] to rules for authentication, key management, encryption and communication and to rules covering the gathering and access to information e.g.,[16], amongst others [26]. In contrast to humans who are permitted to interpret policy, with varying degrees of interpretation, to bring about their desired outcomes, equipment is enforcing faithfully policies that dictate every single action.

The rise of autonomy and the increased availability of artificial intelligence techniques in military equipment makes systems computationally capable of autonomous behavior: implementing complex courses of action, complex reasoning about decisions and learning from experience, including witnessing the behavior of others. The future of military coalition environments therefore raises the question of how policies should be formulated for mixed human-autonomous systems collectives pertaining to different organizations both military and civilian, and belonging to different countries.

Social exchange theory has been applied to explain human behavior in a variety of contexts. We seek to understand how it can be applied within the context of future coalition environments and whether it can form a framework for informing how policy should be set in mixed human-autonomous systems collectives.

At an individual level we seek to understand the behavior of humans when subjected to a policy. Can we explain in which circumstances humans depart from the policies applying to their behavior? Can we understand how to formulate policies to maximize compliance? Can we determine which

[1]Department of Sociology and Criminology, Pennsylvania State University dhf12@psu.edu, mcmillan@psu.edu

[2] Department of Computing, Imperial College London, UK. e.karafili, e.c.lupu@imperial.ac.uk

[3] Department of Computer Science, Purdue University bertino@purdue.edu

policies would usefully need to be implemented in order to avoid poor decision making by humans?

Interactions between human and autonomous systems or between autonomous systems themselves are as yet poorly understood and difficult to predict. What insights from the manner in which policies apply to humans can be used in autonomous systems? If we can explain when humans depart or violate the policies applied to them, can we use this understanding to characterize the circumstances in which *policy relaxation* for an autonomous system can occur? More generally how should we specify policies for human-autonomous systems collectives?

The context of coalitions brings a further complexity in which social exchange theory may prove of particular help to explain policy compliance and guide policy formulation to achieve the desired outcomes. In a coalition environment, the different members of the coalition are subject to both national and coalition policies with different degrees of alignment. Understanding how policy affects human behavior leading to different degrees of compliance/violation with the policy provides useful insights into the decision making processes and how best to orchestrate them. Can the same principles be used to govern policy for autonomous systems or mixed human and autonomous systems teams?

Intuitively, autonomous systems act on behalf of their human owners. Thus, theories that explain human behavior should also apply to the behavior of their autonomous *"proxies"*. This may also make it easier for humans to explain and predict the behavior of autonomous components. But this hypothesis has not been explored before in a coalition context, and both social science theories and engineering models need to be adapted to evaluate its implications. We aim to open the discussion here and take the first steps towards this goal.

In Section II we consider the behavior of individuals when their collaborations are governed by policies and discuss the conditions necessary for them to maintain or alter their commitment to policies. We briefly introduce a scenario in Section III, which enables us to conduct a more general discussion in Section IV and conclude in Section V.

II. Decision Making under Policy

We develop a general social science perspective that captures a central aspect of the way in which people form decisions in a social interchange. Exchange theories predict the choices made when people engage in interactions and decisions to optimize their outcomes when faced with possible alternative choices.

The model we formulate builds upon several principles derived from social exchange theory and its offshoots. Social exchange theory [3], [13] is a particularly comprehensive social science perspective, with roots in economics, sociology, and psychology. Its distinct advantages include its generalizability and the fact that it can be readily operationalized. Concepts based on social exchange theory apply to a wide range of human behaviors [8], [9], [19], although they have not yet been applied to coalition operations. Related social science theories include rational choice, behavioral economics, bounded rationality, and game theory. We propose, and modify, several tenets of social exchange theory to apply these propositions to decisions made in collaborations between humans and autonomous systems (AS) [20]. We base our arguments on previous models applied to interchanges in social relationships [6], [23].

When faced with an external crisis (e.g. terrorist attack) or internal upheaval (e.g. government transitions), when do leaders maintain or change their commitment to a specific policy? Conflict assessment is especially challenging, because decision makers must make choices based on incomplete information - with the "right" answer only identifiable after a conflict is over [10], [1]. Using a social exchange approach, we identify the conditions necessary for individuals to maintain or alter their *commitment* to a critical policy.

The basic tenets of our approach, as applied to human commitment to a policy, include:

- People attempt to maximize the $outcomes$ associated with the decisions they make related to policies, with outcomes defined as anticipated $rewards$ minus the anticipated $costs$ associated with the application of a particular policy.
- $Approval$ of a policy is a positive function of $outcomes$ and a negative function of the actor(s) $expectation$ level for outcomes based on past experience or protocol.
- $Commitment$ to a policy, or a line of policies, is a positive function of $approval$ and $investments$ in that policy, and a negative function of the number of high quality, $alternative$ policies of which people are aware.

We summarize the main propositions as follows:

$$Outcomes = Rewards - Costs \tag{1}$$

$$Approval = Outcomes - Expectations \tag{2}$$

$$Commitment = Approval - Alternatives + Investments \tag{3}$$

We define the terminology in the model as:

$Rewards$	anticipated benefits accrued from pursing a policy,
$Costs$	anticipated drawbacks to pursuing a policy,
$Outcomes$	anticipated degree of positive versus negative results associated with a policy,

$Expectations$	the standard, or level, of outcome desired, based on past experience or protocol,
$Approval$	the level of satisfaction or agreement with a policy,
$Alternatives$	perceptions of other policies that could be applied,
$Investments$	the number and magnitude of resources tied to a policy,
$Commitment$	the intention to follow through with the application of a policy.

A. Comparing Rewards and Costs

Rewards represent a generic, utilitarian concept and can be material, or social. Rewards occur at different levels, such as the individual, the coalition, and can extend to the broader society. Material rewards, for example, can include the acquisition of coalition resources, whereas physical rewards would involve the recovery of coalition troops. People could benefit on an individual basis, as well, with a successful mission resulting in social approval from colleagues, for example, or professional advancement. Yet the overall goal of all choices and behaviors, regardless of unit of analysis, is for coalition forces to engage in actions that in the long-run provide benefits at a societal level.

Costs also can come in different forms and take place at various levels. Coalition and civilian casualties represent the most severe, and obvious, potential sacrifices associated with the application of a policy [11]. Material costs could include elements such as lost equipment and severed communication systems. Social costs involve the disapproval from peers and supervisors that can occur when the chosen policy departs from those preferred by others, and/or when it fails to generate positive results. Precarious lines of action also could result in personal demotions or reassignments, generate conflict that extends to locals, or creates political backlash within a society or globally.

According to the simplest, and most straightforward, economic model of decision-making, individuals will compare the anticipated rewards and costs to determine the expected level of outcome associated with a policy (see Equation 1, above). If the rewards outweigh the costs, then they will approve of that policy. If the net outcome is estimated to be negative, on the other hand, they will be dissatisfied and fail to give their approval. A number of social science models apply this basic framework to compare how rewards and costs predict various types of human behavior, including assessments of war strategies (e.g., [10]), altruistic actions [18], and aggression [8].

B. Including Expectations for Outcomes to Predict Policy Approval

Here we extend this basic cost/reward analysis further by incorporating the concept of personal (or group) expectations. We note that even in cases where outcomes are predicted to be positive, some people may not approve of a particular policy. Instead, we maintain that approval, or satisfaction with a line of action, depends not only on the predicted outcome, but also on the relative expectation level that is held by a person or group (see Equation 2, above). For example, some individuals may have very high expectations for specific outcomes, whereas those of others may be low. People with low expectations are easy to please. Those with expectations that are high, on the other hand, will not readily grant their approval and may only do so when the predicted rewards far outweigh the costs ([21], pg. 81-82). For instance, individuals tend to be more sexually satisfied in their romantic relationships if their levels of comparison are characterized by low rewards and high costs [15].

Personal expectations tend to develop out of past experiences. Individuals who faced a previous situation in which a similar policy was applied with deleterious results, for example, may apply a high standard of comparison and refuse to approve of a policy unless the current, anticipated outcome is unusually positive. Those who never experienced such a setback will be much more likely to express their agreement with the strategy. In a coalition environment, the expected level for outcomes also is likely to be dictated ahead of time by command, rather than left to the individual troops to determine. A coalition may be encouraged to pursue a series of strategies at all costs, for example, or to do so only in the case that a completely successful outcome is predicted. In such cases, the protocol, rather than individuals' experiences, largely determines the level of expectation used in shaping policy approval.

C. Including Alternatives to Predict Final, Policy Commitment

Approval of a policy, which is a function of rewards, costs, and expectations, is one factor that influences the likelihood of individuals making a commitment to a particular line of policy. It is not the only important factor, however. People can be approving and relatively satisfied with a policy, and yet still not commit to following through with it. Or they can be not particularly approving, but decide to go ahead with the policy anyway. Peoples ultimate level of commitment depends on two additional factors, we argue, those of $alternatives$ and $investments$.

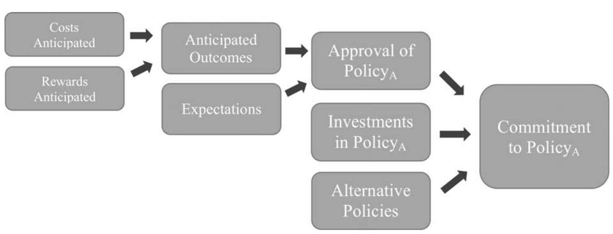
Standard rational choice approaches typically assume there is always an alternative to the status quo - and that the question is purely one of maximizing the utility between the options [12]. Alternatives, however, vary in number and attractiveness. Thus the presence or absence of alternatives, as well as their desirability (costs minus benefits), represent critical aspects of our approach.

In a given coalition scenario, typically there will be a variety of possible policies that could be recommended. The existence of alternative policies is apt to influence individuals levels of commitment. When people are aware of a promising alternative policy, or policies, they are expected to be less committed to the particular policy of interest. On the other hand, when faced with a situation in which there are relatively few, or no, viable alternatives to the proposed policy, then individuals are much more likely to commit themselves more fully to the option at hand [23], [19]. When they face limited alternatives, people may choose to commit to a policy that they do not view very favorably (e.g. *I voted for the "least bad" candidate*).

D. Including Investments to Predict Final, Policy Commitment

In addition, coalitions often invest in particular policies, where investments refer to dedicated resources that cannot be retrieved if the policy is not carried out (see Equation 3 above). Such investments can include specialized equipment, extensive training in unique skills, and the person hours dedicated to planning and strategy. Casualties incurred while enacting policies also can be considered a form of investment. When investments are extensive, high levels of commitment are more likely. Low levels of prior investment in a policy make it easier for individuals to reject it. See Figure 1 for a visual depiction of our model.

Figure 1. Model of Human Evaluation of Policy



E. Human Bias

One main source of human bias in decision-making involves so-called *sunk costs*, where sunk costs refer to investments that cannot be retrieved [2]. People can refuse to ignore prior investments, or so-called sunk costs, in situations where they should. In fact, research finds that individuals often escalate costly actions in cases where they have invested highly and with negative consequences [25], rather than treat these loses as sunk costs that cannot be recovered. Such an investment perspective encourages people to continue to pursue the same actions, despite the fact that doing so may be less effective or result in additional losses.

F. Summary

A coalition will pursue a decision, or a line of action, when the anticipated benefits exceed the costs, when the expected outcome is greater than the standard (e.g., protocol), when the alternative options are worse, and the investments into that line of action are high. Note that operationalization of these concepts is context specific. Particular situations and policies will determine the criteria and the manner in which the concepts of rewards, costs, and so on, are operationalized. The negotiations of mission policies and strategic policies provide context that we will model and to which we will apply elements of our particular variant of social exchange theory.

Theories based on the principles of social exchange are useful in accounting for a range of choices on the part of humans. They aide in explaining seemingly poor or sub-optimal decision-making on the part of humans. For example, the approach helps to explain situations in which individuals remain in extremely unrewarding situations and relationships, such as those that are violent (e.g., [22]). In such situations, individuals who choose to remain in low reward situations typically have relatively few rewarding alternatives, and they have invested highly in the current arrangement (i.e., shared resources that are difficult to divide). In addition, individuals who make the decision to stay in high cost situations, tend to exhibit low comparison levels [15]. That is, they have low expectations for their outcomes in the first place. Due to the costliness of past, similar experiences, their expectations regarding the current situation are relatively low.

Our model could be used to determine how humans would have made policy related decisions in a specific context. Those predictions can then be used to create policy rules that machines can use to implement in various situations. These algorithms can be used in the post-operation analysis model.

III. A HUMAN-AUTONOMOUS SYSTEM SCENARIO IN A COALITION ENVIRONMENT

Consider the hypothetical example of a UK patrol on a reconnaissance mission in an urban environment. Suspicious activity is detected together with signs of armed conflict and explosions. Although, the patrol could investigate the context in detail, the potential risk (and thus costs) are considered high. Instead, it would be advisable to collect preliminary aerial reconnaissance from an unmanned drone to mitigate the risk. A request is placed within a coalition environment. A US automated drone on return from its mission receives the request and needs decide whether to divert from its route to provide assistance. The decision is context-dependent and difficult to configure with declarative rules alone. Amongst others, it may depend on the amount of fuel left, the delays caused, the information carried by the drone that could not be sent yet, or even the risk of revealing the presence of the drone. Rules would therefore need to be specified for all combinations of these parameters and all arising circumstances, which are difficult to predict. A human decision and new tasking order can be requested, but this requires communication to be available and somebody with sufficient understanding at hand to make the decision. Even then, the revised/combined task may arrive too late.

Coalition policies dictate both the timely completion of the mission and the principle of providing assistance to allied forces when in need. $Rewards$ and $costs$ are associated with both alternative courses of action as are $expectations$ of the level of desired outcomes. The different $alternatives$ are the plans of actions that can be automatically derived from either of the two policies and $investments$ can be derived from the relative value of the resources involved. The elements of the model described in Section II are thus available for the drone to derive its commitment to the two policies and make an autonomous decision. More importantly, the decision made would, in essence, correspond to the rational choice a human would make given the same information. It is therefore explainable to the humans involved both at the base commanding the drone and to the soldiers on patrol.

The aerial reconnaissance reveals the presence of casualties in need of medical assistance, which in turn would require the reconnaissance patrol to divert from its mission objectives. Providing medical assistance may require supplies to be delivered through robotic mules, which in turn would require collaboration with both the aerial support and the ground patrol to determine the level of support needed and to navigate the urban environment at the site.

Achieving the best societal outcomes in a specific context requires constantly dealing with emerging circumstances. There is constant exchange of information and services and constant re-evaluation of the policy alternatives in each context. Whilst humans depend on the autonomous systems components of the team, such as drones and mules, these systems also rely on human support to achieve their outcomes. In each circumstance multiple and possibly conflicting policies apply. A common framework that enables both human and automated components to re-evaluate their $commitment$ to the policies provides the necessary flexibility to enable the collaboration while not requiring it to foresee all possible circumstances. Moreover, it enables both human and autonomous components of the system to explain their behavior and have expectations of each other's behavior when all the circumstances are known.

IV. Discussion

What constitutes *policy* when the model proposed in Section II is applied uniformly across both human and autonomous team members? Unlike a traditional system perspective where policies definitively determine the actions to be taken in response to individual events (e.g., [7]), policies for human-autonomous collaborations may describe broad desirable courses of actions and objectives. Non-compliance, or more simply, divergence, from the course of action prescribed occurs when circumstances change and alternative policies lead to better outcomes. Policies are also needed to determine the relative parameters of the model, i.e. how to compute rewards, costs, and investments and their values in given circumstances. Determining the value of providing assistance to wounded victims (civilians or coalition fighters) and the value of the equipment or the information it carries would require policy directives. Evaluating the $expectations$, i.e. the level of outcome associated with a particular policy requires learning from past history and a *policy analytics* framework to be deployed to measure and record outcomes associated with policies in context.

We have considered so far a perspective where each individual makes decisions from their own point of view based on the $comittment$ value they derive. In a military context there is a need to ensure the predictability of the team's behavior. Furthermore, the behavior of the autonomous components needs to be predictable to the human members of the team [17]. This may require disclosing the rewards, costs and expectations of the individuals in a given situation, though not necessarily all the factors considered in their evaluation. This stands in contrast with many existing systems perspectives which tend to exchange information about the context rather than about how it is perceived.

In addition to the human bias (see Section II-E), further differences between human and autonomous systems components must be considered. Amongst others, the policy $alternatives$ need to be explicit and complete for an autonomous system, whereas they are often implicit for humans. Autonomous components also require a precise calculation or estimate of the rewards and costs whereas, this is often implicit and approximate in human behavior. An autonomous system would typically maintain an exhaustive list of alternatives and use all the values in the calculation of the parameters. For humans, choosing between alternatives or outcomes is often an intuitive estimate and humans are often error prone in these evaluations. Accounting for these differences and their impact is also something that requires significant further work.

V. Conclusions

In summary, we develop an initial model of decision-making as it occurs when humans confront policy in teams of humans and autonomous systems. The model begins with a basic, cost-benefit analysis, but then extends beyond that to incorporate additional, key elements that likely shape rational, human assessments of policies. We note that whereas high level policies driving coalitions may not frequently change, policies driving specific actions may have to rapidly change to adapt to different contexts and situations, and our model would be particularly relevant in assessing reactions in these fluid environments. In addition, individuals may have to violate policies that are in place because of unexpected events/situations. Understanding how people react to such policy changes and how to handle violations are critical. Also it remains important to better identify which policies are more likely to be violated by individuals; for example, an individual may more easily violate a policy requiring a particular authorization for the use of some specialized equipment than a policy requiring that a forward operating base must always have two soldiers patrolling the area around the base. Understanding variations in a users perception of the various model parameters (rewards, costs, etc.) with respect to the same policy, but under different situations could be critical to better design and engineer policies, in particular for autonomous systems.

ACKNOWLEDGMENT

This research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence under Agreement Number W911NF-16-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

REFERENCES

[1] Akerlof, George. 1976. The Economics of Caste and of the Rat Race and Other Woeful Tales. Quarterly Journal of Economics 90: 599-617.

[2] Arkes, Hal R. and Catherine Blumer. 1985. The Psychology of Sunk Cost. Organizational Behavior and Human Decision Processes. 35: 124-140.

[3] Blau, Peter M. 1964. Exchange and Power in Social Life. New York: Wiley.

[4] Boutaba, Raouf and Aib, Issam. 2007. Policy-based Management: A Historical Perspective. J. Netw. Syst. Manage. 15, 4 (December 2007), 447-480.

[5] Burgess, M. 1995. Cfengine: a site configuration engine, USENIX Computing systems, Vol8, No. 3 1995.

[6] Cook, Karen S. and Eric Rice. 2003. Social Exchange Theory Pp. 53-76 in Handbook of Social Psychology, edited by John DeLamater. New York: Kluwer Plenum.

[7] Damianou, Nicodemos, et al. "The ponder policy specification language." Policies for Distributed Systems and Networks. Springer Berlin Heidelberg, 2001. 18-38.

[8] Faris, Robert and Diane Felmlee. 2014. Casualties of Social Combat: School Networks of Peer Victimization and their Consequences. American Sociological Review. 79, 228-257.

[9] Felmlee, Diane and Susan Sprecher. 2000. Close Relationships and Social Psychology: Intersections and Future Paths. Social Psychology Quarterly 63:365-376.

[10] Gartner, Scott S. 1997. Strategic Assessment in War. Yale University Press.

[11] Gartner, Scott S. 2008. The Multiple Effects of Casualties on Public Support for War: An Experimental Approach. American Political Science Review 102(1): 95-106.

[12] Green, Donald and Ian Shapiro. 1996. Pathologies of Rational Choice Theory: A Critique of Applications in Political Science. Yale University Press.

[13] Homans, George C. 1961. Social behavior: Its Elementary Forms. New York: Harcourt, Brace & World.

[14] Kephart J.O. and Walsh, W. E. "An artificial intelligence perspective on autonomic computing policies," Proceedings. Fifth IEEE International Workshop on Policies for Distributed Systems and Networks, 2004. POLICY 2004., 2004, pp. 3-12.

[15] Lawrance, Kelli-An, and E. Sandra Byers. 1995. Sexual Satisfaction in Long-Term Heterosexual Relationships: The Interpersonal Exchange Model of Sexual Satisfaction. Personal Relationships 2: 267-285.

[16] Moses, Tim. "Extensible access control markup language (xacml) version 2.0." Oasis Standard 200502 (2005).

[17] Muir, Bonnie M. 1987. Trust between humans and machines, and the design of decision aids, International Journal of Man-Machine Studies, Volume 27, Issue 5, 1987, Pages 527-539

[18] Piliavin, I.M., Rodin, J.A. and Piliavin, J. (1969) Good Samaritanism: An underground phenomenon? Journal of Personality and Social Psychology, 13, 289 -99.

[19] Sprecher, Susan. 2001. Equity and Social Exchange in Dating Couples: Associations with Satisfaction, Commitment, and Stability. Journal of Marriage and Family 63(3): 599-613.

[20] Steinbruner, John D. 1974. The Cybernetic Theory of Decision: New Dimensions of Political Analysis. Princeton, NJ: Princeton University Press.

[21] Thibaut, John and Harold Kelley. 1959. The Social Psychology of Groups. New York: Wiley.

[22] Rusbult, Caryl E. and John M. Martz. 1995. Remaining in an Abusive Relationship: An Investment Model Analysis of Nonvoluntary Dependence. Personality and Social Psychology Bulletin 21(6); 558-571.

[23] Rusbult, Caryl E. 1983. A Longitudinal Test of the Investment Model: The Development (and Deterioration) of Satisfaction and Commitment in Heterosexual Involvements. Journal of Personality and Social Psychology 45(1): 101-117.

[24] M. Sloman and E. Lupu, "Security and management policy specification," in IEEE Network, vol. 16, no. 2, pp. 10-19, Mar/Apr 2002. doi: 10.1109/65.993218

[25] Straw, Barry M. 1976. Knee-Deep in the Big Muddy: A Study of Escalating Commitment to a Chosen Course of Action. Organization Behavior and Human Performance 16:27-44.

[26] Dinesh C. Verma. 2000. Policy-Based Networking: Architecture and Algorithms. New Riders Publishing, Thousand Oaks, CA, USA.