Computational Models for Belief Revision, Group Decision-Making and Cultural Shifts

MURI ANNUAL PERFORMANCE REPORT 2008

AFOSR Grant No. A9550-05-1-0321 (Start Date: 1 May 05)
PRINCIPAL INVESTIGATOR: Professor Whitman Richards, 617-253-5776, wrichards@mit.edu
LEAD INSTITUTION: Massachusetts Institute of Technology
AFOSR PROGRAM MANAGER: Dr. Terence Lyons, 703-696-9542, terence.lyons@afosr.af.mil

NOTE: This performance report covers progress and changes that have occurred after the Base Period Review, namely 16 Dec 07 - 15 Dec 08.

PROGRAM OBJECTIVE: The primary objective of this MURI is (1) to explore how beliefs support and lead to certain actions in one culture but not another, and (2) to develop computational models that further our understanding of the relation between beliefs, decisions, and actions. A key requirement for such models is to distinguish the different roles played by sacred values versus instrumental or secular values, which differ widely across cultures. The goal of these models is to provide formal explanations for how the beliefs of individuals affect group and individual actions, and how groups evolve. Such models are an important step toward understanding and predicting the dynamics and actions of groups. They are fundamental to an understanding of how actions of a group may be altered by belief revision, by either internal or external pressures (including force). They also will be needed for strategic reasoning in negotiations, where beliefs in different cultures may lead to what appears to be irrational proposals, yet are seen as rational in that competitive culture. This MURI initiative supports DoD's need to understand how terrorist cells are created and how they evolve, thus providing models for the underlying seeds, which can be more easily identified early in the development process. In turn, the models may suggest countermeasures that will help revise or manipulate belief structures to reduce the likelihood of militant cells forming within a culture. The research is also very relevant to negotiations between states having different cultures.

CURRENT MURI CONSORTIUM RESEARCH TEAM MEMBERS *:

- **Massachusetts Institute of Technology** (PI) Whitman Richards, Prof. of Cognitive Science, Computer Science and Artificial Intelligence Laboratory (CSAIL), Joshua Tenenbaum, Assoc. Prof. of Computational Cognitive Science, Patrick Winston, Prof. of Computer Science. CSAIL
- **Harvard University,** Avi Pfeffer, Assoc. Prof. of Computer Science
- **University of Michigan,** Scott Page, Prof. of Economics & Complex Systems, Jenna Bednar, Assoc. Prof. of Political Science
- **Northwestern University,** Kenneth Forbus, Prof. of Computer Science, Douglas Medin, Prof. of Psychology
- **CUNY, John Jay Center for Terrorism & Univ. Mich.,** Scott Atran, Prof. of Anthropology, Psychology and Public Policy
- ***Georgia Tech,** Jeff Shamma,, Prof. Electrical & Computer Engineering
- ***North Carolina State,** Jon Doyle, Institute Prof. of Computer Science
- ***Georgetown,** John Mikhail, Assoc. Prof. Law & Philosophy (Ethics)

Consultants: Robert Axelrod, Prof. of Political Science and Public Policy, University of Michigan
Marc Sageman, MD, PhD; Sageman Consulting LLC; Univ. Penn. Adjunct Prof., Dept. Psychiatry;
Rajesh Kasturirangan, PhD., National Inst. Adv. Studies, Bangalore; J. J. Koenderink, Tech Univ. Delft (formerly Prof. of Physics, Helmholtz Institute, Utrecht)

***NOTE: The * indicates PI's added for the option years, replacing the departure of one original member of the MURI team, Brian Stankiewicz, Asst. Prof. of Psychology, Univ. TX, (now at 3M, Minneapolis.)**

FINANCIAL: (as of 1 Nov 08): 3-YEAR BASE PERIOD – FY05: **$669,604 (5 months)** FY06: **$1,139,405** FY07: **$1,183,188 & FY07: 71,181 (additional funding)** FY08: **$551,821 (7 months)** OPTION PERIOD - FY08: **$718,743 (5 months)** FY09 + **not yet authorized as of 15 Dec 08.**

To-Date (15 Dec 2008) Base Period funds of $2,844,000 have been expended and are "off the government's book." This figure takes into account invoices from the subawardees and consultants through the period 31 Oct 08, leaving an unexpended balance of about $700,000. The current commitments to the set of present PI's for the period 1 Nov 08 through 31 Oct 09 are estimated to be $1,976,000. We note that our unexpended funds were accumulated through the departure of Prof. Brian Stankiewicz in Jun 07, and the Stanford sabbaticals of Prof. Page and Prof. Bednar. We have used these funds to add three new PI's to the MURI: Prof. Jeff Shamma at Georgia Tech (multi-agent models and control in networked systems), Prof. John Mikhail at Georgetown (models for moral reasoning), and Prof. Jon Doyle at North Carolina State (the logic of belief systems.) In addition, Dr. Rajesh Kasturirangan has been added to provided further bridges to South Asian culture and violent and non-violent groups; Prof. Jan Koenderink is now engaged (with Richards) in multi-level cellular automata models for layered, heirarchical social interactions. (Part of this work will be presented at the AF review in Jan 09.) Collaborations with all of these new PI's have taken place, beginning 1 Sept 08. The projected expenses for these new PI's and consultants is about $400,000 for the remaining duration of the contract. An additional $150,000 is being allocated to fund three workshops, leaving a $150,000 reserve for new students or post-docs.

SCIENTIFIC APPROACH: As mentioned, a major goal of the MURI is to develop computational models that further our understanding of the relation between beliefs, decisions and actions in different cultures. Our aim is to provide predictive, rather than just descriptive models. For example, rather than simply describing the structure of a militant or non-violent network (Atran et al), we have made strides in understanding the processes that underlie their formation and evolution (Richards et al.) Understanding the origin and nature of shared and revised beliefs among group members is critical (Page & Bednar.) Unfortunately, data are scarce, and hence an important component of our effort is the reliable documentation of the formation of groups such as those engaged in the acts at Madrid, London, Bali, etc. The research led by Scott Atran under our MURI currently provides the most detailed and reliable set of data available, and this work continues with more details and studies of a dozen or so such networks. Additional data continue to be collected in non-violent settings (Medin et al.), such as the Wisconsin Menomenee, the Guatemalan Ladinos, and Amish and Muslim communities. These studies of non-violent communities allow closer examination of moral issues and sacred vs. secular values – the latter being critical in both belief revision and negotiations across cultures (see below). Because the origins of many beliefs rest in the traditions of a culture, we are formalizing how beliefs, actions and values are expressed in stories within a culture (Winston, Forbus). The framework for representing stories will support rapid, automatic analysis of the semantic interpretation of text, such as excerpts from news sources. Implications for future trends can be explored through analogies with traditional stories from the relevant culture. Finally, individuals or groups striving for social change need to have strategies for actions. We are developing and testing new frameworks for understanding strategic reasoning, subject to beliefs, which dictate in part which moves are considered rational (Pfeffer). These complement the more traditional "game-theoretic" approaches, and appear more plausible in many real world, multi-agent scenarios.

IMPORTANT TRNSITION EVENTS

- Atran et al continue to be sought after for briefings to various Government agencies. In 2008, Atran briefed the House Committee on Homeland Security, following up earlier briefings in '07 to the US State Dept, and House of Lords. Sageman has been a consultant to NYPD. Other briefings on militant networks were at NSC and NCTC. These briefings provided details of the evolution of several militant networks, including Madrid, Bali, & Leeds. They document the psychology of the players, the relations among the players, their motivations as well as the network structure. Many important generalizations emerge.
  (See http://www.edge.org/3rd_culture/Atran07/index.html.)
- Tenenbaum's framework and algorithms for inferring structural forms from relational data are now being developed by Navia Systems into a scalable platform for government and public use.
- Winston & Finlayson's Story Workbench platform for semi-automatic annotations of text is now under beta testing. Its Java wordnet interface subroutine JWI has been released: over 1000 downloads occurred in '08.
- Gal's (& Pfeffer) multi-agent Colored Trails platform for studying strategic reasoning is now used by six universities.

Other Scientific Accomplishments in '08
(please see last year's annual report for Option Period review for complete listings for previous years)

- Atran, Medin & Ginges have completed an important study showing negotiations that propose trading sacred for secular assets may create impediments to any further productive negotiations. Hence culturally-sensitive negotiations can not be treated in the same manner as one would by an economist concerned only with monetary assets. This work provides one of the first clear findings in a Middle East context. Resolution of quarrels arising from conflicting scared vales may require concessions that acknowledge the opposition's core concerns (Atran, Axelrod, Davis.)
- Atran & Axelrod have begun the development of an approach to framing (and reframing) of contentious values to promote negotiations in seemingly intractable conflicts and to understand competitive framing in pursuit of political goals.
- Studies of Hindu-Muslim conflict over sacred land in Northern India by Medin's group partially replicate their previous Middle East findings on how sacred and secular values do not intermix but also suggest that, in some contexts, sacred values are somewhat malleable. There is a hint that there may be different degrees of sacred values which link up with the saliency of group identity.
- Page & Bednar have modeled the dynamics of belief revision that would be typified by an individual wishing to join a select group. The model explains two puzzles in the theoretical study of culture: how meaningful cultural signatures can be derived, and how diversity can persist within a culture.
- A representation and model for small group evolution has been developed by Richards & Wormald. The dynamics of the model arises from a competitive interaction among three forces: leadership dominance, team bonding, and diversity. The group's vulnerability to alternative leadership emerges from the study.
- Beliefs underlie actions. Page and Golman have compared basins of attraction for different learning rules for collective actions. These models differ from most research that considers only the stability of equilibria and not their likelihood of attainment. With this newer approach to modeling, we can show that cultural learning and best response learning can have vanishing overlap in their basins of attraction.
- Tenenbaum's group continues to explore the relation between beliefs and actions using a Bayesian Framework. Here the emphasis is on modeling an individual's ability to understand another's plans for actions.
- Dehghani and Forbus have developed a concept map system that enables scientists to construct formal representations of human mental models. For example, using interview data gathered by Medin's group, the system used analogical generalization to classify individuals in different cultures based on their belief systems. Inspection of the nature of the generalizations led to new insights about properties of the groups, subsequently confirmed by manual coding and analysis of transcript data.
- Dehghani and Forbus continue to develop their computational model of moral decision-making. The different impact of sacred versus secular values is expressed via qualitative order of magnitude relationships. The model can derive connections between the choices of a scenario and the system's values either via first-principles reasoning or via analogy with stories involving moral reasoning.
- An important advance in understanding strategic play is Pfeffer's identification of four fundamental reasoning patterns in games to characterize the way information is used and manipulated. Pfeffer continues to explore how the reasoning patterns inform our understanding of negotiation games with conflicting interests. These ideas are also being incorporated into other scenarios, such as the multi-layer cellular automata representation being developed by Koenderink & Richards.
- Craig Joseph at Northwestern continues both theoretical work and empirical research on psychosocial adaptation in Muslim communities (see publications.)

Student and Postodoctoral support over portions of the base period include: MIT: 4 Phd, 2 MS Graduate Research Assistants and 1 PostDoc; UMich: 4 PhD; Harvard: 2 PhD and 2 PostDoc; John Jay: 2 PhD and 1 PostDoc; Northwestern: 6 PhD and 1 PostDoc.

WORKSHOPS

1. **Patterns of Influence (currently being organized by W. Richards, J. Shamma, & H. Peyton Young.)**
2. **Narrative & Culture (Winston, Finlayson, & Richards) Fall '09**
3. **Latent Feature Analysis & Structural Forms (optional)**
4. **Violent & Non-violent Group Formation (optional)**

PUBLICATIONS and PRESENTATIONS: approximately 20+ papers in refereed journals or refereed proceedings, about 15 working papers in progress, and over 20 presentations at national and international meetings.

Selected Publications (07 – 08 only)

Atran, S. Axelrod, R. & R. Davis (2007) Sacred Barriers to Conflict Resolution. Science 317, 1039-1040.

Atran, S. & R. Axelrod (2008) Reframing sacred values. Negotiation Jorl. 24, 221-246.

Atran, S. Magourik, J. & J. Ginges (2008) The radical madrassah factor. CTC Sentinel (Counter Terroriism Center, West Point) . http://www.ctc.usma.edu/sentinel/default.asp.

Atran, S. & D. Medin (2008) *The Native Mind and the Cultural Construction of Nature*. MIT Press, Cambridge

Bartels, D. M. & Medin, D. L. (in press). Are morally-motivated decision makers insensitive to the consequences of their choices? Psychological Science

Bednar, J. and Page, S. (2006) Can Game(s) Theory explain culture? The emergence of cultural behavior within Multiple games. (Rationality and Society, in press.)

Gal, Y. & A. Pfeffer (2008) Networks of Influence Diagrams: reasoning about agents'decision-making processes. Jrl of Art. Intell. Res. 33:109-147

Ginges J., Atran, S., Medin, D. & Shikaki, K. (2007) Sacred Bounds on Rational Resolution of Violent Political Conflict. PNAS 104, 7357

Haidt, J. & C. Joseph (2007) "The moral mind: how five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules." In Carruthers, Laurence & Stitch (eds). "The Innate Mind: Foundations and Future" Vol 3. Oxford Univ. Press.

Joseph, C. & B. Riedel (2008) Islamic Schools, assimilation, and the concept of Muslim Amereican character. In: Being andBelonging: Muslims in the United States ssince 1911" K. Ewing (ed). Russell Sage Fdn Press, NY

Kemp, C. & J. Tennenbaum (2008) The discovery of structural form. Proc. Nat. Acad. Sciences 105, 10687-10692.

Magourik, J., Atran, S. & M. Sageman (2008) Connecting terrorist networks. Studies in Conflict and Terrorism, 31, 1-16.

Page, S E., Sander, L. & C. Schneider-Mizell. (2007) Conformity and dissonance in generalized voter models. Jrl. Stat. Phys. 128, 1279 –1297.

Conference Proceedings (07 – 8 only)

Antos, D. & A. Pfeffer (2008) Identifying reasoning patterns in games. 23rd Conf on Uncertainty in AI.

Dehghani, M. Unsworth, S., Lovett, A. & K. Forbus (2007) Capturing and categorizing mental models of food webs using QCM. 21st International Workshop on Qualitative Reasoning Workshop, Aberystwyth, UK

Dehghani, M., Iliev, R. & S. Kaufmann (2007) Effectsd of fact mutability in the interpretation of counterfactuals. Proc. 29th Annual Conference of the Cognitive Science Society, Nashville, TN

Ficici, S.G. and Pfeffer, A. (2008) Modeling how Humans Reason about Others with Partial Information, Seventh International Conference on Autonomous Agents and Multiagent Systens (AAMAS).

Ficici, S.G. and Pfeffer, A. (2008) Simultaneously Modeling Humans' Preferences and their Beliefs about Others' Preferences, Seventh International Conference on Autonomous Agents and Multiagent Systens (AAMAS).

Forbus, K. , Klenk, M. & T. Hinrichs. (2008) Companion Cognitive Systems: design goals and some lessons learned. Workshop on "Naturally Inspired AI", Wash. DC Nov 08; Tech Report FS-08-06, AAAI Press., Menlo Park CA

Finlayson, M. A. (2008) Collecting Semantics in the Wild: the Story Workbench. Workshop on "Naturally Inspired AI", Wash. DC Nov 08; Tech Report FS-08-06, AAAI Press., Menlo Park CA

Forbus, K., Lockwood, K., Tomai, E., Dehghani, M., and Czyz, J. (2007). Machine Reading as a Cognitive Science Research Instrument. AAAI Spring Symposium on Machine Reading.

Gal, Y. & A. Pfeffer (2007) Modeling reciprocal behavior in human bilateral negotiation. National Conference on Artificial Intelligence (AAAI'07),Vancouver, BC.

Gal Y., Grosz B., Pfeffer A., Shieber S. (2007) The Influence of Task Contexts on the Decision-making of Humans and Computers Proceedings of the Sixth International Interdisciplinary Conference on Modeling and Using Context (CONTEXT'07), Roskilde University, Denmark.

Gal, Y, Kasturirangan, R, Pfeffer, A. and W. Richards (2008) How tacit knowledge guides action. AAAI Workshop on "Naturally Inspired AI", Wash. DC Nov 08; Tech Report FS-08-06, AAAI Press., Menlo Park CA

Joseph, C. (2008) Islam, Virtual Theory and the Cognitive Science of Ethics. Ann. Mtg. Amer. Acad, of Religion, Chicao Il (1-3Nov08.)

Kemp, C. Goodman, N. D. & J. Tenenbaum (2007) Learning causal schemata. Proc. 29th Annual Conference of the Cognitive Science Society.

Pfeffer A., Gal Y., (2007) On Reasoning Processes in Games National Conference on Artificial Intelligence (AAAI'07),Vancouver, BC.

Richards, W. (2008) Modal Inference. AAAI Workshop on "Naturally Inspired AI", Wash. DC Nov 08; Tech Report FS-08-06, AAAI Press., Menlo Park CA

Tomai, E. & K. Forbus (2007) Plenty of blame to go around: a qualitative approach to attribution of moral responsibility. In Proc. Qualitative Reasoning Workshop, Aberystwyth, UK

Tomai, E. & K. Forbus (2007) Narrative presentation and meaning. AAAI Fall symposium on Intelligent Narrataive Technologies, Washington D.C.

For a complete listing of publications, see: http://groups.csail.mit.edu/belief-dynamics/