November 2007

Profiling is Politically 'Correct': Agent-Based Modeling of Ethno-Political Conflict

Barry G. Silverman
University of Pennsylvania, basil@seas.upenn.edu

Gnana K. Bharathy
University of Pennsylvania

Benjamin Nye
University of Pennsylvania

Postprint version. To be Presented at *Interservice/Industry Training, Simulation and Education Conference 2007*, November 2007, 11 pages.

Profiling is Politically 'Correct': Agent-Based Modeling of Ethno-Political Conflict

Abstract

A holy grail for military, diplomatic, and intelligence analysis is a valid set of software agent models that act as the desired ethno-political factions so that one can test the effects that may arise from alternative courses of action in different lands. This article enumerates the challenges of such a testbed and describes best-of-breed leader and follower profiling models implemented to improve the realism and validity of the agent. Realistic, 'descriptive' agents are contrasted to rational actor theory in terms of the different equilibria one would expect to emerge in conflict games. These predictions are examined in two real world cases (Iraq and SE Asia) where the agent models are subjected to validity tests and a policy experiment is then run. We conclude by arguing that substantial effort on game realism, best-of-breed social science models, and agent validation efforts is essential if analytic experiments are to effectively explore conflicts and alternative ways to influence outcomes. Such efforts are likely to improve behavioral game theory as well.

Keywords

political simulation, agent-based models, personality profiling, policy analysis tools

Comments

Postprint version. To be Presented at *Interservice/Industry Training, Simulation and Education Conference 2007*, November 2007, 11 pages.

Profiling is Politically 'Correct': Agent-Based Modeling of Ethno-Political Conflict

Barry G. Silverman, Gnana K. Bharathy, Benjamin Nye
Electrical and Systems Engineering Department
University of Pennsylvania, Philadelphia, PA 19104-6315
basil@seas.upenn.edu, bharathy@seas.upenn.edu, benjamid@seas.upenn.edu

ABSTRACT

A holy grail for military, diplomatic, and intelligence analysis is a valid set of software agent models that act as the desired ethno-political factions so that one can test the effects that may arise from alternative courses of action in different lands. This article enumerates the challenges of such a testbed and describes best-of-breed leader and follower profiling models implemented to improve the realism and validity of the agent. Realistic, 'descriptive' agents are contrasted to rational actor theory in terms of the different equilibria one would expect to emerge in conflict games. These predictions are examined in two real world cases (Iraq and SE Asia) where the agent models are subjected to validity tests and a policy experiment is then run. We conclude by arguing that substantial effort on game realism, best-of-breed social science models, and agent validation efforts is essential if analytic experiments are to effectively explore conflicts and alternative ways to influence outcomes. Such efforts are likely to improve behavioral game theory as well.

Keywords: political simulation, agent-based models; personality profiling; policy analysis tools

ABOUT THE AUTHORS

Barry G. Silverman is Professor of Electrical and Systems Engineering at the University of Pennsylvania where he is also Director of the Ackoff Collaboratory for Advancement of the Systems Approach (ACASA). He holds the BSE ('75), MSE ('77) and PhD (also '77) all from the University of Pennsylvania, is a Fellow of IEEE, AAAS, and the Washington Acad. of Science, and sits on the board of several organizations and journals in the intelligent systems fields. The focus of his research has largely been on aesthetic and cognitive engineering of embedded game-theoretic agents that can help humans improve their learning, performance, and systems thinking in task-environments. Over the years, his lab has produced or is in the process of creating an agent-based model of mind-body duality; patient training games and human physiology simulations; a terrorist campaign and crowd simulator; numerous autonomous and emergent agent tools; several distributed, computer-mediated, human-to-human collaborative systems; 3 role playing games (RPGs); and the AESOP interactive fiction game generator. As a result of all this work, Barry is also the author of over 130 articles, 12 books/proceedings, over 100 technical reports, 7 copyrighted software systems, a boardgame, and several research and teaching excellence awards.

Gnana K. Bharathy's formal academic training has been in the areas of Engineering (process/environmental and information systems), Risk Analysis, and Systems Science and he completed his doctoral work in Systems Engineering. During the course of his dissertation work, Gnana has developed a systems methodology for integrating social system frameworks and modeling human behavior through knowledge engineering based process, and has employed the same to create several models of leaders and followers in situations involving conflict-cooperation. His dissertation has recently been awarded the INCOSE-Stevens award for promising research in systems engineering and integration. Gnana also received the Wharton Risk Management and Decision Process Center's Ackoff award (2005) for carrying out research on Human Decision Processes.

Benjamin Nye is currently pursuing his doctorate in Electrical and Systems Engineering, researching alongside Barry Silverman. Current research work he is involved in focuses on simulating decision making in sets of hierarchal groups, working towards automated exploration of the state space. In addition to a B.S. in Computer Engineering, he also has a significant background in psychology.

Profiling is Politically 'Correct': Agent-Based Modeling of Ethno-Political Conflict

Barry G. Silverman, Gnana K. Bharathy, Benjamin Nye
Electrical and Systems Engineering Department
University of Pennsylvania, Philadelphia, PA 19104-6315
basil@seas.upenn.edu, bharathy@seas.upenn.edu, benjamid@seas.upenn.edu

INTRODUCTION AND PURPOSE

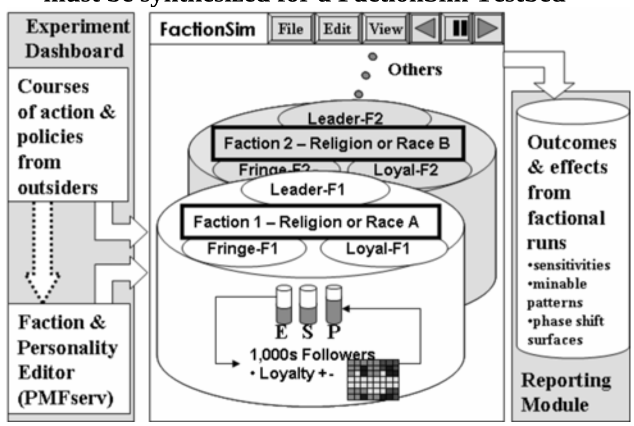
Analytic game theory is the mathematics of strategy, and as such, holds great promise for helping to understand conflicts. At the same time, analytic game theory has a weak record of explaining and/or predicting real world conflict – about the same as random chance according to Armstrong (2002), Green (2002). In the field of economics, Camerer (1999) points out that the explanatory and predictive powers of analytic game theory are being improved by replacing prescriptions from rational economics with descriptions from the psychology of monetary judgment and decision making. This has resulted in 'behavioral game theory' which adds in emotions, heuristics, and so on. In this paper, we pursue the same approach and believe the term 'behavioral game theory' is broad enough to cover all areas of social science, not just economics.

Specifically, the military, diplomatic, and intelligence analysis community would like for (behavioral) game theory to satisfy an expanding range of scenario simulation concerns. Their interest goes beyond mission-oriented military behaviors, to also include simulations of the effects that an array of alternative diplomatic, intelligence, military, and economic (DIME) actions might have upon the political, military, economic, social, informational (psyops), and infrastructure (PMESII) dimensions of a foreign region. The goal is to understand factional tensions and issues, how to prevent and end conflicts, and to examine alternative ways to influence and possibly shape outcomes for the collective good.

PROFILING FACTIONS AND THE FACTIONSIM TESTBED

Our exploration begins by constructing a testbed (FactionSim) that facilitates the codification of alternative theories of factional interaction and the evaluation of policy alternatives. FactionSim is a tool where you set up a conflict scenario in which the factional leader and follower agents all run autonomously. You are the sole human interacting and using a set of DIME actions to influence outcomes and PMESII effects (see Figure 1).

Figure 1 –Models and Components that must be synthesized for a FactionSim Testbed



Factions are modeled as in the center of Figure 1 where each has a leader, two sub-faction leaders (loyal and fringe), a set of starting resources (Economy, E, Security, S, and Politics, P), and a representative set of over 1,000 follower agents. A leader is assumed to manage his faction's E- and S- tanks so as to appeal to his followers and to each of the other tribes or factions he wants in his alliance. Each of the leaders of those factions, however, will similarly manage their own E and S assets in trying to keep their sub-factions and memberships happy. Followers determine the level of the P-tank by voting their membership level (see Sect. 3.2). A high P-tank means that there are more members to recruit for security missions and/or to train and deploy in economic ventures. So leaders often find it difficult to move to alignments and positions that are very far from the motivations of their memberships.

Despite efforts at simplicity, stochastic simulation models for domains such as this rapidly become complex. The strategy space for each leader facing only two other leaders is in the trillions of options, a number impossibly large to explore. As a result, FactionSim's Experiment Dashboard (left side of

Fig.1) permits inputs ranging from one course of action to a set of parameter experiments the player is curious about. On the bottom left is the profile editor of the personalities for the leaders and sub-leaders, and of the key parameters that define the starting conditions of each of the factions and sub-factions. Certain actions by the player that are thought to alter the starting attitudes or behavior of the factions can flow between these two components – e.g., a discussion beforehand that might alter the attitudes of certain key leaders (Note: this action is often attempted in settings with real SMEs and diplomats playing our games).

Game Analysis

FactionSim runs a set of multiple games, G = {G1, G2, …, Gn} simultaneously. Within a faction one may observe games between rival leaders, between leaders and followers, and follower on follower. The across-faction games include attempts to cooperate and/or compete with other factions' leaders and followers, and/or attempts to contain factions aimed at your own downfall. For discussion's sake, consider these as iterated semi-cooperative games (ISCGs). This game formulation is the simplest game one can analyze involving conflicts between (and within) factions. Using it helps to clarify many of the key elements of these conflicts.

Let us next consider how FactionSim's games might be treated by two types of ISCG agents, namely:

Rational Actors: Presumed normative and devoid of psychic concepts as in post-WW II economic theory and intro game theory classes - perfectly informed, purely logical, constant discount rate (i), and motivated by self-interest to maximize their material payoffs. All actors have identical payoff functions where they compute R{E|S|P}as uncontested resources, Q as resources at stake, and CstA as the cost of actions. The expected payoff is thus: Payoff = R + Q – CstA. Mutual conflict or fight-fight is a well-known Nash equilibrium. We know also that if CxCy> FxCy, then mutual cooperation is Pareto optimal and in repeated games, if the agent histories are remembered, no agent is excessively powerful, and agents start with mutual cooperation, then the following is the well-known mixed strategy that will prevail: attack if provoked (tit-for-tat) to deter other leaders from taking advantage, but otherwise cooperate. The subgame perfect equilibrium consists of long periods of cooperation punctuated by occasional conflicts. Ignoring rare conflicts, one may write the predicted payoffs for any given 'rational' agent in alliance with others as:

$$\text{PAYOFFx} \;= \sum\nolimits_{t=0}^{T} CxCy(t)/(1+i)^{t} \qquad (1)$$

Descriptive agents: Swedberg (2001, p.325) states "If sociological game theory is not to end up as an artificial exercise, it is absolutely essential that the beliefs, ideas and experiences of the actors themselves are moved onto center stage". One must profile the individuals involved to find out the inventory of items at stake and to build realistic agent models. We do this with best-of-breed social science instruments (Sect 3). Such actors use these approaches to decide everything from R and Q, to the size of an action, to how to discount (i), to how much they are willing to pay for their gambits (CstA), etc. – one wouldn't even expect to use the same formulas for normative vs. descriptive computations. Aside from material payoffs, these agents attend to moralistic issues driven by their emotional value (emV) and how relationships (ΔKxy) change; they may commit errors and use biased heuristics; and they may see games through a different lens (e.g., settling grievances, fast track to next life). For these agents, the payoff function becomes Payoff = R + Q – CstA - | ΔK | + emV.

Researchers like Macy & Flache (2004) show that as one alters agent aspirations (something equivalent to emV and ΔK), the stable equilibrium [1] collapses and the prediction of fight-fight becomes near-continual. Woods (2004) shows that divisibility of Q as well as emV and ΔK type issues must be elaborated if conflicts are to be settled. Results like this mean that one must reduce the guess work about what drives the resource disputes, moral dilemmas, and social relationship grievances.

PERSONALITY PROFILING MODELS

Profiling of personalities has not yet reached the stage of a mature science with first principles; however, there are best-of-breed profiling instruments with respectable field trials and high inter-rater reliability. These are useful for creating agent frameworks with greater degrees of realism. Such implementations, if done carefully, may in fact improve profiling science. Hendrickson and McKelvey (2002) suggest that social science theories, in general, need to be computationally formalized as agent models to show they are analytically adequate. These models in turn must be subjected to correspondence tests against real world phenomena to verify them (ontologic adequacy). This two step testing process improves the science by revealing the agenda for advancement.

Unlike the evolutionary tradition where personas are mutated, this approach of profiling real personalities within connectionist agent models allows one to watch the generative mechanism and to observe what they do, how they learn and adapt, and what macro-behavior emerges from the actors' micro-decisions. Using profiling instruments reduces the dimensionality to the traits and factors they require, and where these are applied, we can use training datasets, fill in the traits and factors of archetypical as well as real characters, conduct validation tests, and treat these parameters as no longer independent variables clouding the larger political analyses – they exist within encapsulated components and only their inter-relationships to other parts are significant when assessing the whole. This is no different than systems engineering for any complicated device. A crash test of an automobile does not depend on how the pistons fire. We similarly encapsulate other parts of the faction model – e.g., the (E|S|P) resource tanks that we currently model as stacks of poker chips that grow or fall. One can plug in finer resolution models for any given tank without affecting overall system performance. With that in mind, we turn now to the best-of-breed profiling theories we implemented as leader and follower models.

Profiling Leaders

In FactionSim, each leader and follower is modeled within a framework known as PMFserv (Silverman 2005) where the leader's cultural values and personality traits are represented through Goal, Standards and Preference (GSP) trees. These are multi-attribute value structures where each tree node is weighted with Bayesian probabilities or importance weights. A Preference Tree is one's long term desires for world situations and relations (e.g., no weapons of mass destruction, stop global warming, etc.) that may or may not be achieved in the scope of a scenario. In FactionSim agents this translates into a weighted hierarchy of territories and constituencies (e.g., no tokens of leader X in resource Y of territory Z). The Standards Tree defines the methods a leader is willing to take to attain his/her preferences, and what code that others should live by as well. Finally, the Goal Tree holds short term needs the agent seeks to satisfy each turn (e.g., vulnerability avoidance, power, rest, etc.). The GSP tree is a value model editor that allows one to (a) implement leader and follower profile instruments as nodes on the trees and (b) set the weights on the nodes which in turn implements a personality profile (see Figure 2).

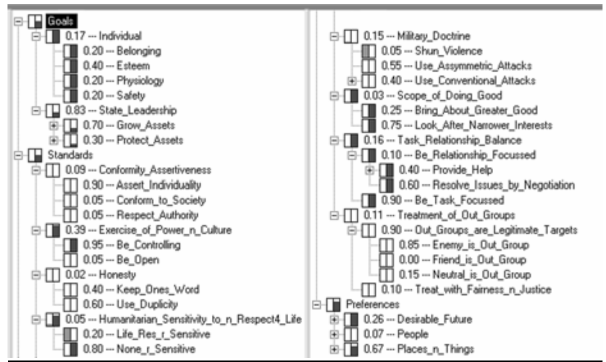


Figure 2 – GSP Tree Structure, Weights, and Activations

Perhaps the best leader profiling instrument is Hermann (1999) who offers a descriptive theory of leader style that is measurable and can be fully implemented in this framework. After two decades of studying over 122 national leaders including presidents, prime minister, kings, and dictators, Hermann uncovered a set of leadership styles that appear to influence how leaders interact with constituents, advisers, or other leaders. Hermann determined that seven traits are particularly useful in assessing leadership style: need for power, control, openness (combines 2 traits), task vs. relationship focus, distrust, and in-group bias.

In Hermann's profiling method, each trait is assessed through content analysis of leaders' interview responses as well as or other secondary sources of information. Hermann's research also has developed methods to assess leadership at a distance, based mostly on the public statements of leaders. Hermann (1999) has developed mean scores on each of the seven traits. A leader is considered to have high score on a trait, if he or she is one standard deviation above the average score for all leaders on that trait.

In order to implement the Hermann instrument as an agent model (GSP Trees), we had to do the following:

- Need to increase power (and its inverse, protection) is both a long term Preference as well as a short term Goal. So it appears on both trees. In the Machiavellian and Hermann-profiled world of leaders, the goal tree reduces to a duality of growing vs. protecting the resources in one's constituency. Expressing goals in terms of power and vulnerability provide a high-fidelity means of evaluating the short-term consequences of actions.
- Most of the other Hermann traits govern personal and cultural norms and thus appear on the Standards tree.

- The UN GLOBE model of leaders (House, 2004) includes several traits like Hermann's but also adds Scope of Doing and Sensitivity to Life (humanitarianism) which we adopt here as well. We also add one further trait, namely Protocol vs. Substance Focus as a continuum to describe the leader's penchant for protocols (e.g., state visits or speech acts such as religious blessings) as opposed to taking any concrete actions.
- Resource management doctrine – We add specific standards that capture the doctrine a leader adheres to when considering his Economic and Security tanks. Beneath each subnode that has a + sign, there are further subnodes, but under the G- and P-trees these are just each faction's resources in each territory with valence and importance valuated weights.

The structure of the GSP trees is shared by all agents. However, the tree weights are unique for each agent and thus capture individual differences that may be determined by culture, ideology, or personality. Other papers discuss how the weights may be derived so as to increase credibility: e.g., see Bharathy (2006), Silverman (2002a,b, 2006b). An example of the weights is the insurgent leader shown in Figure 2. He is low on conformity, humanitarianism, scope of doing good, and treating outgroups with fairness, and high on exercise of power, and asymmetric warfare.

GSP trees are used by the agent for all decisions – e.g., selecting a next game action, determining faction alliance moves, or deciding on a speech act. They give each agent a robust and individual worldview. When contemplating a decision, the agent calculates the subjective expected utility (SEU) it expects to derive from every action available to it, as constrained by perception and chooses the alternative that maximizes SEU. Thus

$$\textbf{Best Response (SEU)} = \mathbf{Max}\{\Sigma\, \mathbf{U}(\mathbf{a_k})^*\mathbf{Pr}(\mathbf{a_k})^*\, \Phi(\mathbf{r_j}) + \psi\}$$

where utilities (U) for next actions, a_k, are the anticipated E|S|P tank gains or losses the actions afford combined with how those affect the nodes of a given agent's GSP trees. $\Phi(r_j)$ is a function that captures the strength of positive and negative relationships one has with agent or object j that are effected or spared by a_k, and ψ handles merging and discounting (decay) prior GSP activations. Probabilities assess the likelihood of success or failure. Also, the GSP tree weights adhere to principles of probability; e.g., all child node insights add to unity beneath a given parent, activations and weights are multiplied up a branch, and no child has multiple parents (independence). Although we use fixed weights on the GSP trees, the succeed and fail reservoirs on each node (see Fig.2) serve to render them dynamic and adaptive to the agent's current needs. Thus, when a given success reservoir is filled, until ψ decays it, that tends to nullify the importance of the weight on that node (or amplify it if the failure reservoir is filled). In this fashion, one can think of a form of spreading activation (and deactivation) across the GSP structure as a game proceeds.

Profiling Followers

We introduce three refinements in order to also be able to model the values and motivations of followers – (1) additions to the GSP trees, (2) a group-affinity profiling instrument, and (3) group transfer dynamics (exit, voice, and loyalty). In keeping with analytic adequacy, each of these refinements is an implementation of a well-respected model drawn from the social sciences. In terms of the Goal tree changes, the leader goals are still there but may be zeroed out if this is strictly a follower, or may be left in at some degree of importance if this is a mid-level leader. For followers in general, where day-to-day existence is a struggle, the first four of Maslow (1987)'s hierarchy of needs is considered a useful representation of the range of short term goals that a person might have to be concerned about. Without regard to order of achievement, these are added to the G-tree under the node labeled 'individual' (see Figure 2). Each of these nodes are activated by lower level branches on the tree pertaining to physiology, to relationship dynamics, or to the affinity instrument described below. In terms of the Standards Tree, we add Conformity Assertiveness as a way to capture Hofstede(2003)'s Power-Distance and Individualism factors (respect authority, conform to society) and the GLOBE study's Assertiveness factor.

For determining an individual's group affinity, one needs an instrument that measures it. The instrument that we have adopted here involves Eidelson and Eidelson (2003) who have developed a five belief ("dangerous ideas") framework for better understanding the psychology of individual-group dynamics particularly relevant to conflict settings. These beliefs are considered particularly important influences on a group member's perceptions of his/her group's current circumstances and future prospects. Eidelson worked with us to help us implement his model within the GSP tree framework. Thus an agent who is profiled within the GSP trees can be seen to be harboring "Dangerous Ideas" to the extent that the five beliefs are present as follows:

•*Vulnerability* (V) Revolves around a sense of living in harm's way amid constant threat and peril.

Computed as the inabilitiy of agent i to make progress toward its objectives (i.e., G- and P-tree successes minus failures).
•*Injustice (I).* Perception of being a victim of mistreatment by specific others or by the world at large. Computed as the amount of agent i's G-tree failures attributable to others who violate agent i's S-tree, minus G-tree successes that others cause.
•*Distrust (D).* The presumed hostility and malicious intent of other individuals or other groups, computed as the amount to which agent i's S-tree is violated by others.
•*Superiority (S).* Conviction of being better than others—morally superior, chosen, entitled. Computed as agent i's perception of disutility or consequences to other agents and groups it dislikes of agent i's G-tree achievements.
•*Helplessness (H).* Refers to perceived inability to influence or control events and outcomes; self-perpetuating because it diminishes motivation. Computed as the world's disutility: ie, sum of all GSP tree failures.

With this framework, depending on the perceiver, two agents may view the same group as superior or inferior, as suffering grave injustices or as exaggerating minor slights, as helpless or capable of effective action, and so on. The agents compute the possibilities. Thus if a viewer sees a group as vulnerable and he doesn't want to be vulnerable, there will be negative activations for remaining loyal to this group and following its action choice policies. Conversely less negative (and possibly positive) activations will be afforded for the member who exits away from this group perceived as vulnerable.

Mathematically, the reader may recall $\Phi(r_{ij})$ from the prior section. Here we examine the case where j is a group (or leader) A and the term refers to the membership, relationship, or strength of affinity of agent i to group A. An agent i can belong to multiple groups at varying strength according to:

$$\mathbf{\Phi(r_{iA}) = (Superiority_A \ x\ GSPcongruence_{iA}) / VID_{Ai}}$$

where, Superiority and VID (=V+I+D) are from Eidelson instruments if available, else derived by GSP trees of agent i in reacting to leader or group A.

Groups are characterized by GSP weights for the average of all members as well as by property lists defined a priori (religion, political system, corruption, maturity, etc.), and salience factors. GSP congruence is estimated using the sum of the means square differences in the GSP nodes. GSPcongruence = 1 – Sqrt[Sum[(wi1 –wi2)^2], which is the correlation of the weights between two GSP trees. If an agent is in Group B, it will not be drawn to a Group C whose GSP archetype is substantially incongruent to its own. If an agent is in a group (or under control of a leader) whose average GSP is greatly different from its own, the agents tend to reduce membership (P-tank contributions) and use Voice to resist the leader or attempt to Exit to another group.

If agent i desires to exit from any group A to join any C, this is governed by the delta in utility of membership in each group plus a cost factor adjusted for transfer rate or demand elasticity. If the delta is positive, or larger than some loyalty factor, exit may occur. Let, this delta be:

$$\mathbf{\Delta\Phi_j = [\ (U(\Phi_C) + COST_{TR}\ ) / TR_{AC}\ \ ] - U(\Phi_A)}$$

where,
$U(\Phi)$ = utility of membership, found from GSP trees
$COST_{TR}$ = cost of migration, land costs, and lost opportunity costs
TR_{AC} =Transfer Rate or group porosity, a measure of ease of entry to or exit from group A to group C,

$$TR_{A\to C} = Salience_{ExitA} \times Salience_{EnterC} \times GSPcongruence_{iC}$$

TR varies between (0,-1) and grows larger as porosity grows. Salience is the extent to which a group permits exiting by ingroup members, and entry by outgroup members. It is the porosity permitted by the group. There is a tuple or value pair that gives both salienceForEntry and salienceForExit. The demand elasticity for exiting a group is 1/TR.

The followers compute their grievance state for each faction, ranging from –4 to +4 an use this to determine their actions. As an example, suppose an agent identifies himself with Faction B, but lives under the rule of Faction A. The top state (GS+4) is total support of a given Group, say A. A faction getting a mid-point grievance scale (GS-0) means that agent is undecided and/or helpless to resist what A wants. At the other extreme of GS-4, the Faction B agent who lives under Faction or Leader A has already joined a resistance faction C working against A. At the extremes on either end, the agent will submit to militaristic commands of the leader of that group, while at the next lower level it will be only willing to go to protests, and verbally and economically support the activities of that group's leaders. Of course, exit from A and joining of C is governed by TR, and a several tick waiting interval to be sure the agent doesn't change its stance.

EXPERIMENTS

The previous sections synthesized social science theories pertinent to faction conflict and implemented them as agent models. Here we present two real world conflicts. Each begins with various validity tests to see if the agents correspond to the real world and if the underlying mechanisms seem adequately represented. A key policy parameter is then systematically varied in each experiment to learn the elasticity of conflict to that course of action.

Elasticity of Conflict in Iraq Due to Outside Support

During the spring 2006, five student teams assembled a total of 21 PMFserv leader profiles across 7 real world factions so that each faction had a leader and two sub-faction leaders. The seven factions – government (2 versions - CentralGov and LoclGov), Shia (2 tribes), Sunnis, Kurds, and Insurgents – could be deployed in different combinations for different scenarios or vignettes. The leader and group profiles were assembled from strictly open source material and followed a rigorous methodology for collecting evidence, weighing evidence, considering competing and incomplete evidence, tuning the GSP trees, and testing against sample datasets [see for example Bharathy (2006)]. A well-known fundamentalist Shia leader's GSP tree was shown in earlier Figure 2.

Validation testing of these models was run at one of the military commands for 2 weeks in May 2006. They assembled 15 SMEs across areas of military, diplomatic, intel, and systems expertise. Within each vignette the SMEs attempted dozens of courses of action across the spectrum of possibilities (rewards, threats, etc.). One interesting COA is reflected in earlier Figure 1 by the vertical arrow on the left of the chart linking the player to the Personality Editor. That is, a popular COA of the diplomats was to 'sit down' with some of the persuadable leaders and have a strong talk with them. This was simulated by the senior diplomat adjusting that leader's personality weights (e.g., scope of doing good, treatment of outgroups, etc.) to be what he thought might occur after a call from President Bush or some other influential leader. The SME team playing the MNC presented their opinions at the end of each vignette. The feedback indicated that the leader and factional models corresponded with SME knowledge of their real-life counterparts. They accepted the profiling approach as best in class and invited us onto the team for the follow on.

Here we show an illustrative policy experiment on 4 factions initially organized into two weak alliances (dyads): (i) CentralGov trying to be secular and democratic with a Shia tribe squarely in their alliance but also trying to embrace all tribes, (ii) a Shia tribe that initially starts in the CentralGov's dyad but has fundamentalist tendencies, (iii) a secular Sunni tribe that mildly resents CentralGov but does not include revengists, and (iv) Insurgents with an Arab leader trying to attract Sunnis and block Shia control. Each faction has a leader with two rival sub-leaders (loyal and fringe) and followers as in Figure 1 – all 12 are named individuals, many are known in the US. This is a setup that should mimic some of the factional behaviors going on in Iraq, although there are dozens of political factions there in actuality. Figure 3 summarizes the outcomes of three sample runs (mean of 100 trials each) over a 2 year window. The vertical axis indicates the normalized fraction of the sum across all security tanks in these factions, and thus the strip chart indicates the portion of the sum that belongs to each faction. Rises and dips correspond either to recruiting and/or battle outcomes between groups. The independent variable is how much outside support is reaching the two protagonists – CentralGov and Insurgents. When CentralGov and Insurgents are externally supported (3A), CentralGov aids the Shia militia economically while the Shia battle the Insurgents. Fighting continues throughout the 2 year run. A take-away lesson of this run seems to be that democracy needs major and continuous outside help, as well as luck in battle outcomes and some goodwill from tribes for it to take root. When only the Insurgents are supported (3B), the CentralGov is crippled by Insurgent attacks and civil war prevails. When the borders are fully closed and no group receives outside support (3C), the insurgency ultimately fails, but the CentralGov becomes entirely reliant upon the Shia group for military strength- a puppet government. These runs suggest the elasticity of conflict with respect to outside support is positive, and with no interference, the country seems able to right itself, although we in the West might not like the outcome. Of course these runs only include 4 of the many factions one could set up and run, plus due to page limits, we only displayed the effects of actions upon the Security Tank, and not other resources of the factions.

Earlier we asserted the ΔK and emV terms of payoff are vital to game theoretic formulations. A human behavior model such as PMFserv is able to help the analyst to generate and understand these, though we omit the strip charts of changing ΔK and emV strengths due to page limits. As contrasted with normative agents, factions in our runs fight almost constantly and are more likely to attack groups with which they have negative relationships and strong emotions. Relationship and emotions also factor into the formation of alliances. For example, across all runs, CentralGov has a friendly relationship towards the Shia, who are moderately positive back. This leads to CentralGov giving aid to the Shia and consistently forming an ally. Likewise the Sunni Secular have slight positive feelings towards the Insurgents, and are more likely to assist them, unless others are more powerful. Finally, some action choices seem to have purely emotional payoffs. For example, from an economic perspective, the payoff from attacking an enemy with zero economy is zero - a wasted turn. Yet in run 2c, when the Insurgents fail, the Shia still occasionally attack them simply because the Insurgents are their enemy. This seems to be a case where emotional are at least as important as material payoffs.

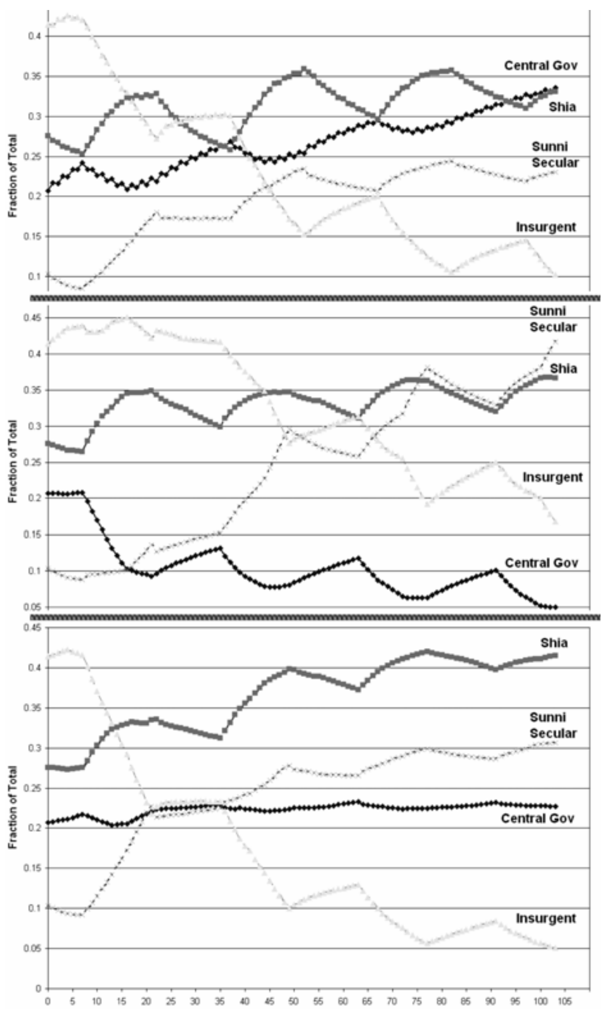


Figure 3 – Military Power of Iraqi Factions Under Alternate DIME actions (mean of 100 runs).

Impact of Leader Action on Follower Choices: FactionSim for SE Asia

While the previous section focused on leader profiles, this Section adds the profiles of followers and the decisions that result. Without naming the actual country or leader, it has a Buddhist majority and government. During the 1990s, the country was relatively stable, however, in the last few years, the rural provinces have seen a rise of Muslim separatist sentiment. The Buddhist Leader branded the separatists as bandits, and sent police from the north to handle protesters in the Muslim provinces. We obtained a database of the events with civilian injury and death and classified the incidents based on the size and intensity of the incident. The incidents were aggregated and plotted against time. The data was then longitudinally separated into 'independent sets' with a training set consisting of Jan-June 2004 while the test set began July 2004 and ran till Dec 2004 ending just before the tsunami.

Table 1 - Muslim Faction Shifting from Relatively Cooperative (GS0-2) to Largely Fighting (GS3 & GS4)

(2 years, average of 30 repeat trials)

Starting State (Avg of Weeks 1 & 2) Muslim Population at Start Is Neutral with Few Grievances Registering		End State (Avg. Weeks 103 & 104)
GrievanceState0 - Neutral	30	6%
GrievanceState1 - Disagree	55	1%
GrievanceState2 - Join Oppost	15	37%
GrievanceState3 - Nonviolent	0	39%
GrievanceState4 - Fight-Rebel	0	17%
TOTAL	100	100%

The training dataset was used to profile and tune the following types of agents for PMFserv (Eidelson was contracted to profile the Muslims in his instrument):

- Buddhist Group-Leader (structure of his GSP trees are in Fig 2) - data indicates harsh, cruel, task, corrupt, wealthy, successful. Sends worst behaving cops down to provinces, never discourages brutality.
- Muslim Group – Rural villagers lead by a local sultan agent and sorted into two archetypical groupings:
- Loyal Muslims - Despite lack of cultural freedom, muslim schools, etc. they are law-abiding, rural family members who want some autonomy.
- Fringe Muslims – tend to be sons of Moderates who were Wahhabi and college-trained, now

unemployed, running religious schools in family homes.

- <u>Separatist Group</u> – not profiled, just a place-holder, initially with no members. Scripted to periodically attack Buddhists.

In PMfserv, we profiled the Buddhist leader, and 161 Muslim agents including 1 Sultan, 80 loyals, and 80 fringe agents. We interoperated these with a cellular automata that is known as the Civil Violence (CV) model (Epstein et al., 2001). CV involves two categories of actors, namely 1,200 villagers and a variable number of cops. Muslim villagers and may be actively rebellious or not, depending on their grievances. 'Cops' are the security tank forces of the Buddhist Leader, who seek out and arrest actively rebellious agents. The main purpose of introducing CV is to provide a social network for the cognitively detailed PMFserv villagers to interact with. The Buddhist Leader examines the state of the world and makes action decisions to assist or suppress villagers (e.g., pay for Buddhist schools, add more cops, reduce cop brutality, etc.). The 161 PMFserv agents then assess their view of the world, react to how cops handle protester events, how their GSPs are being satisfied or not by leader actions, and to their emotional construals. The grievance level, leader legitimacy, and group membership decisions by 160 archetypical villagers in PMFServ are passed to 160 agents they control in the cellular automata. These agents influence neighbors in the population who spread news and form their own view of the situation. The number of Civil Violence villagers in each level of grievance (neutral through Fight Back as shown in the rows of Table 1) are added up and this information is passed back to PMFserv to help determine its starting level of grievance for the next cycle of reactions to <u>Buddhist Leader's</u> actions. The left side of Table 1 shows the starting values as percent of Muslim agents that occupy each Grievance State. By the end of the run, the right side of Table 1 shows the emergence of a majority of the population resisting and fighting (non-violent as well as violent). Specifically, it shows what percent of the population has been shifted from Neutral Grievance to higher states (recall the scale of earlier Section 3): GS0 (neutral) through GS4 (fight back).

We compare this simulated grievance over time to a proxy consisting of the incident severity – a weighted average of actual fatalities and injuries, where injuries are simply counted (w=1), but the weight on fatalities is 100. This was computed from the test dataset mentioned earlier (not the training data). To conduct the comparison, we applied the non-parametric Kendall's Tau measure of correlation. With a two sided test, considering the possibility of concordance or discordance (akin to positive or negative correlation), we can conclude that there is a statistically significant lack of dependence between base case simulated grievance and observed incident severity rankings at a confidence interval of 88%. In sum, the null hypothesis is rejected and real (test interval) incident data and simulation results are correlated.

As to the leadership, in the test dataset, the real Buddhist leader made 52 decisions affecting the population. We sorted these into positive, neutral, and negative actions. In the simulated world, the PMFserv clone made 56 action decisions in this same interval. At this level of classification (positive, neutral, negative), the mutual entropy (M) statistic between the real and simulated agents was less than 0.05, indicating significant correlation between real and simulated agent choices. This high degree of correlation exists only for the aggregate summary of positive, neutral, negative actions. If we try to correlate the precise negative action chosen or when it occurs, the correlations deteriorate rapidly.

An interesting experiment is to see how this outcome is affected by altering the Buddhist leader's policies. To do so, we can alter his personality (e.g., outgroup are targets, sensitivity to life, scope of doing good, etc.) by 15% in either direction. Reducing these is equivalent to what the SMEs in the Iraqi case study attempted when they had Bush call and try to convince a given leader to be more tolerant. Raising these up by 15% is what might happen if he grew more autocratic. Since the Leader's attributes lead directly to shifts in his course of action selections, these three versions of the leader were run to set up a range of potential futures for the followers. Thus, the y-axis of Figure 4 shows increasing losses of civil rights or the Inverse Quality of Citizenship (InvQtyCitizenship) as measured by the Muslim Group's calculated grievances (or VID). The x-axis shows the decision of these agents to retain membership in the Buddhist-lead Government (these are the members of GS0, GS1, and GS2). Agents who leave and join the separatists (GS3 and GS4) are not appearing in this plot. The plot thus shows that as long as conditions are not too intolerable, the entire population cooperates and remains. As conditions worsen, more and more agents exit and membership shrinks. This is what Hirshman (1970) refers to as the demand curve for civil rights. In FactionSim, we are able to fit the following linear regression to this demand curve with an R-square of 0.79

InvQltyCitizenship = 1.35- 0.83 Membership_as_ Fraction

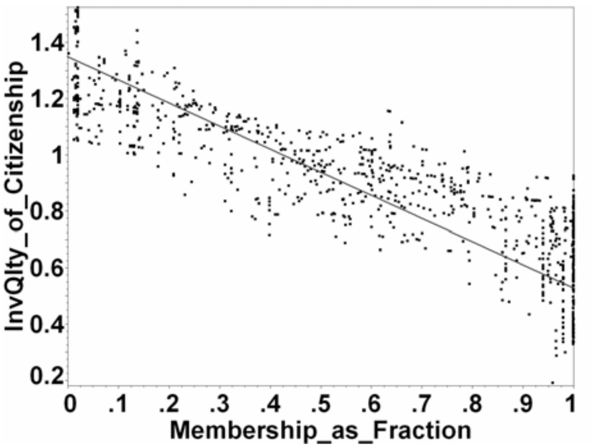


Figure 4 – Derived Demand Curve for Civil Rights by Faction Y's Followers

The Buddhist leader's ingroup bias, financial wealth, narrow scope of helping only his own faction to the north, and willingness to use violent repression seem to combine in the real world (and in our model of him) and make him unable to comprehend this new reality. In the summer of 2005, he had to impose martial law on these provinces to try and quell the separatist movement. In the summer of 2006, with the approval of the monarch, a military junta removed him from power due to his mismanagement of this situation and economic issues.

LESSONS LEARNED AND NEXT STEPS

The primary argument against rational game theory is its poor track record of prediction in matters of real world conflict primarily because it often simplifies the game and agents to the point that they bear little resemblance to the real world. Behavioral Game Theory seeks to overcome this dilemma, however, there is a lack of first principles in the social sciences for modeling agents. The field offers best-of-breed theories and instruments, but these exist in reductive silos (specialties) and few of these have computational implementations. To pass the test of analytic adequacy, we needed to map a number of social science models into a common utility-theoretic formalism and integrate them together into a faction modeling framework – e.g., Hermann, Hofstede, UN Globe, Eidelson, Hirshman, and so on. This synthesis was necessary (and successful) in order to have agent models capable of playing the roles of real world leaders and followers in rival factions. Such an integration, however, exposes the edges of the individual models and identifies an agenda for synthetic research. Did we include/exclude the right set of profiling factors? Does our implementation preserve the original intent? Did we fill in the gaps properly? Can the merged trait set be revalidated? We summarize working with one of these scientists and found him positive about such an agenda, and confirmed that other scientists would be similarly open to a synthetic agenda. Despite having worked reductively to reach their position of prominence, it seems that synthesis is a path forward that many in the field are willing to embrace.

In terms of ontologic adequacy of the current faction model synthesis, this research has tried to explore its robustness and cross-sample fitness. It is worth dwelling a bit on the benefits that were observed and the lessons learned from the case studies of this paper. For one thing, the descriptive agents passed validity assessment tests in both conflict scenarios attempted—our current day Iraqi leader agents were passed after extensive subject matter expert evaluation and the SE Asia leader and followers passed separate correspondence tests (correlations of over 79% on average). Validity is a difficult thing to claim, and one can always devise new tests. A strong test, however, is the out-of-sample tests that these agents also passed. Thus the SE Asian leader and followers were trained on different data than they were tested against. Further, the complete structure of the model of the leaders was originally derived in earlier studies of the ancient Crusades (Silverman et al. 2005) and this was transferred to the SE Asian and Iraqi domains. The only thing updated was the values of the weights for GSP trees and various other group relations and membership parameters – derived from open sources. So the structure of the leader model also survived and passed two out-of-sample tests relative to the Crusades dataset.

This article concludes with two experiments, one for assessing the elasticity of conflict in Iraq with respect to outside support, and the other for determining rate of radicalization of the population (and its inverse, demand for civil rights). These experiments explain what has and is being observed, and hence illustrate the promise of descriptive agents for extending game theory. The title of this article may be a bit of an overstatement, but using profiling as the basis of descriptive agent models does seem to be the correct approach for sociological game theory. 'Correctness' is more about the generative mechanisms inside the agents than whether any given predictions are accurate. If the generative mechanisms are roughly 'correct', one can have trust that experiments on these agents will yield useful insights about the alternative policies that influence them. Both experiments presented here pass initial adequacy tests and illustrate how analysts might use agent tools to explore policy alternatives and to identify parameter sensitivities and robustness. We also hope the results illustrate this as a promising research direction with further research steps warranted.

ACKNOWLEDGEMENT

This research was partially supported by AFOSR, DARPA, and the Beck Fund, though no one except the authors is responsible for any statements or errors in this manuscript.

REFERENCES

Armstrong, J.S. (2002). Assessing game theory, role playing and unaided judgment. *International Journal of Forecasting, 18*, 345-352.

Camerer, C. (1999). Behavioral Economics: Reunifying Psychology and Economics, *PNAS, 96*, pp.10575-577

Eidelson, R.J., Eidelson, J.I., (2003). Dangerous ideas: Five beliefs that propel groups toward conflict, *American Psychologist, 58*, 182-192.

Epstein, J. (2002, May). Modeling Civil Violence: An Agent-Based Computational Approach, *PNAS, 99:3*, pp. 7243-7250

Green, K. C. (2002). Forecasting decisions in conflict situations: a comparison of game theory, role playing and unaided judgment. *International Journal of Forecasting, 18*, 321–344.

Hendrickson, L., McKelvey, B., (2002, May), Foundations of 'new' social science, *PNAS, 99:3*, pp.7288-7295.

Hermann, M. G. (1999). *Assessing leadership style*. Hilliard, OH: Social Science Automation, Inc.

Hirschman, A.O. (1970). *Exit, voice, and loyalty*. Cambridge, MA: Harvard University Press.

Hofstede, G. (2003). *Culture's consequences: comparing values, behaviors, institutions, and organizations across nations.* Newbury Park, CA: Sage Publications.

.

House, R.J., Hanges, P.J., Javidan, M, et al., (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies*, Thousand Oaks, CA: Sage Publications

Macy, M.W., Flache, A., (2002, May 14). Learning dynamics in social dilemmas," PNAS, 99, Sup. 3, 7229-7236

Maslow, A.H., (1987). *Motivation and personality*, NY: Addison-Wesley.

Silverman, B.G., Bharathy, G. (2005). Modeling the Personality & Cognition of Leaders, in 14th *Conference on Behavioral Representations In Modeling and Simulation, SISO, May.*

Silverman, B. G., Johns, M., Cornwell, J., & O'Brien, K. (2006). Human Behavior Models for Agents in Simulators and Games: Part I – Enabling Science with PMFserv. *Presence*, 15: 2, April.

Silverman, B. G., O'Brien, K., Cornwell, J. (2006b). Human Behavior Models for Agents in Simulators and Games: Part II – Gamebot Engineering with PMFserv. *Presence*, 15: 2, April.

Silverman, BG, Bharathy, G., Nye, B (2007a). Gaming and Simulating EthnoPolitical Conflicts in Proc. *Descartes Conference on Mathematical Modeling for Counter-Terrorism (DCMMC)*, NYC: Springer.

Swedberg, R. (2001). Sociology and Game Theory, *J. Theory and Sociology*, 30:3, 301-335.

Wood, E. J. (2003). Distributional settlements and civil war resolution: stakes, expectations, and optimal agreements, under revision for re-submission to *The Journal of Conflict Resolution.*