MODELING THE CO-EVOLUTION OF STATES AND NATIONS

L.-E. CEDERMAN, Harvard University*

ABSTRACT**

It can be hypothesized that wherever states manage to assimilate their peripheries prior to nationalism, nationalist transformations proceed comparatively peacefully. In cases where the cultural penetration is weak, however, the political and cultural maps clash, thus producing tensions that drive national secession, unification, and irredentism. To trace these processes, I propose an agent-based model that embeds nationalist mobilization in a dynamic state system. My preliminary findings confirm the main hypothesis: conflict was found to vary negatively with state-framed cultural centralization. Yet some of the high-assimilation cases feature extreme levels of conflict due to nationalist unification's undermining effect on the balance of power.

INTRODUCTION

This paper employs agent-based modeling as way to cut through the complex interplay between power and culture in world politics. As opposed to standard modeling strategies, which usually treat states and nations as fixed units, the computational framework allows me to endogenize both actor types as distinct, though co-evolving, units. My primary goal is to investigate the effect of specific causal mechanisms on macro-outcomes. More precisely, I investigate the hypothesis that relates nationalist violence to a mismatch between cultural and political borders: wherever cultural penetration manages to centralize political rule prior to nationalism, the transition to the nationalist era is likely to proceed relatively smoothly. By contrast, where states lack the infrastructural capacity to standardize culture, violent transitions involving secession and unification become more likely.

Although still subject to further sensitivity analysis, my preliminary findings confirm the hypothesis: as the strength of state-led cultural assimilation increases, the transition to nationalist politics usually becomes smoother and less violent. In situations characterized by little coupling of cultural evolution and state-formation, however, the geopolitical and cultural maps clash, thus triggering conflictual outcomes as state borders and national communities adjust to the underlying ethnic landscape. But this relationship is not deterministic. In some cases, strong state-framed assimilation removes only partially smaller-scale ethnic cleavages. In other cases, it

---

[*] *Corresponding author address:* Lars-Erik Cederman, John M. Olin Institute for Strategic Studies, Harvard University, 1737 Cambridge Street, 411, Cambridge, MA 02138; phone: 617-495-2616; fax: 617-495-1384; e-mail: cederman@cfia.harvard.edu (Department of Government, Harvard University, from July 1, 2001).

[**] An earlier, more extensive version of this paper was prepared for presentation at the Annual Convention of the American Political Science Association, Washington DC, August 31-September 3, 2000. The current version owes much to the excellent computational advice of Laszlo Gulyas and Nick Collier. I am also very grateful for the generous support of the Robert Schuman Centre for Advanced Studies, Florence.

provokes unification that undermines the preexisting balance of power, thus paving the way for potentially catastrophic turmoil.

I rely on a new implementation of my previous computational research that focused primarily on state-formation (Cederman 1997). On the theoretical side, the extended model goes well beyond the earlier version by endogenizing both states and nations. At the heart of the new model, there is a geopolitical engine that generates emergent state systems through competitive pressures, including internal and external combat, resulting in conquest and secession. Within this environment, a multidimensional ethnic "landscape" is embedded. Based on cultural identities, nations form and take on behavioral importance for the development of the system's structure. These consequences include national secession from multinational states, national voluntary unification into culturally homogenous polities, and irredentism, involving both processes at the same time.

Methodologically, the new geocultural model also represents significant progress over earlier modeling solutions. In contrast to the latter, which was programmed in Pascal, the current framework is based on RePast, a comprehensive, Java-based simulation package developed at the University of Chicago.

The paper is organized in the three main sections. The first section provides an overview of the model. Then follows a section presenting a sample run with weak assimilation. The third section presents a contrasting run with strong state-formation. Finally, a concluding section discusses the heuristic insights gained from this exercise. It should be emphasized that this paper offers only a flavor of the model. A full technical description together with replication results can be found in Cederman (2000).

OVERVIEW OF THE MODEL

While this study focuses on studying the consequences of nationalism, it is first necessary to create both a cultural landscape and a geopolitical environment in which national identities can develop. Thus, before modeling the repercussions of nationalist transformations, I start by modeling the ethnic backdrop of state-formation as in pre-modern Europe.[1] I assume that the system, while far from perfect as a model of the pre-modern environment, initially consists of a large number of independent actors, which is subsequently followed by a phase of state formation. This process produces a balance-of-power equilibrium that prevents a limited number of states from conquering each other. During this pre-nationalist phase, state-led cultural assimilation takes place, but nationalist activity does not occur until nationalist mobilization is triggered. Assuming that a geopolitical equilibrium has been reached, the era of nationalism then begins. At this point, culture starts to matter for political mobilization. Given suitable conditions, nations will start to appear. What follows is a co-evolutionary process that links nation-building with changes in the state system.

Before turning to the model's behavior, it is useful to summarize the three main phases of the model graphically together with the interaction effects operating in each period (see

---

[1] To avoid conceptual confusion, it is important to keep apart the definitions of states and nations. For the present analytical purposes, it is sufficient to rely on Max Weber's (1946) classical distinction between states as sovereign organizations exercising legitimate control over a bounded territory and nations as cultural communities that strive to possess their own state.

Figure 1). All simulation runs start with a pre-modern period, which sees the creation of the initial grid of states. This is also when the age of migrations creates cultural landscape. After this

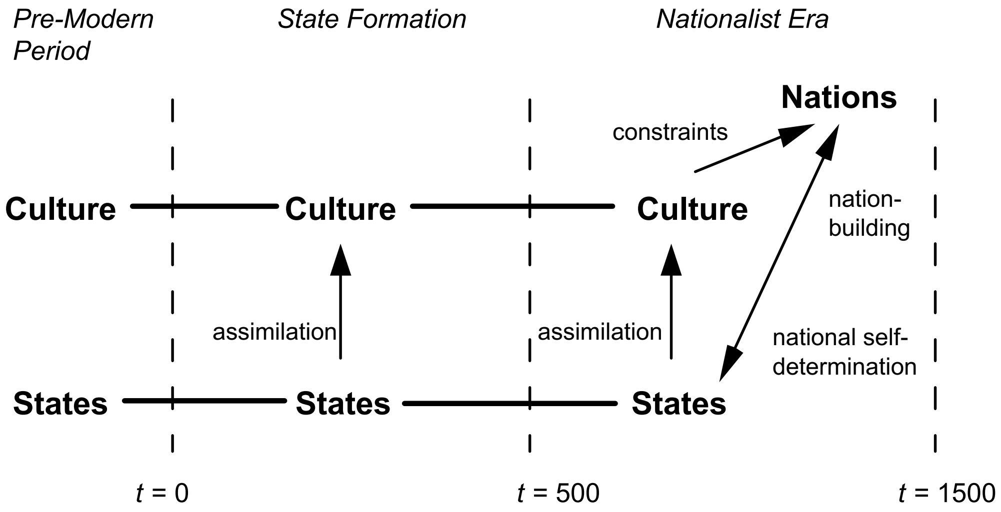


FIGURE 1 The Three-Phase Logic of the Model

stage, the era of state-formation unfolds. Most of the action at this point concerns interactions among states. To the extent that assimilation takes place, it affects the cultural identity of the states (and provinces). In the opposite causal direction, there is no impact at all, since culture is assumed to matter only once nationalism appears after time step 500.

From this point, the causal picture becomes more involved. Under nationalism, capitals still assimilate the culture of their provinces, but, in addition, a third type of entity emerges, namely nations. Their emergence can be attributed to political decisions on the part of states and provinces. Within the constraints of the cultural landscape, they decide either to create new nations or to join already existing national communities. Unlike the early modern period, culture starts to matter for political action, but only indirectly, through national identities. National self-determinations implies that national affiliations profoundly affect the behavior of the states. As illustrated by the sample runs, culturally “unhappy” national minorities resort to national secession, split nations come together, and states take irredentist action. This bi-directional link between states and nations constitutes the co-evolutionary nexus that is the main focus of our attention.

The present framework is based on an extended and stylized version of my previous “emergent polarity” model, developed in Cederman (1997, Chap. 4). As in that model, the world initially consists of a grid of unitary states (for earlier frameworks of this type, see Bremer and Mihalka 1977; Cusack and Stoll 1990). To give room for cultural evolution, however, I use a 15 × 15 grid for all runs rather than a smaller grid size (see Figure 2). The figure reveals the square arrangement of the initial system’s 225 statelets. The black boundaries mark the states’ borders and the square dots represent “capitals.”

From the very beginning, all actors acquire cultural identities and retain these whether they are sovereign or not. The shading in Figure 1 denotes the cultural differences among the local sites. While the darker squares correspond to the cultural border areas characterized by great differences, the brighter ones refer to “plains” in the cultural landscape. Each actor has a string of eight symbols that is assigned in a simulated “age of migrations” prior to the first time period (cf. Axelrod 1997, Chap. 7). This setup phase creates a synthetic cultural map that can be arbitrarily tuned.

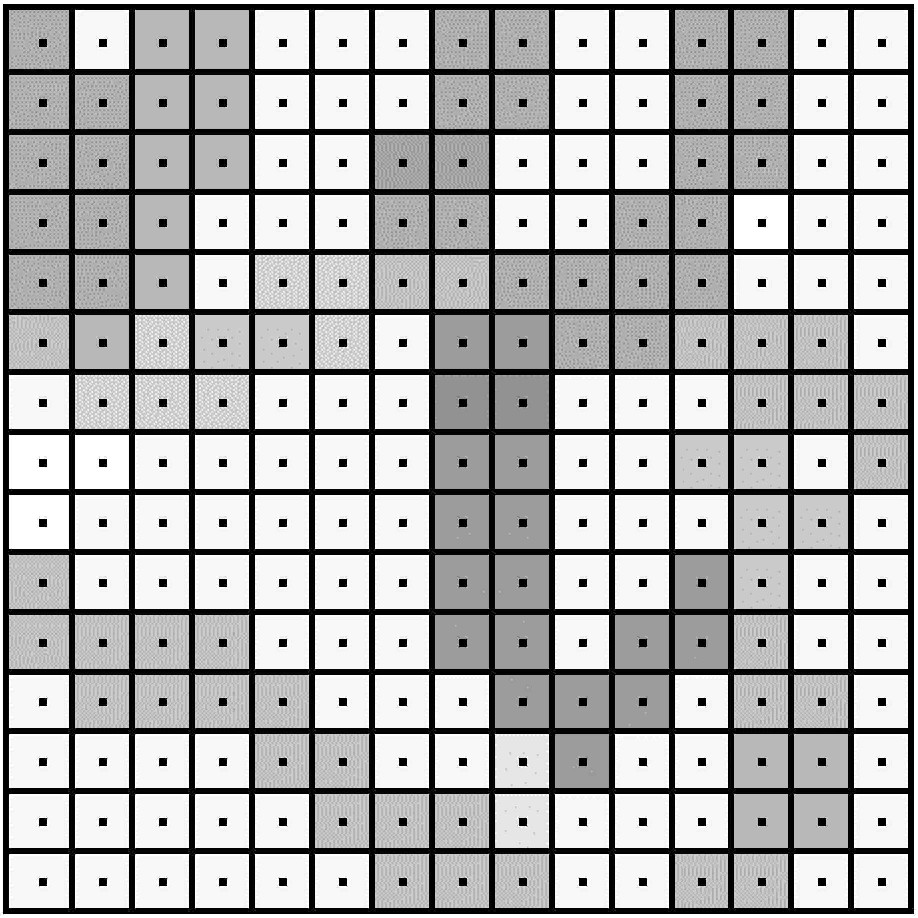


FIGURE 2 The Initial State of the Pre-Nationalist World

Two parameters drive this process: the number of “ethnies” that participated in the cultural settling of the landscape and the “smoothness” defining the cultural similarity within each tribe’s area of population. In our example, ten randomly located ethnic groups fill the whole grid with a trait-by-trait similarity of 0.90. This process guarantees that there will be both dialectal nuances as one moves from province to province as well as more abrupt ethnic cleavages.

Once unleashed, these actors will start interacting locally with their immediate territorial neighbors. Interaction can be of two types: either the actors peacefully exist side by side or they engage in combat. Conflict is initiated according to a simple decision rule, which depends on the local power balance between any two states. Roughly speaking, the actors play a “grim trigger” strategy with another. This means that they normally reciprocate whatever their neighbors did to them on their respective fronts. But whenever the power balance in their favor exceeds a preset threshold that governs the offense dominance of the system, they launch unprovoked attacks. The results reported in this paper are based on a fairly sharp threshold of 2.2 with some added noise.[2] The latter feature makes the decisions less than fully deterministic. In essence, this rule implies that all states consider attacking adjacent states once they get a little more than twice as powerful as any of them.

Once conflict erupts between two states, neither of them will resume cooperation until the battle is over. Combat ends with the victory of either party or, in some cases, with a stalemate. As in the decision rule for unprovoked attacks, the criterion for victory is a stochastic one with the same threshold at 2.2, though much more noise is involved here than in the decision function. In other words, the actors start to consider offensive behavior once their superiority becomes slightly more than a factor of two. Normally, the actor can count on winning, but because of the Clausewitzian “fog of war,” this may take a few iterations during which unpredictable things may happen. Most importantly, the attacking state may weaken its other fronts so much that opportunistic neighbors launch new offensives against the original aggressor.

Victorious parties gain the local territory fought over in any specific battle. If conquest destroys the supply lines from the capital under attack to its provinces, those parts that have been cut off regain independence. Should the capital itself be invaded, the entire state collapses. Thus, in offense-dominated systems, conquering states can grow quite quickly. Figure 3 illustrates the situation at time step 57 for the particular system depicted in Figure 1. At this point, the system’s polarity has been drastically reduced. Larger states with emergent boundaries have started to appear, though there are still a few smaller territories left waiting to be absorbed.

The figure also illustrates that action sometimes goes beyond interstate warfare, because in addition to such “horizontal” exchanges, the model features “vertical” two-level action. Graphically, rebellions of this type are marked with crosses, whereas small lines sticking into the neighboring sites denote interstate attacks. Even after their losing sovereignty, all provinces retain some resources and may thus rebel against the capital. This is likely to happen if the capital gets involved in too many other fronts, as illustrated by the state slightly northwest of the grid center. This actor is in deep trouble since it is under both internal and external attack.

The internal fronts are treated very much as any interstate combat theater: the same rules of deterrence and combat apply, with the exception that capitals have no reason to trigger conflicts with their peripheries. Should a rebellion occur, however, they always respond. While victorious rebellions are rewarded with secession, failures have no other consequences than the province remaining inside the state (see Cederman 1997, Chap. 5).

Shortly after the current time-step in the sample system, the neighboring actors around the collapsed state carve up the post-imperial remainders, and the system settles with seven surviving states (see Figure 4). In the following, these states will be referred to by letters ranging from A to G.

Despite the noise built into the decision rules, this is a rather robust equilibrium because even after 1500 time periods, it still remains intact. This is unsurprising given that the current model relies on a very simple territorial rule of resource updating. In fact, the resource

---

[2] This value was chosen since it generates stable geopolitical with moderate polarity equilibria quickly.

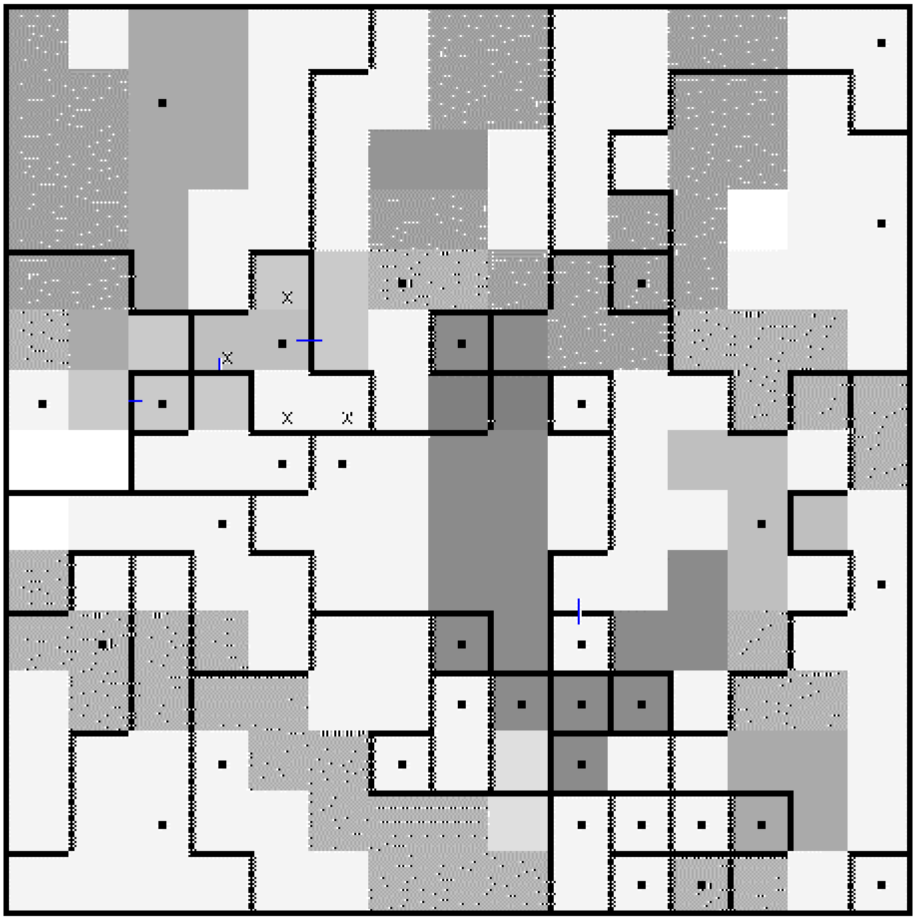


FIGURE 3 Secession Attempts and Interstate Combat at Time 57

"metabolism" of the system depends directly on the extent of the states' territorial holdings.[3] Each square in the grid yields a fixed amount of resources. A distance-dependent "tax rate" yielding, at most, 70% defines the power balance between the capitals and their conquered provinces.

THE NATIONALIST PHASE WITH WEAK NATION-BUILDING

Once having created a stable geopolitical environment without ruling out the possibility of changes, we are ready to introduce nationalism. Rather than treating nationalism as a constant "law," I model the phenomenon as a macro-historical process that exogenously hits an area at a specific point in history. To simplify things, I assume that this point occurs after 500 steps. The pre-national stage of the model was calibrated to generate a reasonably stable multipolar equilibrium well before this point.

---

[3] This mechanisms represents a simplification of the resource scheme used in Cederman (1997, Chaps. 4, 5). In that framework, a stochastic "harvest" was allocated to each state in each round. Rather than including accumulated resources, the current system lets a state's resource level be a direct function of the territory it controls.

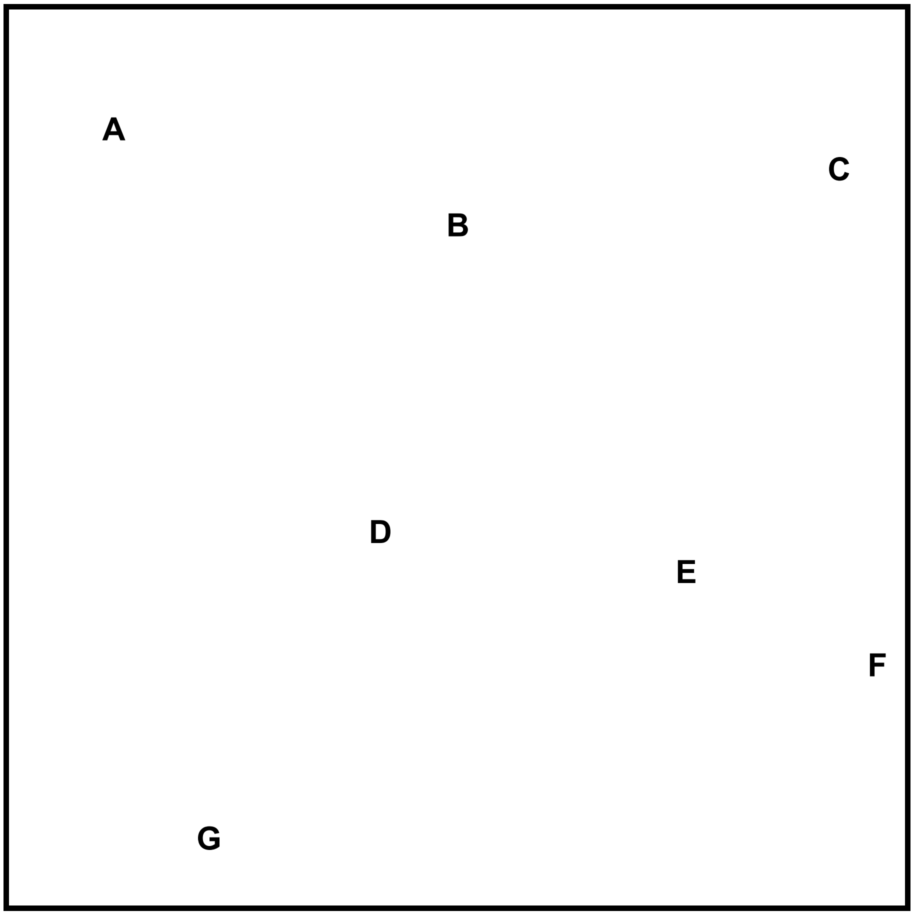


FIGURE 4 The Geopolitical Sample System in Equilibrium at Time 500

Nations are represented as distinct groups to which any state or province can belong if the nation's identity fits the actor's own culture. While both capitals and provinces are eligible members of nations, the probability of launching a nationalist movement depends crucially on the geopolitical status of the territory in question. Thanks to their resources, capitals have a much higher likelihood of founding their own nations, but provinces may sometimes create nationalist platforms in opposition to their respective capitals.

Figure 5 illustrates national mobilization in our sample system. At time period 536, precisely 36 steps after the onset of nationalism, three nations have formed, each one marked with a number surrounded by gray (green) boundaries. Of the seven capitals, three have their own nations. The most interesting development relates to nation 1, which started to develop within state D but soon spilled over into the territory of state A. This was to be expected given the shape of the cultural landscape. Capital D is located in a cultural basin of high similarity that intersects the borders with states A, B, and G. The settlements of 1-nationals in state A are far away from the capital, so it is not surprising that that they started a nationalist secession campaign (see the crosses).

The settlements of nation 1 inside state A violate the most basic principle of nationalism: namely that each nation should possess its own state, or, for short, there should be national

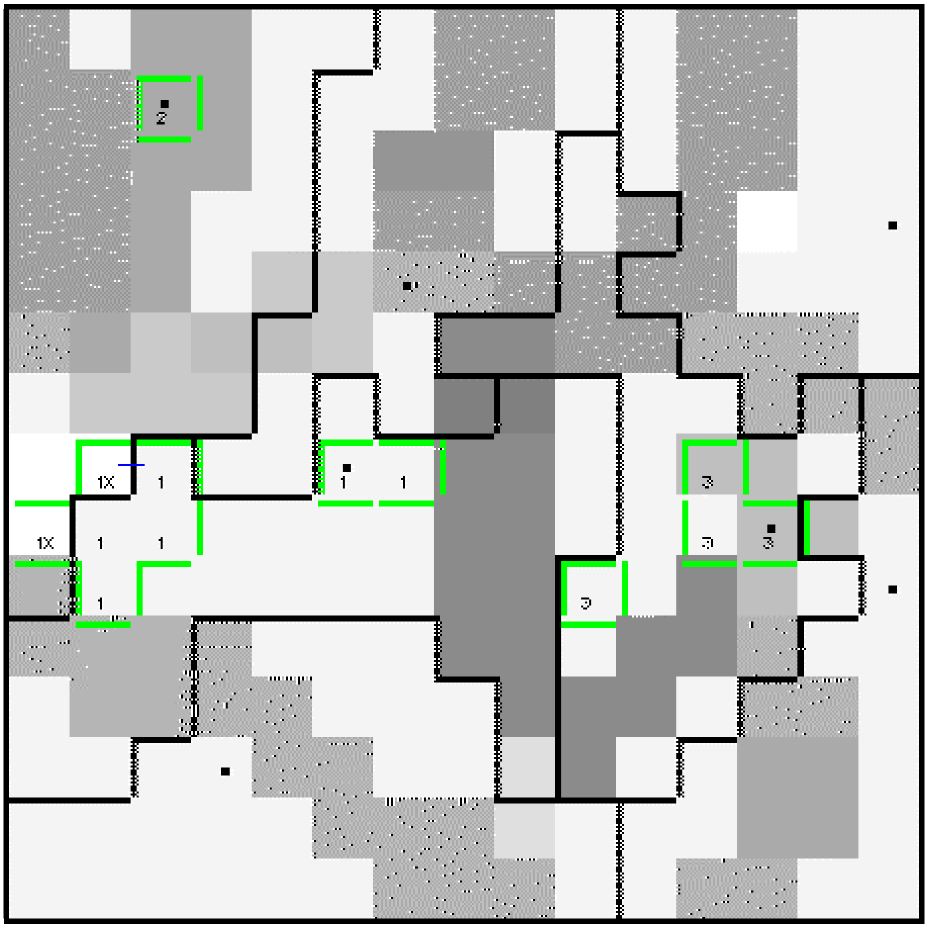


FIGURE 5 Irredentist Action in the Era of Nationalism at Time 536

self-determination. The discrepancy can be rectified both internally and externally. As pointed out by Myron Weiner (1971), these two phenomena often appear together. Whereas the former possibility corresponds to nationalist secession, the latter relates to irredentism. The behavioral rules of the nationalism extension turn all mobilized capitals and provinces into nationalist actors without entirely depriving them of geopolitical "appetite."

In our sample history, the co-nationals' secession campaign obliges state D, which is also a member of nation 1, to take irredentist action in order to help its kin across the border. The resulting interstate warfare, which involves the forces of capital A against those of state D, compounds the secessionist civil war, pitting the provincial nationalists (marked by crosses) against capital A.

In general, nationalist capitals seek to liberate their nationalist kin in other states if these populations do not enjoy "home rule." Provinces that belong to a nation modify their strategy such that they try to jointly break out of "foreign rule." This calculation implies that the national communities' decision making can be coordinated and their resources pooled within the national community. Postulating a more uncertain and risk-taking mode of decision making characterizing nationalist politics, I have set the superiority threshold at 1.5 for both internal and external action with considerably more uncertainty than in the non-nationalist case. Without abandoning entirely power considerations, states and provinces acting on behalf of their nations are thus more trigger happy than when acting on their own behalf. Moreover, there is an automatic obligation to come to the rescue of any kin group fighting a third party. On the whole, these rules are more prone to drag nationalist actors into armed struggle than the purely geopolitical strategy.

Interstate combat between A and D continues for several rounds, the whole time combined with rebellious activity on the part of the nation-1 settlements in state A. But because the balance is fairly even between the two camps, there are no territorial changes for the time being. At time 568, however, fighting diffuses to state B as consequence of nation 1's mobilization spreading to the province within B immediately west of capital D. This development triggers immediate irredentist action on the part of D (see Figure 6). As this province secedes, it first becomes independent, but thanks to the unification mechanism, it quickly joins state D. Completely unprepared to fight, state B starts to crumble. In time period 572, there is a wave of rebellions in the northern part, reflecting the longer distance from the capital than in other parts of the territory. Moreover, state B's territorial neighbors A and C profit from its weakness.

Nevertheless, while losing some territory, B manages to retain its sovereignty. At time 605, it is not the existence of this state, but of state D, that is threatened (see Figure 7). As a result of continuous irredentist fighting on its western front, it has just lost a nation-1 province to

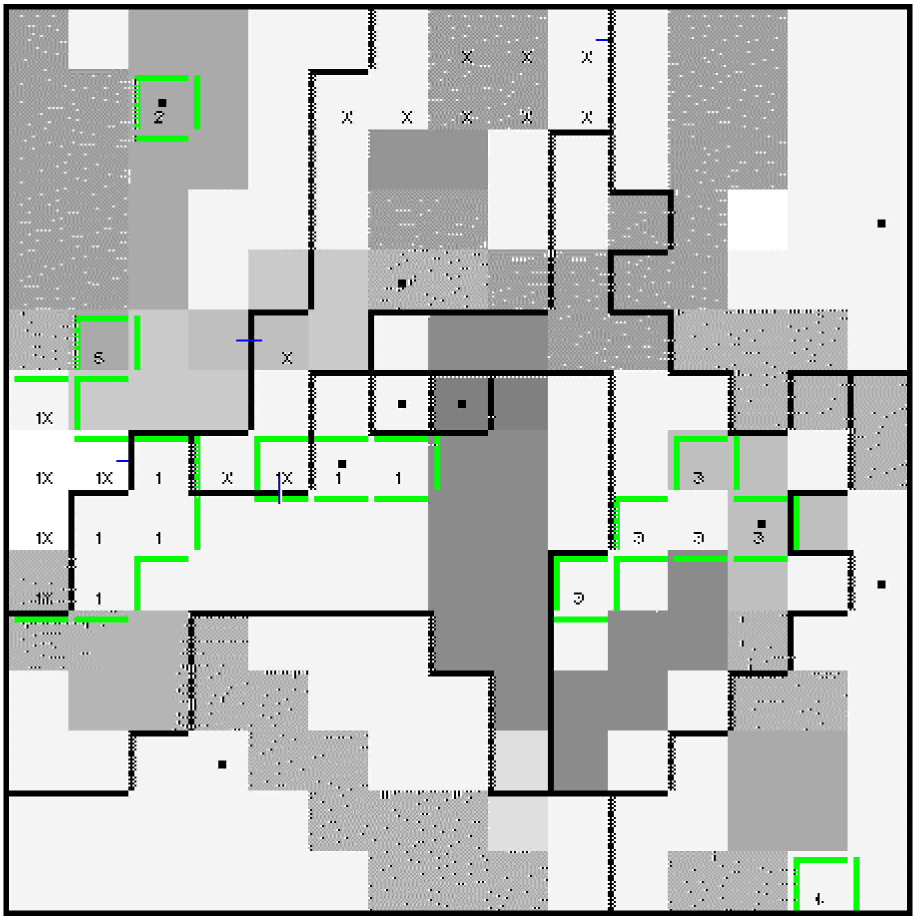


FIGURE 6 State B under Attack at Time 572

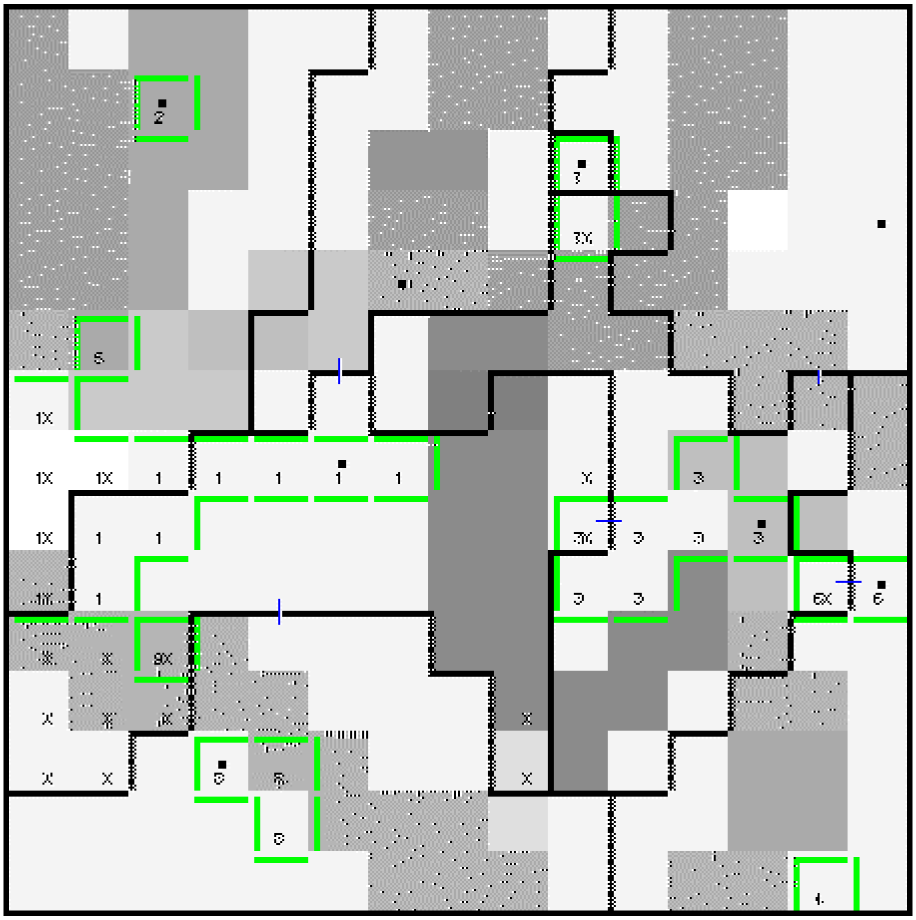


FIGURE 7 State D under Attack on Several Fronts at Time 605

state A. All the other sovereign neighbors have started wars against D, having diluted its power on too many fronts. In fact, on its eastern front, state E is attempting to redeem a newly mobilized group of nation-3 activists.[4] Finally, in the southwestern and southeastern parts of the country, there are regional rebellions.

Where does all this activity lead? Figure 8 portrays the long-run equilibrium after 1,500 iterations. This situation features four surviving states: A, C, E, and G. Unsurprisingly, state D never recovered from the "feeding frenzy" by its internal and external enemies. As a result of this development, nation 1 loses its sovereign political representation and becomes a national minority within state A. Note that some of the sites within this community have been converted to nation 2, the state-carrying identity of A. When state B finally collapsed, the capital of A was engaged in interstate conflict that provoked a rebellion on the part of the nation-1 minority. But once state B disappeared, state A turned all its power against the insurgents and crushed them. Under normal circumstances, national communities can never be destroyed, but in

---

[4] State E needs to watch out, however, because state C has initiated a campaign against it, and state F is acting on an irredentist claim to a province inhabited by nation 6.

the event of a central power's subjugation of a nationalist secession attempt, the core sometimes gets the chance to "ethnically cleanse" the province through migration from one of the co-nationals residing within its territory. Such a transition also shifts the national identity to that of the capital. This is exactly what happened in our example. Indeed, the cultural plain that was previously monopolized by nation 1 has become fragmented by the nation-2 colonies.

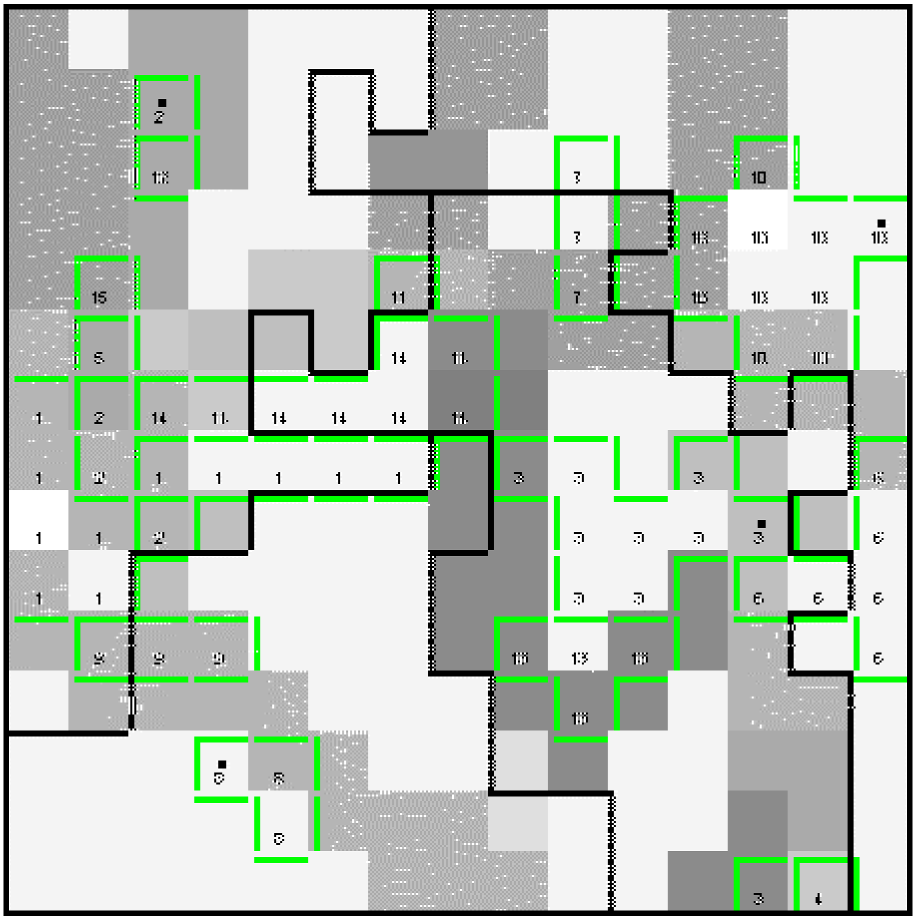


FIGURE 8 The Sample System in Equilibrium at Time 1500

Other nations have faired equally badly. The once state-controlling nation 6 has been split between states C and E after unsuccessful irredentist challenges to E's power. A few national communities are "unhistorical," since they formed too late to gain a state on their own (see nations 7, 13, and 14; each nation's number reflects its temporal appearance). The lucky nations are those that grew up around the capitals. There are two major basins of national activity (see nations 3 and 10). The cultural conditions in states A and G are less favorable for state-framed national mobilization. Capital A is surrounded by a "rugged" cultural region, preventing it from spreading the state's national identity 2 other than in the cleansed areas. In the case of G, it does not help to have the capital located in such a basin, because too specific a national identity prevents nationalism from catching on beyond the most immediate area.

AN ALTERNATIVE NATIONALIST HISTORY WITH STRONG STATE-LED ASSIMILATION

The assumption that the capitals have no influence over the cultural landscape within their states does not capture the conditions of nation-building in all areas. True, in some cases, as in Eastern Europe or the Third World, states have left a rather modest imprint on the cultural landscape (Schieder 1991; Gellner 1997). Even in the weakest cases of nation-building, however, one would expect some cultural spill-over beyond what the cultural background offers. In some prominent cases, states have indeed had a very strong impact on culture through centralized assimilation, even if it takes time for assimilation to penetrate large territories (E. Weber 1976). In early-modern Western Europe, such processes proceeded mostly unintentionally, as a side-effect of administrative standardization, military service, or commerce (Mann 1992). The model includes a parameter controlling the states' capacity to assimilate the culture of their provinces both before and after the onset of nationalism. As indicated in the introduction, the cultural assimilation parameter plays a central role as independent variable in this paper, and it is subject to controlled manipulation in the replications reported on below.

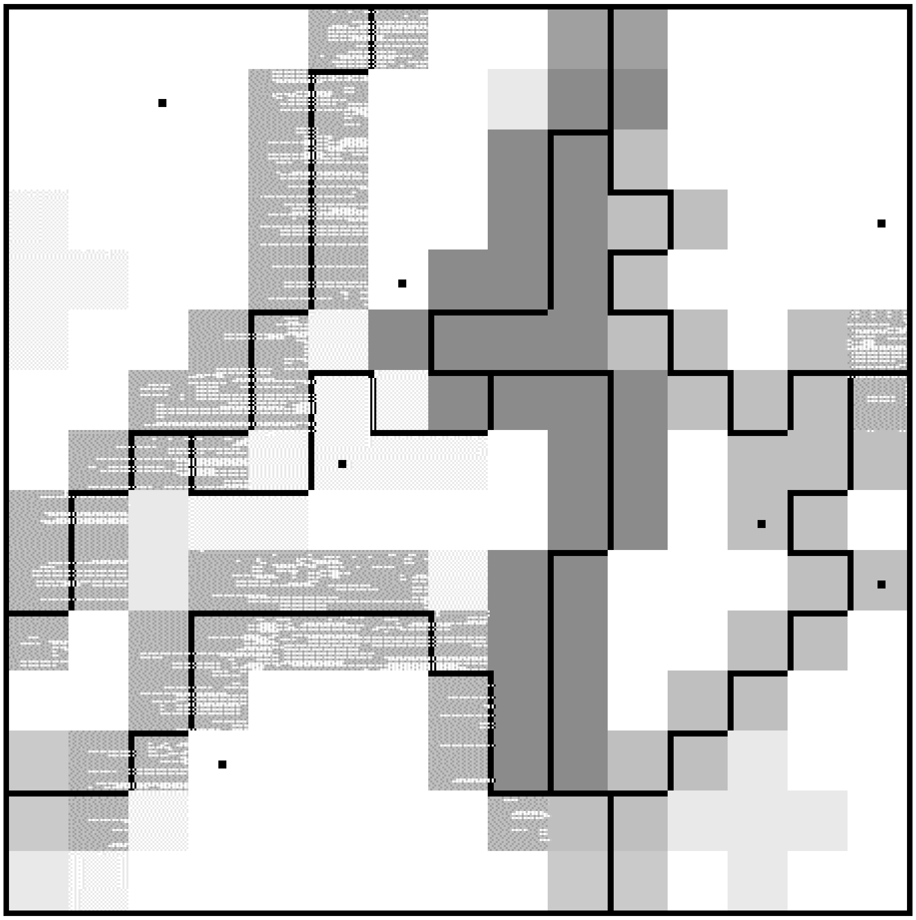


FIGURE 9 The Assimilatory System at Time 500

What would happen if we were to replay the "tape of history" with a high level of state-framed cultural penetration? As already stated, one of the main purposes of this paper is to investigate the geopolitical repercussions of different levels of cultural assimilation. So far, the grid has resembled eastern, rather than western, Europe, or perhaps even the Third World. Agent-based modeling makes it possible to rewrite history massively and to study the systemic effect of complex historical processes such as nationalism.

Starting with an identical system as in Figure 2, the states are now allowed to standardize culture quickly from the very beginning. Figure 9 shows the result of this changed specification at time 500. Compare this new setting to that of Figure 4 (as well as Figures 5 through 8), in which the ethnic landscape coincides with the initial cultural map. In the present case, clear boundaries coinciding with state borders have started to emerge. (Recall that whereas darker shades correspond to cultural border sites, the brighter ones denote areas with similar culture.)

Once nationalist mobilization is triggered, it will follow very different lines from those characterizing the culturally decentralized system (see Figure 10). In the counterfactual sample run, the activity starts in capital E and spreads quickly throughout its provinces at time 527, shortly followed by a similar process in state D. Thanks to the coincidence of political and cultural boundaries, national mobilization campaigns remain safely inside the states' borders and are therefore less likely to cause geopolitical havoc.

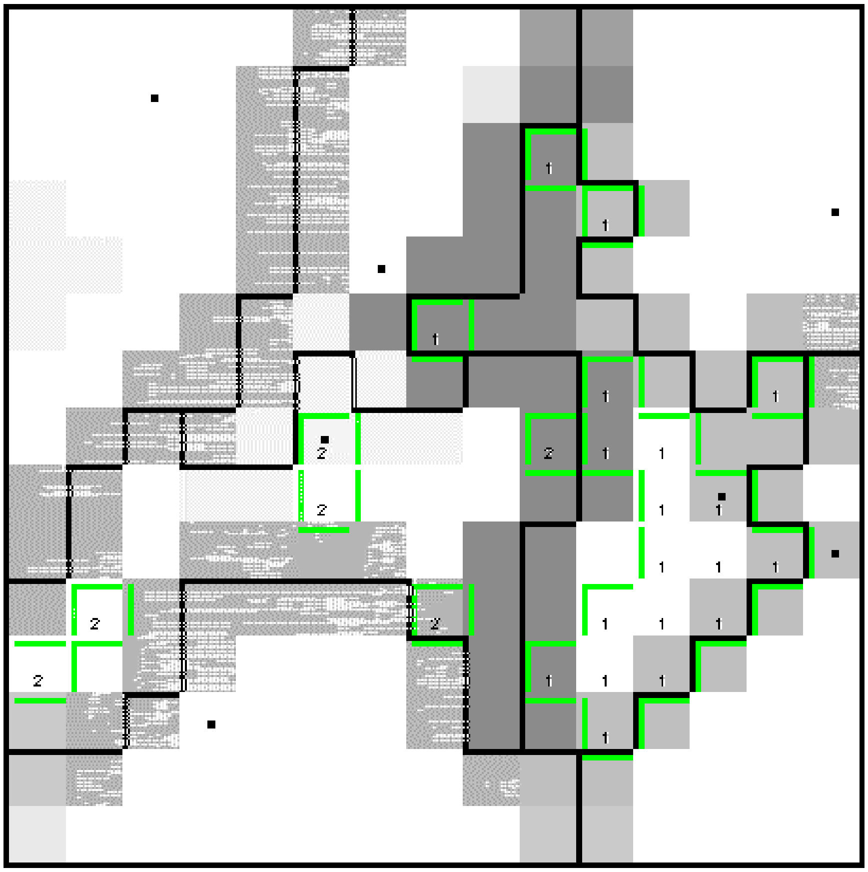


FIGURE 10 State-Framed Nation-Building Progresses Quickly at Time 527

Yet, even with high assimilation speeds, the balance of power can be undermined even in culturally centralized systems. The most obvious threat is unification in cases where the cultural border between two states is insufficient to stop border-transgressing mobilization. A less obvious source of instability derives from the differentiated resource extraction rule. It is assumed that taxation of co-national provinces proceeds without distance discounting.[5] Thus, uneven timing of nationalist mobilization, as, for example, Napoleon's pioneering use of *la levée en masse*, could upset the balance.

All the same, in our sample run, no such disturbances occur. Focusing on time period 1500, Figure 11 illustrates how the entire system has undergone a nationalist revolution without this affecting the geopolitical map. As opposed to the outcome of the culturally fragmented run reported on in Figures 5 through 8, states B, D, and F are saved by state-led assimilation and the lack of external irredentist temptations or irredenta located inside their territories. Moreover, as an emergent result, the fit between states and nations is perfect. All nations own their own states, and the states are ideal nation-states without any minorities.

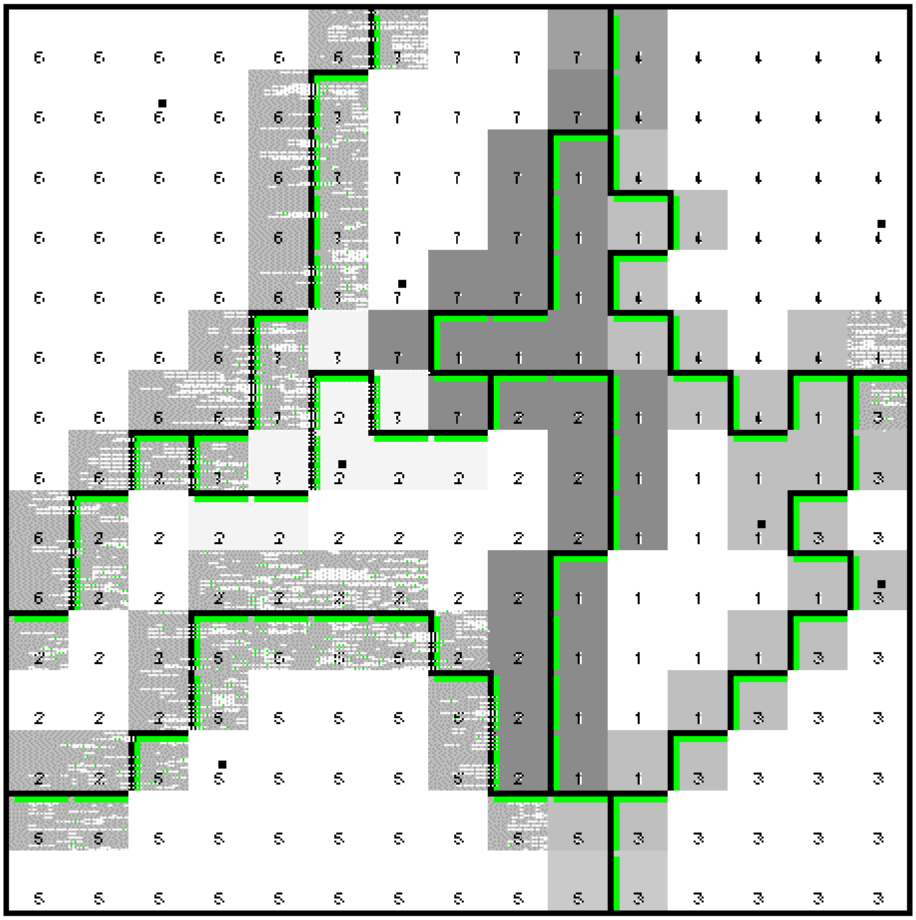


FIGURE 11 State-Framed Nation-Building Produces Perfect Nation-States at Time 1500

---

[5] The higher level of loyalty through the printed word or modern mass media justifies the absence of a distance gradient (cf. Anderson 1991).

CONCLUSION

Notwithstanding lack of comprehensive sensitivity analysis,[6] this study has generated a number of insights. First of all, it highlights the value added of a truly systemic perspective that views nationalist conflict as a side-effect of fundamental sociological transformation. Whereas most analysts have contented themselves with treating states and nations as given entities while focusing on the behavioral interactions between them, the computational framework allows us to trace the macro effects of nationalism, including structural transformations of both cultural and political boundaries. While far from the only factor influencing such processes, state-led assimilation has played the central role as the main determinant of this geocultural process.

States' cultural penetration, which includes ostensibly "soft" factors such as education and language policy, explains why states and nations coincide in some historical cases, whereas in others they do not. Endogenizing such a process puts the analysis on a more secure, constructivist basis than has been proposed by IR scholars.

Agent-based modeling helps us put the systemic effects of geopolitics and nationalism into perspective. In analogy with Schelling's (1978) famous segregation model that traces residential clustering of two initially interspersed ethnic groups, it is not too fanciful to view the entire international system as a sorting mechanism, where national self-determination acts as an institutionalized motivational rule to reduce the overall "frustration" of the system (i.e., the deviation from the nationalist idea of one-nation-one-state). At the level of mechanisms, however, there are important differences. Whereas migration drives convergence in Schelling's framework and the more general class of Tiebout models (Kollman, Miller, and Page 1997), systemic adjustments in world politics happen through state-transforming events, such as secession, unification, and irredentist conquest, and the nation-altering processes of assimilation and nation-building.

The present geocultural model also clarifies some nontrivial aspects of constructivist nationalism theorizing. Rather than assuming an essentialist one-to-one correspondence between cultural cores and national identities from the outset, the separation of cultural landscapes from nationalist mobilization facilitates analysis of the conditions under which such emergent outcomes become likely. As we have seen, in addition to preexisting ethnic conditions, state-led assimilation plays a key role in the generation of national identities. Yet, there is no reason to expect cultural identities to be easily "malleable," especially once nationalist mobilization has taken off. This is ultimately an empirical issue that has to be determined in particular cases. Nor is there any need to postulate that nationalism is fundamentally an irrational force. While I have assumed nationalist politics to be characterized by higher levels of uncertainty and risk-taking behavior, all the simulations in this study were based on a power-sensitive decision rule. Yet perhaps the most important contribution relates to the systemic context of nationalism. So far, most of the specialized literature on nationalism has adopted a one-country focus while ignoring geopolitical interaction effects. By providing a fundamentally co-evolutionary design that problematizes states and nations as distinct entities, the current model places constructivist theories of national mobilization within an ecological context of state interactions.

---

[6] Cederman (2000) reports on systematic replication results from this model.

REFERENCES

Anderson, Benedict. 1991. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso.

Axelrod, Robert. 1997. *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. Princeton: Princeton University Press.

Bremer, Stuart A. and Michael Mihalka. 1977. “Machiavelli in Machina: Or Politics Among Hexagons.” In *Problems of World Modeling*, ed. Karl W. Deutsch. Boston: Ballinger.

Cederman, Lars-Erik. 1997. *Emergent Actors in World Politics: How States and Nations Develop and Dissolve*. Princeton: Princeton University Press.

Cederman, Lars-Erik. 2000. “The Co-Evolution of States and Nations: Modeling Nationalist Conflict as a Side-Effect of Macro-Historical Change” Paper presented at the Annual Convention of the American Political Science Association, Washington DC.

Cusack, Thomas R. and Richard Stoll. 1990. *Exploring Realpolitik: Probing International Relations Theory with Computer Simulation*. Boulder: Lynnie Rienner.

Gellner, Ernest. 1997. *Nationalism*. New York: NYU Press.

Kollman, Ken, John H. Miller, and Scott E. Page. 1997. “Political Institutions and Sorting in a Tiebout Model.” *American Economic Review*.

Mann, Michael. 1992. “The Emergence of Modern European Nationalism.” In *Transition to Modernity: Essays on Power, Wealth and Belief*, ed. John A. Hall and Ian Jarvie. Cambridge: Cambridge University Press.

Schelling, Thomas C. 1978. *Micromotives and Macrobehavior*. New York: W. W. Norton.

Schieder, Theodor. 1991. *Nationalismus und Nationalstaat: Studien zum nationalen Problemen im modernen Europa*. Göttingen: Vandenhoeck und Ruprecht.

Weber, Eugen. 1976. *Peasants into Frenchmen: The Modernization of Rural France, 1870-1914*. Stanford: Stanford University Press.

Weber, Max. 1946. *From Max Weber: Essays in Sociology*. New York: Oxford University Press.

Weiner, Myron. 1971. “The Macedonian Syndrome: An Historical Model of International Relations and Political Development.” *World Politics* 23: 665-683.

TRANSMISSION OF CULTURAL TRAITS BY EMULATION: A VECTOR VOTING MODEL

R.G. REYNOLDS, Wayne State University*
R. WHALLON, University of Michigan
S. GOODHALL, Compuware

ABSTRACT

The goal of this research is to assess the impact of Culture on decision-making behavior. Specifically we are concerned whether the emergence of human culture provided humans with an adaptive advantage over non-human primate counterparts in terms of hunter-gathering capabilities. Reynolds has proposed several mathematical models of hunter-gatherer and primate decision making based upon differences in human and primate cultural traits [1], these were labeled the cultural and vector voting algorithm models respectively. In this paper an agent-based implementation of the vector voting model using Swarm is presented. Learning takes place there by emulation. The performance of this model in a variety of landscapes is compared with that of a random walk model. The results suggest that the vector voting model can produce a variety of emergent patterns that can be considered adaptive and that reflect human foraging patterns as well.

INTRODUCTION

The goal of this research is to assess the impact of Culture on decision-making behavior. Specifically we are concerned whether the emergence of human culture provided humans with an adaptive advantage over non-human primate counterparts in terms of hunter-gathering capabilities. Reynolds has proposed several mathematical models of hunter-gatherer and primate decision making based upon differences in human and primate cultural traits [1], these were labeled the vector voting and cultural algorithm models respectively. In the vector voting model each individual's vote was based upon their own knowledge and knowledge was not shared between individuals. The decision made by the group was a consensus based upon the weights and opinions of the members and was based upon patterns of interaction seen among primate groups. In the cultural algorithm model the individuals knowledge was pooled and used by a central decision maker to produce a decision.

The basic context in which these two decision-theoretic models were compared was a two-dimensional cellular space divided into R discrete sub-regions or cells each of unit area. The task facing the model groups was to compute the answer to various spatial predicates or queries about the region based upon the agents current knowledge. The models were analyzed theoretically and it was shown that the ability to form a collective intelligence through the pooling of knowledge had some distinct advantages.

---

* *Corresponding author address*: Robert G. Reynolds, 432 State St., Wayne State University, Detroit, MI 48202; e-mail: 220057@wayne.edu.

In particular, predicates such as the best direction within a region in which to forage, the one containing the most resources, was limited by the maximum area over which each individual had knowledge. On the other hand, pooling of that knowledge theoretically allowed a group to make these decisions over the entire region.

However, even in a social system where knowledge is not directly pooled, learning can take place. For example, Tomasello and Call [2] state that there are many similarities between how humans and primates understand their social worlds. Each has its own cultural system. In an extensive survey of primate cognition studies they conclude that, "all primates live in basically that same type of social world, in which they individually recognize conspecifics and appreciate both the vertical (dominance) and horizontal (affiliative) relationships that hold between group members. They also have the ability to predict the behavior of conspecifics in many situations based upon combinations of cues and insights, and in some cases to affect the behavior of groupmates via various social and communicative strategies."

They state that the basic difference between primate and human cultures is that in the latter the "intersubjectivity of human linguistic symbols — and their perspective nature as one offshoot of this intersubjectivity — means that linguistic symbols do not represent the world directly, in the manner of perceptual or sensory-motor representations, but rather they are used by people to induce, to construe, certain perceptual conceptual situations — to attend to them — in one way rather than another.

Thus, learning in a primate social system relies heavily on **emulative learning**. An individual watches another perform an action and observes the state changes that result. Thus, learning in this context is directly associated with sensory motor activities relative to objects in the environment. While humans can acquire knowledge in this way as well, they are able to support the **imitative learning** of concepts. With imitative learning "an individual understands others as intentional agents, like the self, that have a perspective on the world that can be followed into, directed, and shared."

The idea is that even when a group makes a decision based upon the knowledge of each individual without pooling, the physical results of that decision can be observed by everyone and learning can take place in an emulative fashion. The question of interest in this paper is what additional behaviors emerge from a group that uses the vector voting approach along with an emulative learning process.

Here we propose an agent-based implementation of the vector voting model using the Swarm simulation environment. First, we describe the vector voting model and the emulative learning method that serves as the basis for primate cultural transmission here. Next, we describe the results of running this model in a variety of environments with a variety of social configurations, and then we summarize our conclusions.

IMPLEMENTING CULTURAL TRANSMISSION IN THE VECTOR VOTING MODEL

The vector voting model represented how primate groups made consensus-based decisions. One of those decision was the direction in which to move during the day. It was shown in [3] that it was not possible for a vector voting model to always select the direction with the maximal amount of resources within a region R. However, theoretical limits aside, how does the vector voting model perform when coupled with emulative learning, the type of learning that is frequently observed in primate cultures?

In emulative learning the observer makes a cognitive connection between what action is performed and the state changes it produces. For example, a primate can observe another rolling over a log and exposing a number of insects. That action can be viewed as producing a state change which can be stored in memory. In our model, the result of a directional decision produces a trajectory through the landscape. As a result of that path each individual has an opportunity to get fed. In each cell that the group enters the resources there are divided among group members by various strategies such as priority or fixed order access or equal sharing of resources.

At the end of the day each individual can store a memory, not of the decision, but of the result. An individual does this by associating a visible landmark with the degree to which they were fed (satisfaction scale) the day they saw that landmark. A memory can have more than one landmark attached to it and different individuals in the group can associate different landmarks with the memory. Each individual has a maximum number of memories that it can store, and a memory is forgotten after a certain number of days (memdepth) unless it is used again.

Emulative learning using memories associated with the icons encountered relative to a group's previous decision can now be used to impact future decisions. Here, our region R has a number of landmarks (numLandmarks) which are distributed through the space. The cellular space is of size N by M and a cell can have at most one landmark assigned to it. Each day after the group decides on a direction to move based upon the memories associated with the landmarks currently visible from their location (visibility).

Each individual effectively pools the satisfaction scores for the memories associated with the visible landmarks in each direction. Here the scores are represented by a preference scale from –5 to +5 where the + direction represents satisfaction and – dissatisfaction. The direction with the highest score is the direction of choice for an individual. That choice is weighted by their status in the group. Each member then moves in the direction of their choice and is observed by the others. The group moves in the direction which achieves the highest consensus.

Group size can change based upon the extent to which individuals are fed. Individuals who have not gotten sufficient resources over a given period die and a group is removed when all of its individuals have died. On the other hand, if the group has been able to feed all of its members over a given period it can add a new individual up to a maximum group size. When it reaches that size it can fission into two new groups.

EXPERIMENTAL RESULTS

The vector voting model augmented with emulative learning was applied to a variety of different environments, each with a different resource distribution pattern. Figure 1 gives an example of a patchwork environment for a given set of 10 runs for each of two different configurations. A table with the set of parameters for each of the competing configurations is given in Table 1. The goals of these experiments are threefold:

1. To compare the model to a baseline random walk through the environment.

2. To observe any emergent patterns of foraging behavior that correspond to those exhibited in primate and hunter-gatherer groups.

3. To observe the relative survivability of groups in terms of various cognitive and social parameters, e.g., the number of memories an individual can have.

While a number of experiments have been conducted, we will summarize some of the results here.

When compared to a random walk the vector voting model invariably plateaued out at a number of surviving individuals that was below the carrying capacity of the environment but substantially above that of the random walk model which converged to a zero population. That exact location of the equilibrium point is a function of a number of model parameters.

Figure 2 gives the number of surviving individuals for the run associated with the environment in Figure 1 and compares the vector voting model to the random walk. As exhibited there, not being able to use knowledge in making a best direction choice will ultimately cause the system to collapse.

Surviving group also exhibited certain patterns of foraging behavior observed in primate and human groups. Specifically, each group began to forage in a area associated with positive landmarks in a cyclic fashion. The size of the territory reflected the spacing between. Figure 3 shows the distances between groups for 10 runs of the vector voting model versus that of the random walk. Notice that the distribution of distances between groups is much more focused than that for the random walk.

It is also interesting to note that those surviving groups for random walk are larger than the vector voting group. This is due to the fact that more individuals foraging can compensate for the lack of specific information. This is shown in Figure 4.

A final observation can be made in terms of the memories. Substantially fewer memories were needed to produce the behaviors above for vector voting groups than were used. This can be seen in Figure 5. Overall, random search produced more bad memories than memory based search with the vector voting model. Notice that in the surviving groups both positive and negative memories were associated with the territorial and cyclic foraging behavior of surviving bands.

CONCLUSIONS

In this paper we augmented the vector voting model of consensus-based decision-making in primate groups with emulative learning. Emulative learning was the basic form for cultural transmission used here. Extensive runs of this model suggest that the consensus based approach based upon icon memories was sufficient to produce regional stability below the carrying capacity of the region, territorial behavior, and a cyclical foraging patterns. It should be mentioned however that the territories could shift over time as different groups disappeared from or were added to the environment. As such, it was flexible to changes in group numbers within the region.

The next step in our project is to implement the Cultural Algorithm paradigm which allows the individuals to pool their memories. Theoretically, this sharing of information will allow the group to make regional decisions about certain predicates more reliably than in the vector voting model. In particular, the pooling of information via a belief space will allow the group to decide on predicates that are not tied to particular locations (such as norms) within the environment. These are called position invariant predicates. Being able to do this makes imitative learning, a form of learning largely unique to the human species, possible [4]. Both activities are influenced by the structure of the language used to perform the pooling and within which the learned information can be articulated. This will be the subject of a subsequent paper.

ACKNOWLEDGMENTS

Dr. Reynolds' research is supported by NSF grant IIS-9907257.

REFERENCES

[1] Reynolds, R.G., (1978) "On Modeling the Evolution of Hunter-Gatherer Decision-Making Systems," *Geographical Analysis*, Vol. X:31-46.

[2] Tomasello, M. (1999) "The Human Adaptation for Culture," *Annual Review of Anthropology*, 28:509-529.

[3] Reynolds, R.G., and Zeigler, B.P. (1979) "A Formal Model for the Operation of Consensus-Based Hunting-Gathering Bands," in C. Rebfrew, K. Cooke (Eds.) *Transformations: Mathematical Approaches to Culture Change*, New York, Academic Press, pp. 405-418.

[4] Pedone, R., and Conte, R. (2000) "The Simmel Effect: Imitation and Avoidance in Social Hierarchies," paper to be presented at Workshop on Cultural Transmission, Sienna, Italy.

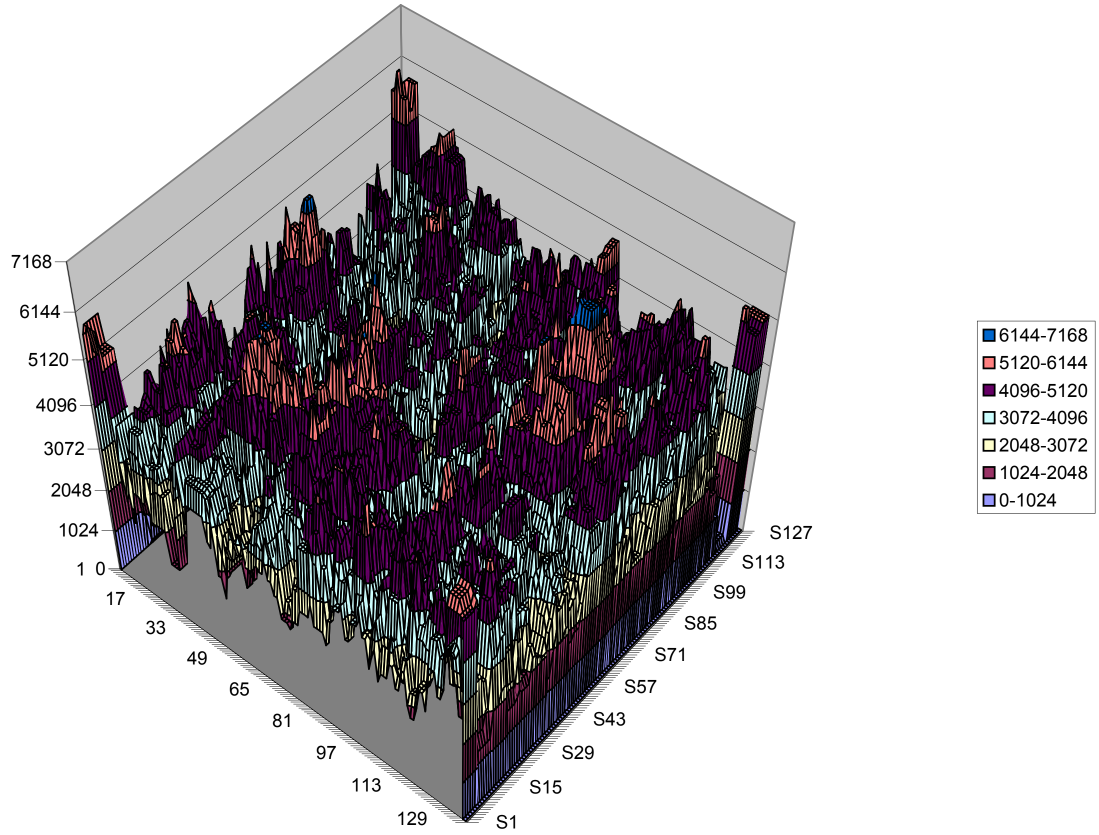


FIGURE 1 A patchwork environment, representing the yields for each cell in the example region

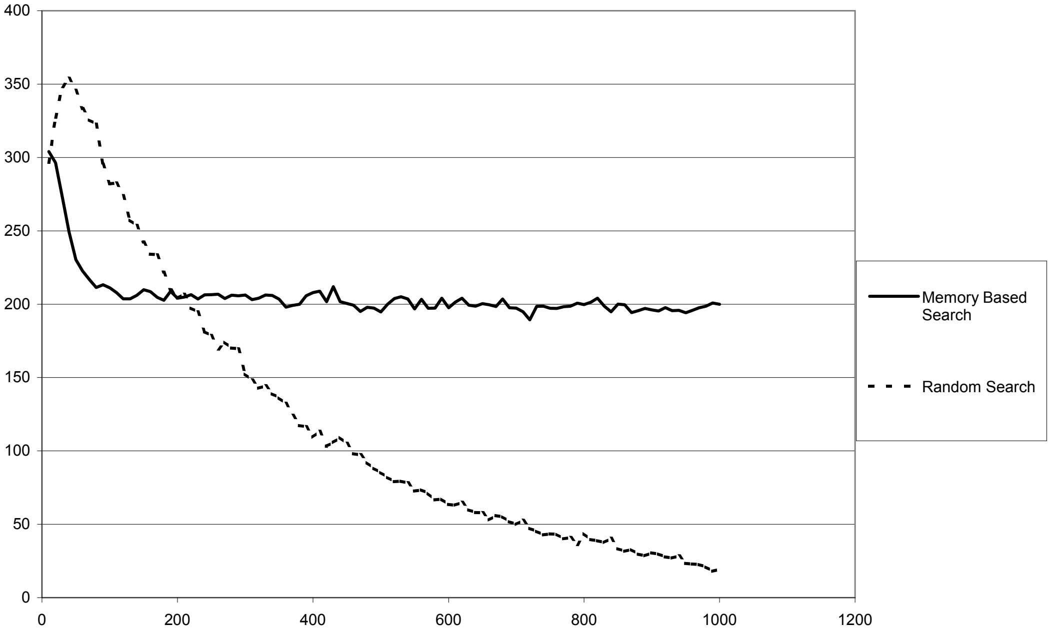


FIGURE 2 The number of surviving individuals in the vector voting model versus the random walk model, summarized over 10 runs each

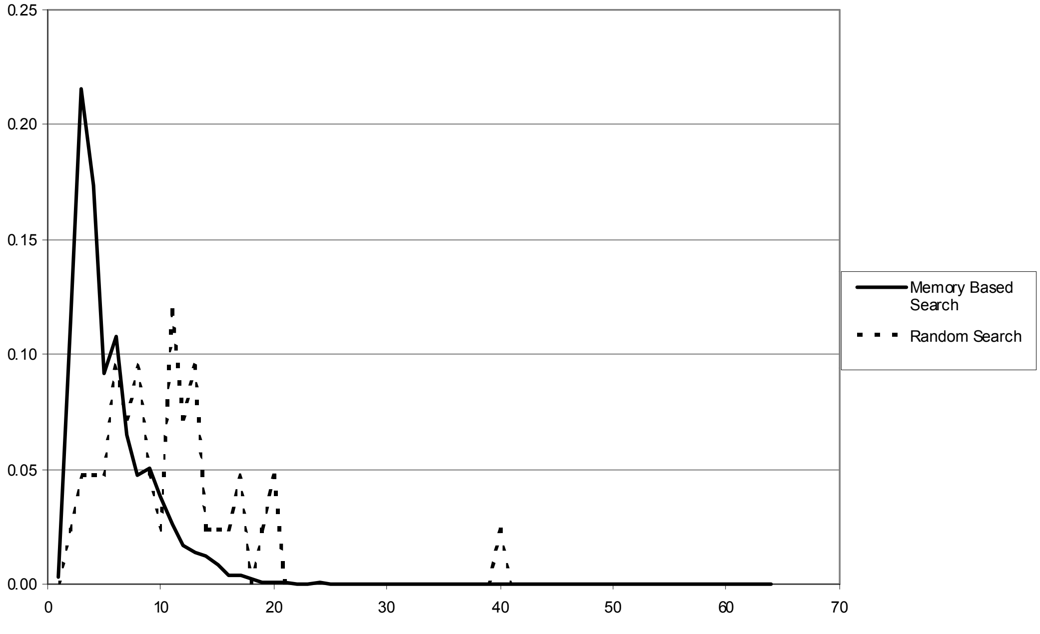


FIGURE 3 Distance between groups at the end of the simulation in the vector voting and random walk models

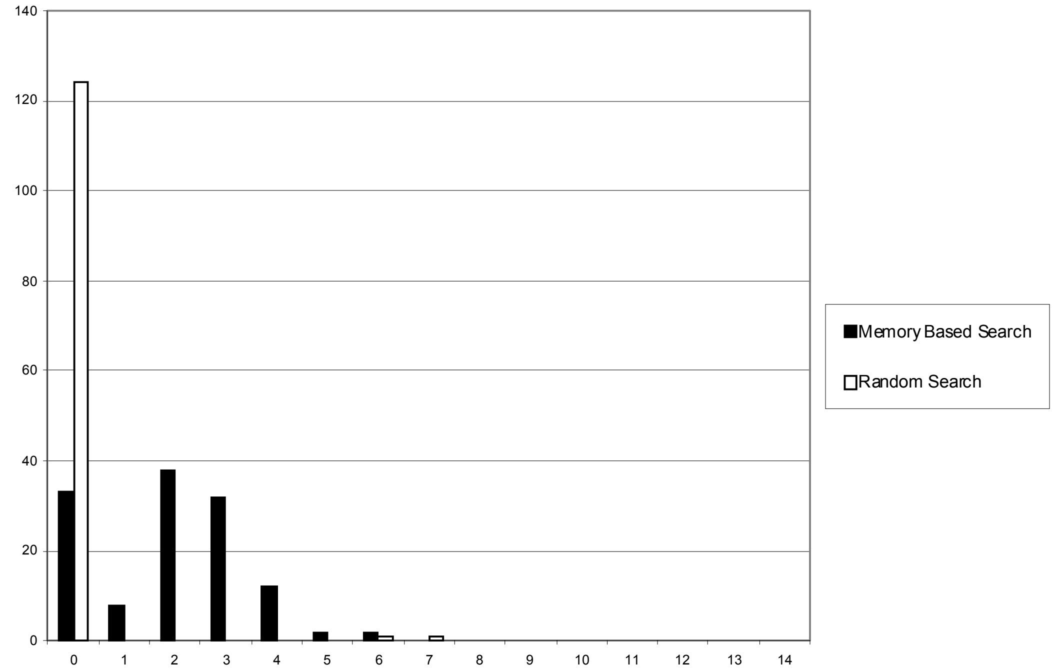


FIGURE 4 Size of the surviving groups at the end of 1000 time steps for the vector voting and random walk models

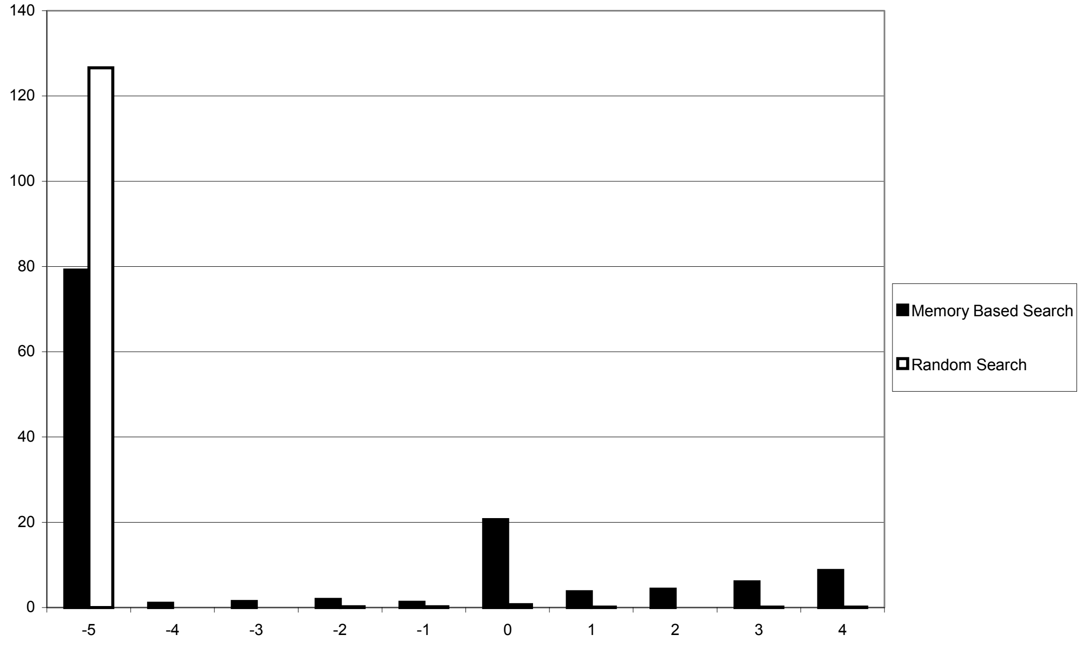


FIGURE 5 The number of memories of different types that are held by groups at the end of 1000 time steps in the simulation for the vector voting and random walk models

TABLE 1 Parameters That Can Be Set for the Vector Voting Simulation System (reflecting two specific configurations selected for testing)

ParmName	Parameter Description	Parm Value1	Parm Value Description1	Parm Value2	Parm Value Description2	Mismatch
bandSize	Maximum size of each band (number of members)	12		12		
bandSplit	Size at which band can split	0		0		
bandStart	Size of each band at the beginning of the run (number of members)	3		3		
baseMDReq	Base minimum daily food requirement for an adult	1500		1500		
bounceMode	Behavior mode when a band reaches the edge of the space or collides with another band	1	make new decision on bounce	1	make new decision on bounce	
CAAcceptance Interval	Cultural Algorithm Acceptance Interval (days): this is the interval at which the CA component accesses the data in the model space	10		10		
CAInfluence Interval	Cultural Algorithm Influence Interval (days): this is the interval at which the CA component provides feedback to the model space	100		100		
CAMode	Selects the CA algorithm to be used	1	basic landmark history	1	basic landmark history	
caption	Trial caption	Memory Based Search		Random Search		YES
consumeMode	Mode of consumption during foraging	0	fixed order	0	fixed order	
daysAhead	Number of days of consumption that an individual can stockpile to recover from shortfalls or to get ahead	+2.0000000000000000e+00		+2.0000000000000000e+00		
daysToDie	Number of days without MDR before an individual dies	7		7		

TABLE 1 (Cont.)

ParmName	Parameter Description	Parm Value1	Parm Value Description 1	Parm Value2	Parm Value Description2	Mismatch
decayDays	Number of days for a dead band to decay and disappear from the display. Until this interval expires, the square occupied by the band is blocked to other bands.	7		7		
diagBand	A band number to trigger diagnostic outputs that are limited to a selected band	0		0		
diagnosticDest	Destination for diagnostic outputs 1 = > console 2 => log 3 => both	0		0		
diagnosticInterval	Diagnostic interval in days: interval at which diagnostic outputs will be written	10		10		
diagnosticMask	Diagnostic mask: a bitmap selecting various classes of diagnostic outputs	0		0		
flipMode	Select a yes no parameter to be flipped for each trial	0	No Parameter Flipping	0	No Parameter Flipping	
foodDistFactor	First computational factor for food distribution. Used differently by each mode. Not always used.	+5.0000000000000000e-01		+5.0000000000000000e-01		
foodDistFactor2	Second computational factor for food distribution. Used differently by each mode. Not always used.	+5.0000000000000000e-01		+5.0000000000000000e-01		
foodDistMode	Food distribution function	6	patchwork around landmarks	6	patchwork around landmarks	

TABLE 1 (Cont.)

ParmName	Parameter Description	Parm Value1	Parm Value Description 1	Parm Value2	Parm Value Description2	Mismatch
foodDistRange	Food distribution range in squares. Usage varies by algorithm. Currently used only for mode 6, patchwork around Landmarks, to control the size of the patches.	12		12		
foodMaxValue	Maximum quantity of food that can exist in a single cell	16384		16384		
forageMax	Number of squares to forage in Mode 1	0		0		
forageMode	Forage mode 0 = en masse while traveling 1 = individual after traveling (unimplemented)	0		0		
lmDistMode	Landmark distribution mode	0	random (seeded corners)	0	random (seeded corners)	
lmPerMemory	Controls how many landmarks are attached to a memory when it is formed	2		2		
logDataInterval	Interval to write to data log (days)	0		0		
memDepth	Memory depth in days, number of days after a memory is formed before it is forgotten	64		64		
memMode	Memory formation mode	0	use n closest landmarks in any direction	0	use n closest landmarks in any direction	
numBands	Number of bands to generate	128		128		
numLandmarks	Number of landmarks to generate	128		128		
quitAfter	Terminate the mode after this many steps. Results in program failure. Superseded by batch mode operation.	0		0		

TABLE 1 (Cont.)

ParmName	Parameter Description	Parm Value1	Parm Value Description 1	Parm Value2	Parm Value Description2	Mismatch
randomDays	Number of days of random movement before using memory (during this interval, bands do not starve)	0		0		
regenRate	Rate of food regeneration per day (≤ 1.0)	+5.0000000000000003e-02		+5.0000000000000003e-02		
reincarnateDays	Number of days before a dead band is reincarnated	9999		9999		
reproduceDays	Number of days where the band is fully fed before it can add a member	10		10		
searchMode	Search mode	1	memory based vector voting	0	random	YES
seedProb	Probability of food in a cell (random distribution). Also controls the minimum amount of food. Not currently used in other modes.	+1.0000000000000000e+00		+1.0000000000000000e+00		
starveMode	Mode for determining starvation	1	deficit > MDR * days to die	1	deficit > MDR * days to die	
stepsPerDay	Maximum number of steps to move in one day	4		4		
stopEvery	Interval (days) at which the model will pause	100		100		
trialName	File name prefix for the trial (yyyymmdd_ hhmmss). This is not specified externally but is generated and exported at run time.	20000729_160227		20000729_151713		YES
visibility	Default landmark visibility in squares: for how many squares is a landmark visible	32		32		

TABLE 1 (Cont.)

ParmName	Parameter Description	Parm Value1	Parm Value Description 1	Parm Value2	Parm Value Description2	Mismatch
vision	Default vision in squares: how many squares can an individual see a landmark	32		32		
worldXSize	X dimension of the model space	128		128		
worldYSize	Y dimension of the model space	128		128		
zoomFactor	Swarm zoom factor (0 = automatic) 1-	0		0		

DISCUSSION: POLITICS

J. PADGETT, University of Chicago, Moderator

[Presentation by Lustick]

Michael North: I think what you presented here is very interesting. What I would like to see in future models is environments that change from very rational to irrational, because I perceive that as happening in society right now in many cases — where the environment itself is changing partway through the evolution of history.

Ian Lustick: First we wanted to find out a little bit about the world without that [variability]. One of the areas that we are working in is globalization, and in that context we will change the parameters of the global state and then watch what happens in the local state under varying conditions, so that would be a [way of looking at what you suggest].

[Presentation by Bendor]

Claudio Cioffi-Revilla: I found your presentation very stimulating because you've taken an important step in the direction of an alternative to the rational choice paradigm. It seems to me that there is a fundamental flaw in the rational choice paradigm. In this instance, it assumes that people vote because they believe that they're going to determine the election. I doubt there is any voter in the world who imagines that he will individually be able to determine the election. If that's the test that the rational choice theory offers, it's misleading about the way people behave and what they believe they can or cannot do.

It's probably true that most people vote because they think they *should* vote, because they feel an obligation to vote. The notion of obligation is important, because, like the notion of aspiration, which you're using, it looks in an entirely different direction for the motivation for voting. Regarding obligation, there's some very interesting literature on deontic logic. Did you look in that direction and decide not to pursue it? There are aspects of deontic logic that are very formalized.

Jonathan Bendor: That's a very good question. We have thought about that. Right now, the only way the model can represent obligation is in a fairly crude way by allowing the cost of voting to go negative to reflect a sense of obligation, a civic duty to vote. And let me just briefly report to you what happens. Suppose people have a negative cost of voting because they feel guilty if they don't vote. What do you think will happen? From a game-theoretic point of view, if people have negative costs of voting, then voting is a dominant strategy — you can't lose by voting. So you should see complete turnout. But we don't get complete turnout. Why? Because the aspiration levels rise when you have negative costs of voting. You can still care about the collective, electoral outcome, you can still be disappointed if you turn out and lose; therefore the model cannot stabilize at a propensity of one. I do think the question of obligation is a very important direction of research.

Maurits van der Veen: I have two related questions, one theoretical, one empirical. The first is, most people don’t vote in 1,000 elections over their lifetime. So I would be interested to see how quickly your model gets to the equilibrium outcome of about 50% turnout. And second, the model gives some suggestion that as people get older and presumably have voted in more elections, their turnout should essentially approach 50%, and that should be empirically testable.

Bendor: It is testable, but note that turnout may approach 50% by declining from above that value, rather than rising to it. As a matter of fact, a colleague of ours who studies Eastern Europe suggested to us that voting participation rates after the fall of communism began to decline after a while, and we have a result in the paper that explains that phenomenon. When you start with very high aspirations that the government can’t deliver, participation rates will tend to fall.

About the rate of time to convergence: we make no real-time claims in the paper, because, as in the linguistics paper that was given yesterday [Satterfield], we have no reason to believe that the speed-of-adjustment parameters reflect, or are calibrated to, the real world. We simply don’t know.

van der Veen: But every period is one vote experience.

Bendor: Yes, but how long it takes to converge to 50% depends critically on the speed-of-learning parameter, on the alphas and the betas.

van der Veen: Right. Well, can you tweak the speed of learning so that the [times] become realistic?

Bendor: Yes. First of all, the starting case of “all shirk” is not realistic. That’s just a thought experiment, as it were, to show that you cannot make the system stay at very low levels of participation — even if you start it there, you would see a breakout of participation.

One commentator on an earlier version of the paper, someone who studies elections, advised us that rate of convergence is not an issue in stable democracies. Turnout rates change very slowly, and the new generation is likely to “inherit” propensity rates that are quite close to their parents’ — so we don’t see a breakout of participation. Empirically, that would be a very unusual situation. We were simply demonstrating a theoretical property.

Michael Heaney: I’m Michael Heaney, University of Chicago. I wonder if you’ve thought about putting a couple of other features into the model: one to reflect the heterogeneity of people’s psychic involvements in the election, and another that would allow for variations in the social pressure to vote. I see these two factors as interacting. Some people are always going to vote, because they are psychically involved, they’re interested; other people don’t care; and there is a large degree of variation in between. And there’s the way people are connected to social networks. So people like myself who are always talking about politics are putting social pressure onto others. We’d expect that if a person had low psychic involvement combined with little social pressure, they’d be unlikely to vote. I think that including these variables in the model would be a mechanism to explain why some people vote and some don’t. Your model seems to assume that everyone is psychically involved in the election.

Bendor: Right now we can distinguish between people in the two factions. We can have heterogeneity and exogenous parameters within a faction. It's time-consuming, but it's relatively straightforward to reprogram the simulation. I think the network idea is definitely a worthwhile thing to do, and will yield estimable predictions.

[Presentation by Cederman]

Miles Parker: Miles Parker, Brookings. I think this is a very nice, rich model. I'm wondering if you've looked at the flip side of the nationalism issue, say, national bifurcation or new identities arising within nations. I'm thinking of the pre-Civil War South, for instance.

Lars-Erik Cederman: Actually, the current system doesn't allow for splits of nations as opposed to states. Once a national community has been formed, the nation can change only if it's completely broken up by a state capital waging a war against the periphery. But what you're saying is entirely right. Our model is a simplification. It's entirely possible to include rules that would allow both for mergers of split nations — smaller nations, say, that may be close to each other in culture — as well as for breakups of existing nations.

Parker: I have another question. It may just be happenstance, but I thought I noticed an episodic or punctuated flavor to the periods when a lot of changes happened. Have you noticed that?

Cederman: Very much so. This model is very close to the [Stephen Jay] Gould perspective of punctuated equilibria. This is what makes it so hard to study nationalism with micro-level theory, because within small timeframes nothing may happen. Much like what happens before an earthquake, there is a tendency for tensions to build up over time. And because of the threshold-type logic that I have built into this system for mobilization and conquest, you get periods of extensive change — a bit like in the Schelling model again. One family moving to another neighborhood may cause a chain of changes. So, absolutely, this episodic tendency is built into the system.

Lustick: I have a question about research strategy. Every time you get an intuition from another kind of theory, say mobilized populations, or secessionist theories, then the model becomes more complex. And when something doesn't happen the way you expected, it's much more difficult to determine why. After my presentation, I got a great suggestion to vary the environment. You can now vary the environment during a run. The number of opportunities to insert things that are not only plausible but that we also know are relevant is outrunning our ability to explore the space that we have created. So how do you balance it?

Cederman: There's no simple answer, but I would say first of all that you have to make a choice from the beginning, because it is a tradeoff between tractability and realism. I've been driven very much by my own intuitive or theoretical knowledge of nationalism that I've drawn mostly from the qualitative literature. My model is an attempt to get things down in a more formalized, "cleaner" fashion than in the qualitative literature.

We're not done — this is just a starting point. But let me emphasize that when you decide how complex a model should be, you always have to be concerned about the underlying theoretical assumptions. There's always a minimum level of simplicity below which you cannot go without violating the most fundamental assumptions of the qualitative literature. But our model is far from the very simple rational choice models that have become influential in the literature. I'm not saying that those models are not useful; I'm just saying that our model offers another perspective.

I believe there is a need to do more robustness checking of this model. I should emphasize that the goal of a model should not be to sweep the entire parameter space, because that's impossible. This is a very complex model. A full parameter sweep may also be a Holy Grail that we're not interested in. If you can get reasonably close to stylized scenarios from history and you can say something meaningful about possibilities, if you can get a feel for things you didn't expect, then you can go back to the historical record and look with this new intuition in mind. For instance, I didn't expect national unification to be as destructive as it was. So this is a qualitative and more heuristic use of simulations. I think that people have misunderstood the value of simulation in this regard — they have confused simulation with the hubris that existed in the older literature.

[Presentation by Reynolds]

David Sallach: Was the environment structured in regions? In other words, could the agent say it is dry in the west or cold in the north? I'd also like to ask about the extent of generalization or classification in the learning process.

Robert Reynolds: We can put in any function that we want to. For example, in one, the middle is awful, but if you go along the edges, it's good.

Sallach: So they were learning regional patterns?

Reynolds: They were learning regional patterns. For these examples, the regional patterns are pretty obvious to us. The patchwork environment that I gave you actually has regional patterns in it, but I can't see them. In fact, the resources are organized probabilistically around the landmarks, and we can adjust how much a focus a particular landmark is. So we can make landmarks in the north more focused for resources than ones in the south. They are learning those types of patterns, but it's very hard individually to see them.

Sallach: So it's iconic patterns as well.

Reynolds: Exactly.

Environmental Processes

THE PROCESS-INTERFACE-TOPOLOGY MODEL: OVERLOOKED ISSUES IN MODELING SOCIAL SYSTEMS

H.V.D. PARUNAK, Environmental Research Institute of Michigan*

ABSTRACT**

The Process-Interface-Topology (PIT) approach to modeling social systems considers the Processes executed by individual agents and by their environment, the Interfaces between participants, and the overall Topology of their interconnection. The PIT perspective focuses attention on important details that the conventional bipartite discussion of individual agents and agent organization sometimes overlooks. This paper identifies some of these details and argues that effective modeling of social systems must include integrated, disciplined analysis of all three aspects.

INTRODUCTION

Discussions of the design and engineering of multi-agent systems commonly focus on two aspects: the design of the individual agents, and design of their interactions. For instance, the Gaia methodology (Wooldridge, Jennings et al. 2000) analyzes a system in terms of roles and interactions, and produces a design around the services that individual agents perform and the acquaintances of each agent. The binary distinction characterizes our own ontology of agent applications (Parunak 1996; Parunak 1998), and is reflected in Jennings' distinction between the knowledge level and social level in agent systems (Jennings 1994). Holarchic (University of Hannover 2000) and compositional (Brazier, Jonkers et al. 1998) approaches emphasize that more layers can exist than just "agent" and "system," and discuss interactions as the mechanism by which entities at one level form an entity at a higher level.

This ontology is fine as far as it goes, but overlooks some key issues.

(1) A real social system is made up not only of agents, but also of an environment in which those agents exist and through which they interact. This environment must be taken into account in analyzing, modeling, and engineering social systems. The case for this position has been articulated for some years in the embodied cognitive science community, most recently as the principle that one must design an agent's ecological niche along with the agent itself (Pfeifer and Scheier 1999).

(2) Design of agent interactions is typically restricted to symbolic communication protocols. Other forms of interaction, often relying on nonsymbolic physical actions, can also be important.

---

* *Corresponding author address:* H. Van Dyke Parunak; ERIM, P.O. Box 134001, Ann Arbor, MI 48113-4001; e-mail: vparunak@erim.org.

(3) The behavior of the community emerges from the interactions of individual agents with one another and with the environment in a way that is not immediately obvious from the definitions of the individual components. Recent results in complexity science (Watts 1999) suggest that the topology of agent interaction is critical to the nature of this emergent behavior.

To accommodate these concerns, we have begun to analyze systems that we wish to model using a tripartite model, focusing on the *Processes* that take place in the system (whether in individual agents or in the environment), the *Interactions* among agents and between agents and the environment, and the *Topology* of those interactions. The successive sections of this paper discuss each of these areas, with special emphasis on details that one might be likely to overlook from the traditional bipartite perspective.

PROCESSES

A process P_i may be defined formally as

$$P_i = <V_i, R_i> \qquad (1)$$

where

V_i is a set of variables whose assignments change over time, and

$R_i \subset V_i^+ \times t \rightarrow V_i^+$ is a set of rules governing how those changes take place over time.

$\boldsymbol{P} = \bigcup_i \{P_i\}$ is the set of all processes.

Processes come in many flavors. The variables may be numerical or symbolic. Time may be continuous or discontinuous. The rules may be expressed computationally (in terms of rewrite rules) or as a system of differential equations. In classical AI, processes usually concern symbolic variables manipulated computationally in discontinuous time, while physics focuses on numerical variables whose values over continuous time are defined by differential equations, but some impressive examples apply physics-like processes to cognitive problems (Port and vanGelder 1995; Schöner, Dose et al. 1995).

We have elsewhere argued for the definition of an agent as a bounded process (Parunak 1997). Most current discussion on the architecture of individual agents focuses on defining that process in terms of computational mechanisms such as rule-based reasoning, multi-threaded object architectures, or even the integration of differential equations. Clearly, anyone who constructs an agent-based model of a social system must define the behavior of the individual agents. By drawing the analyst's attention to "processes" rather than "behaviors," we wish to emphasize that the system's behavior depends not only on the behaviors of individual agents ("bounded processes"), but also on processes that are not "bounded," but form part of the environment in which agents interact.

Social norms and governmental regulations are an important example of such processes in a social system, as are acts of God such as weather, earthquakes, and the natural deterioration of transportation facilities over time. In modeling a social system, sometimes it makes sense to instantiate an agent representing the "environment" to handle such influences. But we must first recognize that the influences exist, and that centralizing them in an "environment agent" is an implementation compromise. Agents that correspond to real bounded entities in the domain are qualitatively distinct from agents that serve to encapsulate diffuse, distributed influences, and we must be aware of the possibilities of unexpected dynamics in their interaction. Two examples of the atypical nature of environmental agents come to mind:

(1) "Environmental" agents may not respond to actions taken by "real" agents as predictably as other "real" agents (Ferber and Müller 1996). Error-correcting protocols and automated retransmission mechanisms can guarantee that a message sent by one agent reaches another, thus changing the state of the recipient (at least with respect to its mailbox). But a robotic agent might push on a boulder all day long and never make any difference to the environment.

(2) The environment is usually spatially distributed, and influences are often bounded in space and time. Thus an agent-based model of the environment may need to consist of a network of agents rather than a single environment agent. Experimental evidence (Parunak, Brueckner et al. 2000) shows that lumping spatially distributed processes into a single agent can distort the result of the overall model.

Control theoreticians are accustomed to making this distinction when they analyze the separate and combined effects of the dynamics of the controller and the dynamics of the plant (the entity being controlled). We recently encountered this effect in a project providing routing control for military aircraft, reported in detail elsewhere (Parunak, Brueckner et al. 2000). Briefly, we constructed a pheromone-based mechanism for guiding friendly aircraft through a space populated with targets that they should attack and threats that they should avoid, and then analyzed their performance as we varied the composition of friendly and adversarial forces. The outcome showed strong nonlinearities. As we studied these irregularities, we hypothesized that they were largely due not to our control mechanisms, but to the rules that governed the outcome of combats between the two forces. We were able to verify this hypothesis by constructing an abstract model that embodied only the dynamics of the environment, and comparing the performance landscape it produced with that resulting from the behavior of agents embedded in that environment.

Thus analysis of a social system to be modeled using agents must include a comprehensive review of the processes that are observed in the domain. Conformably with conventional approaches to agent-based modeling, many of these processes will be localized within bounded entities. In fact, it is reasonable to argue that only domains in which most processes are so localized are really good candidates for agent-based modeling. However, it is rare to find a domain in which all processes are bounded, and development of a faithful model requires that the analyst pay special attention to unbounded environmental processes as well.

INTERFACES

One important distinction between a monolithic software system (a "single-agent system") and a multi-agent system is the fact that agents (and more generally, processes) interact. The interface between two processes specifies how changes in one process's state variables affect the evolution of the other processes' variables. Formally, an interface I_j among a set of processes is itself a process that includes the union of the other processes, as well as additional rules R_I specifying the coupling across the original processes:

$$I_j = \left\langle \bigcup_i V_i, R_I \cup \bigcup_i R_i \right\rangle \qquad \textbf{(2)}$$

where

i is an index ranging over the processes in the interfaced set,

$R_I \subset \left(\bigcup_i V_i \right)^+ \times t \to \left(\bigcup_i V_i \right)^+$ (the "interface rules") is a set of rules spanning the variables of different processes and governing how those changes take place over time, and

$\boldsymbol{I} = \bigcup_j \{I_j\}$ is the set of all interfaces I_j.

In conventional MAS's, interfaces are restricted to inter-agent protocols. Such protocols are an essential component of the interfaces in modeling a real social system, but others must be considered as well. For example, in a market economy, protocols may support the negotiation over the nature and prices of services and goods to be exchanged. With the advent of electronic cash even the payment for those services and goods can be embedded in a protocols, and some services (notably information services) can be delivered through a protocol as well. But if a purchaser is buying a lawn-mowing service or a new car, at some point a physical transaction must take place that falls outside the neat definition of an information protocol. If this transaction is not represented, the overall model will be defective.

We experienced this problem in our early work on YAMS (Yet Another Manufacturing System), a control system for a flexible manufacturing system that distributed tasks across multiple manufacturing workstations (Parunak 1987). We carefully designed the protocol through which the workstations negotiated for allocation of tasks, taking care to construct a formal proof showing that the protocol would not deadlock. Then we implemented the agents, installed the system on physical machinery, and turned it on. Before very long, the system deadlocked! Our proof was not in error. However, the formal analysis covered only the movement of electronic messages among the agents representing the workstations. In the real world, physical parts also moved among the workstations, and their movement (and the information they conveyed by moving) was not included in our analysis.

Agent-based systems for information services can sometimes equate interfaces with protocols. Agent-based systems that model social systems cannot. Social systems engage people in interactions with one another and with the real world that include a variety of physical as well as informational influences, and our models must incorporate these aspects. For implementation reasons, we may need to model (say) the shipment of a car as a special message in an electronic protocol, just as we may need to model an environmental process as an agent. As with processes, so with interfaces, we need to recognize that such an implementation is a compromise, and may obscure important distinctions in the system we are trying to model.

TOPOLOGY

Interfaces induce a graph-like structure over the set of processes:

$$T = \langle \boldsymbol{P}, \boldsymbol{E} \rangle \tag{3}$$

where

$\boldsymbol{E} = \{E_1, ..., E_m\}$

$E_j \subset \Pi\boldsymbol{P}$ is a multi-edge connecting the processes in I_j.

It is commonplace in designing agent-based systems to construct a diagram more or less isomorphic to UML's acquaintance diagram, indicating which agents interact with which other agents. Just as agents do some but not all of the work of processes and protocols do some but not all of the work of interfaces, so acquaintance diagrams satisfy only part of the need for attention to topology. At least two defects encourage us to generalize the concept.

The first defect reflects once more the rude intrusion of the environment into situated agent-based systems, such as those that are intended to model social systems. To a first approximation, the topology of a system in which agents both influence and sense a shared environment is a star with the environment at the hub and agents at the ends of spokes, augmented by direct agent-to-agent communications. A serious weakness of this first approximation is its inattention to spatial and temporal locality. Typically, an agent can influence and sense only a limited area in its environment, and influences in the environment propagate at a finite speed, frequently decreasing in intensity as they propagate and as they age. Insect pheromones are a canonical example of environmentally-based interaction. One insect can sense the pheromones deposited by another only if it comes close enough to the original deposit, within a time window determined by the evaporation rate of the original deposit. Thus the environment links agents whose trajectories in space-time come close enough together. If agents can move, the topology of the system is a function of time. Synthetic pheromone infrastructures (Brueckner 2000; Parunak 2000) suggest one way to implement such a topology. The environment is modeled as a network of places, among which agents move. A place provides a number of services to agents that occupy it: agents can deposit pheromones on a place, and sense pheromone strengths at the place and its immediate neighbors. In addition, the place evaporates pheromones over time, and propagates them to its neighbors. Places themselves are naturally represented as (non-mobile) agents, subject to the caveats in Section 2. A static graph is clearly inappropriate to describe the topology of such a system. At a minimum, one might capture the topology by presenting one acquaintance diagram showing the interconnectivity of the places, a second representing the connectivity of an arbitrary place with one or more arbitrary agents that are resident on it at a given time, and a representation of the constraints on agent movement and the propagation and evaporation of influences that agents can exert on places.

A second defect of the acquaintance diagram as a representation of the topology of an agent-based system is its inattention to the dynamical implications of interconnectivity. Recent work (Watts 1999) shows that processes interacting through graphical structures can behave very differently depending on graph-theoretic features of those structures, features such as characteristic path length and clustering coefficient. Analysts and designers of multi-agent systems and models should recognize this work by paying attention to the potential for topology-dependent dynamics in their systems. Agent populations should be tested for their behavior on a range of topologies, including random, regular, and intermediate structures, to determine how the overall behavior may vary with connectivity.

SUMMARY

Classical analysis of agent-based systems is bipartite, focusing on individual agents and their community relations. The Process-Interface-Topology (PIT) model looks at issues that are orthogonal to this classical division (Table 1), and encourages us to identify details that might otherwise be missed, such as the role of an active environment, the possibility of indirect agent interaction through a shared environment, the implication of agent mobility on locality of interaction, and the influence of system topology on overall behavior.

TABLE 1 Comparing the Models

		Tripartite PIT Model		
		Processes	**Interfaces**	**Topology**
Classical Bipartite Model	**Individual Agents**	Agents are bounded processes.	Agent I/O may include sensors and actuators as well as digital communications.	Agent mobility can cause topology to change over time.
	Community	The environment, which spans multiple agents, may also support processes that impact the behavior of the system as a whole.	Agents interact both directly (classical protocols) and indirectly (through interactions with a shared environment).	The dynamics that emerge from processes on a graph can vary nontrivially based on details of the topology.

When the time comes to implement an agent-based model or system, the software engineer must still work in terms of individual agents and community mechanisms. In model construction, the physical environment itself must be represented computationally. We have identified several work-arounds that can accommodate these constraints, as long as care is taken to avoid certain pitfalls. Specifically,

- Environmental processes can be instantiated as agents. These agents may need to respond to actions from other agents differently than ordinary agents would. If the environment has spatial extent, an adequate representation will typically consist of a network of agents rather than a single environment agent.

- Physical interactions may be modeled as digital protocols. These protocols must model the inconclusive nature of physical actions (in contrast with the deterministic nature of digital communications).

- Acquaintance graphs are a starting point for capturing topology, but should be constructed with attention to the implications of environmentally-based interactions and agent mobility for variation over time, and should be analyzed for graph-theoretic characteristics that may impact the emergent dynamics of interacting agent and environmental processes.

REFERENCES

Brazier, F. M. T., C. M. Jonkers, et al., 1998. *Principles of Compositional Multi-Agent System Development*. Proceedings of the 15th IFIP World Computer Congress, WCC'98, Conference on Information Technology and Knowledge Systems, IT&KNOWS'98, Chapman and Hall.

Brueckner, S., 2000. Return from the Ant: Synthetic Ecosystems for Manufacturing Control. *Computer Science*. Berlin, Germany, Humboldt University Berlin.

Ferber, J. and J.-P. Müller, 1996. *Influences and Reactions: a Model of Situated Multiagent Systems*. Second International Conference on Multi-Agent Systems (ICMAS-96).

Jennings, N. R., 1994. *Cooperation in Industrial Multi-Agent Systems*.

Parunak, H. V., S. Brueckner, et al., 2000. *Distinguishing Environmental and Agent Dynamics: A Case Study in Abstraction and Alternative Modeling Technologies*. Engineering Societies in the Agents' World (ESAW'00), Berlin, Germany.

Parunak, H. V. D., 1987. Manufacturing Experience with the Contract Net. *Distributed Artificial Intelligence*. M. N. Huhns. London, Pitman: 285-310.

Parunak, H. V. D., 1996. Applications of Distributed Artificial Intelligence in Industry. *Foundations of Distributed Artificial Intelligence*. G. M. P. O'Hare and N. R. Jennings. New York, John Wiley.

Parunak, H. V. D., 1997. "'Go to the Ant': Engineering Principles from Natural Agent Systems." *Annals of Operations Research* **75**: 69-101.

Parunak, H. V. D., 1998. What can Agents do in Industry, and Why? An Overview of Industrially-Oriented R&D at CEC. *Cooperative Information Agents II*. M. Klusch and G. Weiss. Berlin, Springer. **1435**: 1-18.

Parunak, H. V. D., 2000. Adaptive control of Distributed Agents through Pheromone Techniques and Interactive Visualization. H. V. D. Parunak, J. Sauter and R. S. Matthews, ERIM CEC.

Pfeifer, R. and C. Scheier, 1999. *Understanding Intelligence*. Cambridge, MA, MIT Press.

Port, R. F. and T. vanGelder, 1995. *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge, MA, MIT Press.

Schöner, G., M. Dose, et al., 1995. "Dynamics of behavior: Theory and applications for autonomous robot architectures." *Robotics and Autonomous Systems* **16**: 213-245.

University of Hannover, 2000. Holonic Manufacturing Systems. *HMS, IMS*, University of Hannover.

Watts, D., 1999. *Small Worlds : The Dynamics of Networks Between Order and Randomness*. Princeton, NJ, Princeton University Press.

Wooldridge, M., N. R. Jennings, et al., 2000. "The Gaia Methodology for Agent-Oriented Analysis and Design." *International Journal of Autonomous Agents and Multi-Agent Systems* **3** (Forthcoming).

AN OBJECT-ORIENTED, AGENT-BASED, SPATIALLY EXPLICIT ENVIRONMENTAL MODEL: A DISCUSSION OF THE APPROACH TO IMPLEMENTING THE SYSTEM

C.C. REWERTS, Engineering Research and Development Center*
P.J. SYDELKO, Argonne National Laboratory
J.H. CHRISTIANSEN, Argonne National Laboratory
J.E. DOLPH, Argonne National Laboratory
A.-M. SHAPIRO, Engineering Research and Development Center
T.N. TAXON, JR., Argonne National Laboratory

ABSTRACT

In this paper we discuss the implementation of an object-oriented, agent-based, spatially explicit simulation of the red-cockaded woodpecker (*Picoides borealis*) (RCW). The goal of this particular modeling effort is to provide a tool for Army land managers who must balance the mission for military training with the protection of the RCW. The goal of the authors is to explore, develop, and use tools for ecological modeling in a wider frame of reference, with this particular effort as a component of the overall ambition. The purpose of this paper is to describe our design goals, why we chose the tools we used, and how we employed them to accomplish the task.

INTRODUCTION

An environmental modeler makes many choices when addressing a problem. The real world, the living inhabitants, and their processes are abstracted to fit the model and modeling environment. To complicate matters, all aspects of modeling technology are in a constant state of change. As technology changes, the modeler is faced with a number of moving targets when choosing the tools, data, and methods from which to build a model of the environment. The environmental processes in the real world are a symphony of interactions. As modelers abstract the real world, they will, for practical reasons, need to focus on a limited number of players and interactions. Meanwhile, other modelers may have focused on other aspects of the real world in their models. Thus, for example, could a wildlife model, a habitat model, and a model of human encroachment be connected to work together?

This paper explains the criteria we used when designing an environmental model that eventually will be multifaceted and applied to assist in land management decisions. We wanted the model to represent entities and processes as they are understood in the real world. We began with an agent-based population model of a rare woodpecker, with plans to include aspects of the environment such as habitat quality and human activities. As we expanded the model, we wanted the option of including models developed by other people, even if the other models were written in a different programming language or ran on a different computer on the Internet. We knew

---

* *Corresponding author address*: Chris C. Rewerts, U.S. Army Construction Engineering Research Laboratory, Engineering Research and Development Center, P.O. Box 9005, Champaign, IL, 61826-9005, e-mail: Chris.C.Rewerts@erdc.usace.army.mil.

that if we integrated other models or more of our own submodels, we wanted to do it in such a way that the various models and processes could be updated through time without requiring a complete overhaul of our system. We wanted a spatially explicit model, to take advantage of geographic information system (GIS) data and functionality. There are several potential users of our model, each of whom will have various data formats or goals for the model; consequently we wanted ubiquitous flexibility. Finally, we wanted a method of modeling such that the effort invested in answering one specific problem could be applied as much as was practical to other, unforeseen problems.

BACKGROUND

The red-cockaded woodpecker (*Picoides borealis*) (RCW) is a resident of the old-growth longleaf pine woodlands of the southeastern United States. The RCW has been listed as a federally endangered species since 1970. The decline in RCW populations has been attributed to loss of habitat and alteration of habitat, primarily due to suppression of fire (USFWS 1985, Walters 1990, Jackson 1994).

United States law places specific requirements on federal land-owning agencies, such as the Department of the Army, for management of federally listed endangered species. The Army cannot harass, harm, kill, or disrupt the natural behaviors of listed species, including the RCW. The Army must take proactive steps to enhance the numbers of RCWs on its lands. Any significant change to activities or construction projects on Army lands must be coordinated with the regulatory agency, the U.S. Fish and Wildlife Service. Because the RCW is known to exist on many Army installations throughout the southeastern United States, its management has been coordinated among many locations to ensure uniform compliance with the law. Formal procedures and decisions are necessary that can benefit from long-term projections of population numbers and responses of the species to human activities. We anticipate that our simulation model will assist the Army in meeting these legal and policy requirements.

The RCW has highly specific habitat requirements. It is the only woodpecker known to excavate cavities for nesting and roosting in living pine trees (Ligon et al. 1986). Furthermore, these pine trees are typically infected with heart-rot fungus (*Phellinus pini*) and tend to be the older trees in the forest (Hooper 1988, Hooper et al. 1991). Older and infected trees can be more easily exploited for cavity excavation than younger trees (Hooper 1988). RCWs depend on their cavities for roosting, nesting, and rearing their young (Ligon 1970). Limited availability of appropriate cavity trees has been thought to be a major factor in the bird's decline (Hooper 1988).

The RCW is also unique in that it is a cooperative breeder. Some of the young males remain on their natal territory for up to seven years as nonbreeding helpers (Ligon 1970, Lennartz et al. 1987, Walters et al. 1988). If the breeding male within the natal territory dies, helpers often inherit the natal territory and become breeders. Helpers may also disperse to a nearby territory to become breeders. Other young males, and nearly all females, disperse during their first year to search for a breeding vacancy (Walters et al. 1992). Because of the RCW's relatively short dispersal distances, the spatial distribution of clusters of cavity trees appears to play an important role in the population dynamics of the species (Engstrom and Mikusinski 1998).

Army land managers must provide high-quality military training while supporting endangered species management and other natural resources objectives. As part of these efforts, environmental data collection and field research have occurred over many years at the larger Army installations. In recent years, the Army has invested in the development of simulation models to help assess the impact of human activities on natural resources. To date, scientific data and computer technologies are often applied to one objective or one facet of the mission at a time. As interdisciplinary approaches and integrated management become more common, decision support systems must be able to combine the various aspects of the ecosystem with the land use mission and land management activities. The Army has launched a large initiative called Land Management Systems (LMS) to promote the integration of technological support capabilities and to solve the varied technical hurdles to increased integration of dynamic simulation models. We hope that our work also helps to meet this Army need.

This paper describes the expandable object-oriented modeling framework we used to implement an agent-based, spatially explicit population model for the RCW. The framework chosen, called the Dynamic Information Architecture System (DIAS), meets our numerous design criteria. It provides dynamic interaction between the population model and other future models and applications. We will project RCW populations dynamically (through space and time) and then expand the model to include other natural processes and various land use and land management influences that are acting within the ecosystem. An earlier application of DIAS produced the Object-Oriented Integrated Dynamic Analysis and Modeling System (OO-IDLAMS). OO-IDLAMS was developed to demonstrate the advantages of an object-oriented architecture approach to integrated natural resources decision support (Sydelko et al. 1998).

APPROACH

We implemented a dynamic RCW population model developed separately by Lechter et al. (1998) within DIAS as the first step in providing a flexible, robust simulation tool for the Army. The Agent-Based Model (IBM) approach was chosen for implementation over a more traditional aggregate population modeling approach, because of its ability to (1) describe the population traits with distributions rather than mean values, (2) represent agent performance and local interactions, and (3) provide a mechanistic rather than a descriptive approach to modeling (DeAngelis and Rose 1992). Furthermore, the model described by Letcher et al. (1998) is spatially explicit, accounting for the importance of spatial distribution to RCW population dynamics.

Currently, the model is being validated using Fort Benning, Georgia, as a case study. Plans are currently underway to incorporate a forest growth model, a RCW foraging and nesting habitat model, and military activity projections within DIAS to further provide modeling and simulation to support decisions impacting the RCW at Fort Benning.

Implementation of RCW Population Model

The dynamic behaviors (or processes) described in the Letcher et al. (1998) model were broken out into five distinct "submodels" for implementation in the DIAS RCW application: breeding, competition, dispersal, fledgling role change, and mortality. These Letcher et al. submodels address specific behaviors (DIAS "Aspects") of specific classes of objects (DIAS "Entities").

We utilized Lechter's model by capturing individual birds, their population groups, and their territories as object entities, and by assigning separate, specific properties and behaviors to each entity.

The Entity object classes developed for the RCW application (RCWIndividual, RCWTerritory, RCWPopulation) contain the attributes that describe the state of the environment throughout a simulation. These Entities also contain links to the processes (the Letcher et al. [1998] submodel behavior) via their Aspects.

Design Criteria for RCW Model

The following sections describe characteristics of DIAS that met our design criteria for this application.

Modularity

Creating a modular, compartmentalized structure allows us to more closely mimic the associations and relationships that exist in the real world and provides a level of modeling flexibility that is well beyond what we could achieve using traditional modeling practices (GIS-based or procedural). By defining the object entities and articulating how they interact through process code, we provide a foundation upon which we can logically articulate and include other relationships and connections in the future. While the application development is very specific (in this case, RCW population dynamics), the application component development is generic. This avoids the pitfalls of "monolithic" models where development proceeds as a series of submodel additions, one on top of another, resulting in a highly unwieldy and inflexible model structure that is time-consuming and costly to expand.

Furthermore, because the RCW Entities and their associated behaviors are developed within a framework that already houses a diverse array of environmental and non-environmental objects (Hydrologic, Atmospheric, and Vegetative), making the connection to other models and processes is easier. We do not have to build everything from the beginning — we can utilize existing objects from the DIAS library and add attributes and behaviors as new applications dictate.

It is important to note that this modular approach allows the user to readily plug in alternative RCW behaviors (submodels) without time spent recoding or reworking the existing application. One would simply substitute a new algorithm for one of the behaviors (write a new method for the Process object), but the connections between the Entities and the events that trigger behaviors remain the same. Thus, altering models or substituting a new model is not disruptive to the overall system, and it is therefore more efficient and more cost-effective to develop applications in this type of framework.

Systems that serve environmental managers need to include opportunities for updating and changing information as knowledge is advanced through related research and monitoring. Overall, DIAS provides an efficient framework within which to bring together disparate data and software for integrated resource planning. It is flexible and robust enough to readily assimilate new models into an existing application.

Code/Object Reusability

Entity class objects need only be designed and built once. Any future application or model that requires use of RCWs can utilize these existing objects as is, or by simply adding any additional attributes or process linkages that may be required for a new application. In this way, objects continue to mature, but are not recoded, as is often the case with more conventional model development. This obviously points to the need to thoroughly design Entity objects up front and to make certain that they are generic enough to be utilized by diverse domains. The reusability of objects and code can and does represent significant cost savings. Oftentimes, applications can be more quickly developed in the DIAS framework because of the benefits of object/code reusability.

Expandability

Our intention was to design the RCW model application so that future implementations could predict RCW populations by incorporating various land use and land management influences acting in the ecosystem. Again, because the approach taken was modular, and the behaviors are “contained” and distributed to the appropriate Entities, we were able to build a system that will not require recoding to incorporate management impacts. The RCW state — whether Individual, Territory, or Population — will be affected by changes in the environment, regardless of what is producing the environmental change. These cause-effect linkages are built into the system. For instance, a change in vegetation whether produced by fire, management practice or disease, will produce a resulting change in RCW habitat that could, in turn, affect the bird. A multitude of influences can impact the natural environment. These influences change the state of the environmental objects first (directly), and then this effect is propagated down through the birds (or any other object in the simulation) as appropriate. In this sense, linkages are not “hard-coded,” but occur naturally, reflecting our current scientific understanding of interrelationships. The birds register to receive events that are “of interest” to them. If an environmental parameter changes, regardless of what changed it, and this parameter is of interest to the bird, the bird will receive notice of the change and will respond accordingly.

Adding in management influences to the existing application will simply be a matter of articulating the management practice/procedure and coding a model that reflects the practice (most likely a CourseOfAction type [COA] model). As management affects the state of the environment, these changes will impact the RCW.

Expressiveness of DIAS COA Objects

We used the FACET (Framework for Addressing Cooperative Extended Transactions) object suite to code the breeding behavior as a COA model. FACET objects are used to represent complex interactions between objects, or “agents,” in a simulation. For this application, the agents are individual RCW birds. The breeding behavior COA is basically a flowchart of agent actions, and acts in the same manner as a Process object to implement behaviors of Entities. While this particular COA is not very complex, it illustrates the ability of DIAS to handle social process models that involve cooperative behavior between agents.

FACET COA type models will be critical for the development of land use and land management plans to assess the interplay of human impact on the RCW environment.

System Dynamics and Feedback Mechanisms

Traditional GIS-based systems are static in nature and do not lend themselves very well to dynamic, inter-process modeling. In general, most models and applications are designed to operate independently, even though effective decisions call for assessing several components of the ecosystem simultaneously, in terms of their relationship to each other as well as how they affect management decisions. DIAS provides a framework for developing applications that address inter-process dynamics in a highly realistic way. DIAS allows us to articulate the dynamics of an ecosystem much more closely to the way in which we understand them to operate in the real world. At the same time, it does not impose a single worldview (one discipline's view) on the development of the application.

DIAS has an event-driven simulation environment. Events in a simulation invoke the behaviors of the Entities at the appropriate time, following the dynamics of the system, as we have articulated them. We specify what objects will be in the simulation (the "playing pieces"). These objects carry their state as attributes. At any time during a simulation, we can pause and assess the state of the environment by evaluating these attribute values. Any process operating in the simulation can potentially modify the state of each object. These processes can be associated with internal or external models, natural/physical models, or management plans. In this way, the DIAS inter-process flows are realistic and reflect the dynamics of the system as we understand them.

CONCLUSION

We have described using DIAS, an object-oriented, event-driven architecture, as an enabling technology to construct a flexible framework for simulating RCW populations, and eventually, habitat and management factors that will affect the populations. Dunning et al. (1994) point out that spatially explicit population models are not a panacea for predicting the locations of individual animals with a high degree of accuracy. Rather, simulation efforts are building a bridge to link ecological research with applied fields such as wildlife management and conservation biology (Turner et al. 1995). Modeling exercises can serve to collect knowledge, data, and theory about ecological processes to improve understanding and communication; serve to screen potential management actions; and help to identify gaps in knowledge or data needed to assess the impacts of human activities

REFERENCES

DeAngellis, D.L. and K.A. Rose. 1992. Which individual-based approach is most appropriate for a given problem? In DeAngelis, D.L. and Gross, L.J. (Eds.). *Individual-based Modeling* *and Approaches in Ecology: Populations, Communities, and Ecosystems*, Chapman and Hall, New York, New York USA, pp. 67-87.

Dunning, J.B., D.J Stewart, B.J. Danielson, B.R. Noon, T.L. Root, R.L. Root, R.H. Lamberson, and E.E. Stevens. 1995. Spatially explicit population models: Current forms and future uses. Ecological Applications 5(1): 3-11.

Engstrom, R.T and G. Mikusinski. 1998. Ecological Neighborhoods in Red-Cockaded Woodpecker Populations. Auk 115(2): 473-478.

Holling, C. S., editor. 1978. *Adaptive Environmental Assessment and Management,* John Wiley, New York, New York, USA.

Hooper R.G. 1988. Longleaf pines used for cavities by red-cockaded woodpeckers. J. Wildl. Manage. 52: 392-398.

Hooper, R.G., M.R. Lennartz, and H.D. Muse. 1991. Heart rot and cavity tree selection by red-cockaded woodpeckers. J. Wildl. Manage. 55: 323-327.

Jackson, J.A. 1994. Red-cockaded woodpecker (*Picoides borealis*). The birds of North America, number 85. The American Ornithologists' Union, Washington, D.C., USA, and The Academy of Natural Sciences, Philadelphia, Pennsylvania, USA.

Lennartz, M.R., R.G. Hooper, and R.F. Harlow. 1987. Sociality and cooperative breeding in red-cockaded woodpeckers. Behav. Ecol. Sociobiol. 20: 77-88.

Ligon, J.D. 1970. Behavior and breeding biology of the red-cockaded woodpecker. Auk 87: 255-278.

Ligon, J.D., P.B. Stacey, R.N. Conner, C.E. Bock, and C.S. Adkisson. 1986. Report of the American Ornithologists' Union Committee for the Conservation of the Red-cockaded Woodpecker. Auk 103: 848-855.

Letcher, B.H., J. A. Priddy, J.R. Walters, and L.B. Crowder. 1998. An individual-based, spatially-explicit simulation model of the population dynamics of the endangered red-cockaded woodpecker, *Picoides borealis*. Biological Conservation 86: 1-14.

Turner, M.G., G.J. Arthaud, R.T. Engstrom, S.J. Hejl, J. Liu, S. Loeb, and K. McKelvey. 1995. Usefulness of spatially explicit population models in land management. Ecological Applications 5: 12-16.

Walters, C. J. 1986. *Adaptive Management of Renewable Resources,* McMillan, New York, New York, USA.

Walters, J.R., P.D. Doerr, and J.H. Carter III. 1988. The cooperative breeding system of the red-cockaded woodpecker. Ethology 78: 275-305.

Walters, J.R. 1990. Red-cockaded woodpeckers: a "primitive" cooperative breeder. Pages 67-102 in P.B. Stacey and W.D. Koenig (Eds), *Cooperative Breeding in Birds: Long-Term*

Studies of Ecology and Behavior, Cambridge University Press, Cambridge, United Kingdom.

Walters, J.R., P.D. Doerr, and J.H. Carter III. 1992. Delayed dispersal and reproduction as a life-history tactic in cooperative breeders: fitness calculations from RCWs. American Naturalist 139: 623-643.

DISCUSSION: ENVIRONMENTAL PROCESSES

D. SALLACH, University of Chicago, Moderator

Robert Reynolds [to Chris Rewerts]: In your habitat model, are different aspects of the habitat considered agents as well? For example, your birds are part of a food chain: they have predators and they prey on others. And the Army itself could be viewed as an agent, allowing the model to predict the impact of various maneuvers on the bird population.

Chris Rewerts: Yes, although the predator-prey issue is not a major one in this case, if you maintain the habitat properly. For instance, if the birds are in big, open areas, then, yes, kestrels come in and get them. Another predator is a snake that climbs up the tree and gets into the hole. The birds take care of that by making the sap run, which is an irritant to the snakes. Another problem is other creatures who try to use the holes. The site staff manages that problem by putting restrictor plates on the holes so only the birds can fit through — they're not afraid to use hardware out there.

The modeling of training is something that I thought I would be doing soon. As it turns out, though, the field studies done so far show that the impacts of the noise, maneuvers, and smoke on the birds is minimal. The Army has management guidelines — 1996 is the latest iteration of those — that say how it can train around the nesting sites. Units are required to stay 200 meters away, I believe, from a nesting site during maneuvers.

Modeling Army training is a bit problematic, because, for security reasons, they don't want to tell us much about what they do. So we look at Army training the same way we look at any other animal: we take measurements then extrapolate upon the landscape. Another project I worked on did just that — it looked at impact probabilities for different types of training events. We could say, for example, "In this footprint of area they're making this type of noise and they're using smoke and obscurance," and we could characterize the time and the spatial location of the training attributes. So, yes, we can get to that point, but so far we've prioritized in different directions based on what we've been learning about the effects of training.

Pamela Sydelko: I'm Pam Sydelko from Argonne. I've been working with Chris on this project. One of the interesting things is that there are many different objectives for managing the land. We run into this issue with the military, but it also applies to any land-managing agency. One of the things the Army has a real struggle with, at Fort Stewart or Fort Benning, for instance, is that primarily they are just managing for this bird. In some installations, the Army may be managing for a bird but also has a water quality or erosion problem. And there are many stakeholders that have to be considered.

I'm interested in how we can represent those different stakeholders as agents, to feed into our modeling suites that look at long-term satisfaction with land management decisions and the tradeoffs involved. There's only so much money you can spend in an installation to manage all of these things. If you make erosion better, you might not have enough money to maintain habitat for an endangered species. Stakeholders drive many of these decisions. And the stakeholders are as dynamic as the environment itself. For example, we're putting dynamics, such as a drought, into the environment as we do this modeling. But what if all of a sudden the red-cockaded woodpecker is not the issue anymore, but an endangered tortoise instead. Your management plan wasn't concerned with soil compaction during training because that's not an issue for the bird, but it's important to the tortoise. It would be helpful to have some agent-based modeling ideas about this type of dynamic — and this would apply for national park and forest management, as well — because these changes in policy and focus happen all the time.

Rewerts: That's a very good point. It sounds as though we're focusing on one species to exclusion. But really our goal is to move toward ecosystem management, even though that science is still wide open.

Christopher Langton: I agree that that's a very important point. We tend to believe that all we have to do is get the science right, and then we can go to the policymakers and say, "Here's the truth," and they'll say, "Fine." But it's not that easy. The policymakers have to consider the special interest groups. This has come up in some models for the Columbia River, where people are trying to get salmon to come back into the river. A solution may be known, but implementing the solution would create conflicts with the special interests. And so it's incredibly important to consider, and probably even put into our models, all of the special interest groups that are working with the policymakers on these issues.

Benjamin Schoepfle: I'm Benjamin Schoepfle from Argonne. Pam's comments are very important. I've done similar work on military bases in the South helping to locate corridors for tanks in relation to a species of woodpecker. One issue is that many of the base commanders have short tenures of about two years, so a new commander may not know the history, especially the finer nuances of these very spatially-tuned policies. I think a context of history is important for modeling: it's important not only to look toward the future, but also to replicate the past. Modeling can give a historical record, and record-keeping is extremely important because many managers are on the job for such a short time.

Randal Picker: I'm Randy Picker of the University of Chicago Law School. To follow up, it's not obvious to me that that an agent-based framework is the right way to capture interactions between policymakers — who are the people I spend a lot of my time with — and the kinds of models you are building. The agent-based framework strikes me as particularly good for dealing with population situations, but I'm not sure it's the right tool when you've got principally five actors who interact with each other, maybe on a repeat basis, in a very small setting.

Rewerts: There's a couple of different ways of looking at this, and I'm going to pick the easy one. One reason to model is for consensus-building. All models are based on assumptions. As you put these things together and see how they operate, you're testing your theories of what you think is actually going on out there in the ecosystem. If everybody agrees that you've got a valid theory being played out in this simulation, you can make policy decisions based upon it. If you're not using a tool like this to help make decisions, you're just flying by the seat of your pants. But you're talking about whether or not to put various decision-makers in as agents?

Picker: Exactly. It's not a question of whether your model does a good job of capturing what's going to happen to woodpeckers. I suspect it does, and it seems to me the Army's doing exactly the right thing in looking to you. But my question is how to get at how some other group is going to react to that, whether the Sierra Club is going to decide, "Woodpeckers aren't important; it’s snail darters that we care about.” I don’t think an agent-based model is necessarily the best way to get at that.

Sydelko: Actually, I think one of the things that we’re talking about is not so much the tail-end arguments of the policymakers looking at the decisions made by the red-cockaded woodpecker model and deciding whether it does what they expected. What would be nice is more of a front-end engine, a suite of models, that actually balances these many factors and concerns in ecosystem management. For example, we have vegetation models that show training impacting vegetation, the woodpecker reacting to the vegetation changes, and erosion occurring because of the vegetation changes.

It’s a challenge to make a management plan that 20 years out is going to satisfy all conditions. Many of these conditions are generated not by land managers, but by other parties. It would be helpful to have a series of scenarios based on various players’ perspectives and what they value. We had a workshop with a group of land managers and military people to try to determine what they value most about the land. Is it protecting endangered species or controlling erosion? Is training most important? We wanted to see whether we could generate a finite set, perhaps even using genetic algorithm tools, of good possible scenarios, run them through the model, and determine how well we meet the criteria. This might give us an idea of where to start. There’s a lot of choices to be made about land management, not only about *what* to do but *where* to do it. If we had some idea of where we could use the agents and consensus-building to drive the scenario-building, that would be worth looking at.

Energy and Infrastructures

BUILDING ELECTRICITY MARKET PARTICIPANT STRATEGIES WITH ADAPTIVE AGENTS

C. RICHTER, Alstom ESCA Corporation, Bellevue, Washington*

ABSTRACT

Strategies that reduce risk and increase profit are important to electricity market participants. Financial instruments that hedge against uncertainty and manage risk are important components of such strategies. One such vehicle is the fixed transmission right (FTR) on a transmission path. The FTR holder is entitled to the transmission congestion rents collected on that path. However, if congestion occurs in an unanticipated direction, the holder of an FTR is obligated to pay. An interesting compromise derivative is the FTR option. For a premium, the FTR option gives its holder the right to the congestion rents if they are positive, without the obligation to pay if they are negative. This paper builds on previous work [13] and discusses how adaptive agents can be used to develop strategies that help the market participant value and utilize FTR options as part of a comprehensive market strategy.

INTRODUCTION

Electricity markets are being re-regulated to promote competition [3,4,5]. Entry barriers to new market participants are being removed; formerly monopolistic utilities have been forced to separate their generation, transmission, and distribution into independent companies known as GENCOs, TRANSCOs, and DISTCOs, respectively. New market entities are emerging (e.g., energy service companies (ESCOs) that purchase wholesale electricity and repackage it for resale to the end-consumer). Most of these market participants are profit-based and benefit from well-designed bidding strategies. The strategies must each be designed for a specific set of market rules, which vary from region to region.

One promising technique for developing strategies involves intelligent or adaptive agents. In addition to the promise these techniques hold for developing profitable strategies for market participants, the same techniques can be used by market regulators and operators to test the market structure, find loopholes that might allow gaming, and improve the stability of the market. The technique begins with building a simple model of the market participants. Part of the rules making up the agent must be adaptable. A training, learning, or evolutionary process (e.g., a genetic algorithm, weight updating) presents the agents with various sets of inputs and rewards or punishes them based on their performance, as judged by a variety of metrics.

Electricity flows according to the laws of physics over a network that is subject to congestion. Real-world market operation, pricing, and product offerings in many markets are designed to reduce the impact of transmission congestion. Simulations using agents to model market participant behavior or develop strategies for a competitive electricity market should include a model of the electricity network if

---

* *Corresponding author address*: Charles W. Richter, Jr., Alstom ESCA Corporation, 11120 NE 33rd Place, Bellevue, WA 98004; e-mail: charles.richter@alstom.esca.com.

they are to produce results of interest to many of the companies seeking to make profit in the new world of electricity.

ELECTRICITY NETWORKS AND COMPETITIVE MARKETS

In some areas, organized exchanges centralize the trading of electricity products. Some areas calculate a region-wide price for electric energy, while others develop locational marginal prices (LMPs) for each node in the network [6]. Some exchanges remove contract default risk by becoming an intermediate partner to all trades. Centralized exchanges facilitating physical (intended for delivery) power trading must interface with an independent system operator (ISO), responsible for secure electricity network operations. Other regions rely on decentralized bilateral trading. Bidding and operational strategies for bilateral trading may vary widely from those used for centrally allocated markets.

Electricity markets should promote economic efficiency and secure power system operation. This means balancing the simplifying assumptions required for creating a liquid market with the reality of power system operation. Power flows through a network according to Kirchoff's laws, making it difficult to predict how a particular transaction might impact the network without considering all other transactions. Additional complexity comes with the market treatment of supportive services (e.g., reactive power and transmission) related to the physical flow of electricity, which may vary widely with the market implementation.

Real-time electricity prices can be largely impacted by transmission congestion. Transmission congestion can prevent the transport of electricity over the network, thereby causing the price of electricity at the demanding end to rise, and the price at the sending end to fall. Congestion costs are often measured by the differences in sending and receiving end LMPs. This amount is collected from those transporting electricity across the congested path, thus providing an incentive to shift generation to the other side of the congested transmission path.

To reduce transmission congestion uncertainty, interested participants can purchase fixed transmission rights (FTRs) from an injection point to a delivery point. The FTR cannot guarantee physical delivery, but can make the holder financially whole by entitling the holder to congestion rents collected along that path in that direction. If network congestion requires a supplier to pay large collection rents to deliver the product, the FTR allows that supplier to offset the charges. Unfortunately, if other transactions on the network cause the power to flow in the opposite direction (counter-flows), the FTR holder is obligated to pay the negative congestion rent.

To further reduce the transmission congestion risk, many market participants have expressed an interest in FTR options. Options on FTRs allow the holder to collect positive congestion rents, but do not obligate them to pay negative congestion rents. Several markets regions are on the verge of adopting FTR options. Several more markets are considering the sale of flow gate rights, which share the nonobligatory nature of FTR options. The FTR option should be considered a valuable part of market strategies.

FIXED/FINANCIAL TRANSMISSION RIGHTS

To illustrate how FTRs are utilized, an example with and without an FTR is presented. Losses are ignored in this example. Suppose that generator *G* is able to generate electricity for a cost of \$10/MWh. *G* wants to lock-in a price for generation and has bilaterally contracted to supply 100 MW for an hour to load *L* for $12.50/MWh. *G* (Node A) and *L* (Node B) are connected through a meshed network as shown in Figure 1. If the network is not congested and LMPs at nodes A and B are $10/MWh, then *G* generates the 100 MWs, and *L* consumes 100 MWs. *G* collects $250 more from the bilateral contract than it would have collected from selling to the spot market. *L* paid $250 more than it would have paid from the spot market. Even though in this scenario *G* profited $250 and *L* "lost" $250, both *G* and *L* are happy because they locked in their price ahead of time, thus reducing their exposure to the uncertainty of the energy spot price.

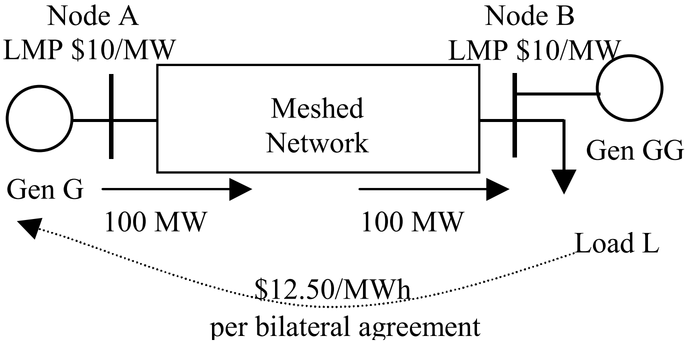


FIGURE 1 Simplified network, without congestion

However, if the network between Node A and Node B is congested, as shown in Figure 2, the LMPs at both of the nodes will be different. *G* may not be able to physically deliver its generation, and may have to purchase generation at Node B to meet its obligation to *L*. Because generation at Node B could be quite expensive, *G* is now exposed to considerable risk.

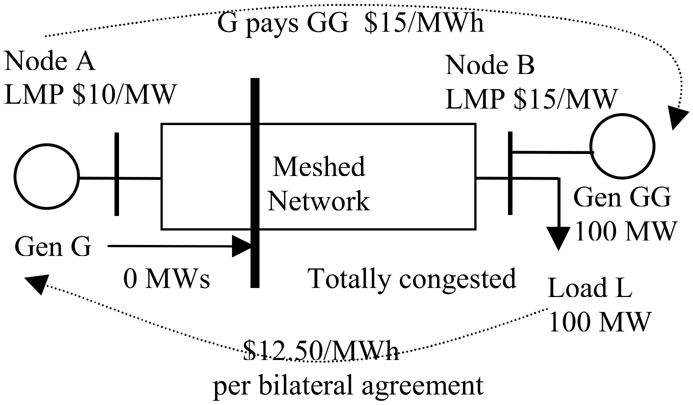


FIGURE 2 Simplified network with congestion (decentralized dispatch)

An example will help illustrate this exposure to price uncertainty. If the path from A to B was completely congested, as shown in Figure 2, and LMP at B was $15/MWh, then *G* would be forced to purchase power at a higher LMP at B (e.g., $15/MWh) and deliver it to *L,* who would pay only $12.50/MWh. This means that *G* would be losing $2.50/MWh. If LMP at B was lower than at A, then congestion is in the opposite direction and *G* will generate from its own unit. Note that in a less congested scenario, *G* may be able to provide some power from its own generator and would purchase the remainder from *GG*. Table 1 shows *G*'s profit in a decentralized environment for the bilateral transaction with *L* under scenarios with varying LMPs at Node B. In the table, *G*'s profit is calculated as:

$$profit \ = revenue \ - total \ costs \ ,$$

where

revenue = *L*'s payment to *G*,
total cost = (*G*'s gen cost) + (*G*'s payment to *GG*),
G's gen cost = ($10/MWh) × (*G*'s gen), and
G's payment to GG = (LMP@B) × (*GG* MWs).

TABLE 1 *G*'s Profit for a 1-h 100-MW Bilateral Contract under Four Scenarios

LMP A ($/MWh)	LMP B ($/MWh)	*L* Pays *G* ($)	*G* gens (MW)	*G*'s gen ($)	*G* pays *GG* ($)	*G*'s total Cost ($)	*G*'s "profit" ($)
10	20	1,250	0	0	2,000	2,000	-750
10	15	1,250	0	0	1,500	1,500	-250
10	10	1,250	100	1,000	0	1,000	250
10	5	1,250	100	1,000	0	1,000	250

In a centrally dispatched market, each generator is dispatched by the ISO, which is essentially running a power flow that considers physical network limits, bids, and offers. The LMPs are a result of the central dispatch and allocation process. The ISO generally would not consider any bilateral agreements between *G* and *L* in the centralized dispatch. The ISO would collect and distribute the money associated with the centrally allocated transaction. Table 2 demonstrates results for the same size transaction and the same nodal price scenarios when coordinated by the ISO.

TABLE 2 Settlement for *G* to *L*, 1-h 100-MW Transaction under ISO

LMP A ($/MWh)	LMP B ($/MWh)	*L* Pays ISO ($)	*G* Gens (MW)	ISO pays *G* ($)	ISO pays *GG* ($)	*G* pays *L* ($)	*G*'s cost ($)	*G*'s profit ($)
10	20	2,000	0	0	2,000	750	750	-750
10	15	1,500	0	0	1,500	250	250	-250
10	10	1,000	100	1,000	0	-250	1,250	250
10	5	500	100	500	0	-500	1,000	250

Table 3 demonstrates generation and payments allocation in the presence of congestion-induced transmission limits. The ISO is clearing (or not clearing) offers submitted by generator *G* at Node A (100 MW offer @ $10/MWh) and by generator *GG* at Node B (100 MW @ $15/MWh). Load L requires 100 MW in this example. The table illustrates generation levels and settlement under different levels of congestion-limited transmission. When LMP A is greater than LMP B, *G* is not responsible for any congestion payments (CP's). When the LMPs at all nodes are the same, there is no congestion in the network. However, when LMP A is less than LMP B, *G* is liable for congestion payments (implicitly defined in the previous table) [(LMP B – LMP A) × MW] as *G* has not hedged against this congestion.

TABLE 3 Effects of Congestion on *G*'s Profit without FTR

Cong Limit A-B (MW)	Supply Cleared @ *G* (MW)	LMP @ A ($/MWh)	Supply Cleared @ *GG* (MW)	LMP @ B ($/MWh)	*L* Pays ISO ($)	ISO Pays *GG* ($)	ISO Pays *G* ($)	*G* pays ISO CP ($)	*G*'s Profit ($)
100	100	10	0	10	1,000	0	1,000	0	250.0
75	75	10	25	15	1,500	375	750	375	125.0
50	50	10	50	15	1,500	750	500	250	0.0

To hedge against paying the uncertain LMP at B under congestion, *G* can chose to purchase an FTR of 100 MW (in this example, the price is assumed to be $1/MW) to ensure that it is hedged. *G*'s profit (including the cost of purchasing FTR and the profit of selling electricity at $2.50 more than generation costs) is shown in Table 4. LMP A is kept the same, while LMP B is reduced from $15 to $5. The congestion limit is assumed to be 100 MW.

TABLE 4 *G*'s Profit with an FTR

LMP @ A ($/MWh)	LMP @ B ($/MWh)	FTR (MW)	FTR ($/MWh)	*G* Pays ISO CP ($)	ISO Pays *G* CR ($)	*G*'s Profit ($)
10	15	100	1.0	500	500	150
10	10	100	1.0	0	0	150
10	5	100	1.0	500	0	-350

When LMP B is greater than LMP A, *G* has a profit of $150, as it collects $500 in congestion rent and has paid $100 to purchase the FTR, and has a profit of $250 for the bilateral contract with *L*. When there is no congestion, then it also stands to make a profit of $150. However, when LMP B is greater than LMP A, then *G* stands to lose $350 because of the counter flow in the opposite direction of the purchased FTR. Hence as the price differential increases, and the LMP A is larger, then *G* will start to make large losses although it has purchased an FTR and hedged against congestion. *G* is always financially liable for any congestion charges because of counter flow in the opposite direction of the FTR purchased. FTRs come with their own risk. If power flows in an unanticipated direction (counter-flows), the FTR holder is obligated to pay the negative congestion rent.

FIXED TRANSMISSION RIGHTS OPTION

A fixed transmission rights option (FTRO) gives its holder the right to the congestion rents when power flow is in the indicated direction, without the obligation to pay when power flows in the wrong direction. A party "writes" an FTRO in return for the premium, or FTR option purchase price, and is then obligated to collect/pay congestion rents when the flow is opposite the direction anticipated.

Figure 3 shows the profitability of an FTRO to the congestion rents from point A to B. The purchaser has paid a premium (e.g., $1) to the FTRO writer, thereby giving the purchaser the right to positive congestion rents (e.g., $0). The purchaser has reduced risk by limiting losses to the premium. The right side of Figure 3 shows what happens from the FTRO writer's point of view. The FTRO writer receives the premium for assuming the risk and is obligated to pay the congestion rents when they are negative.

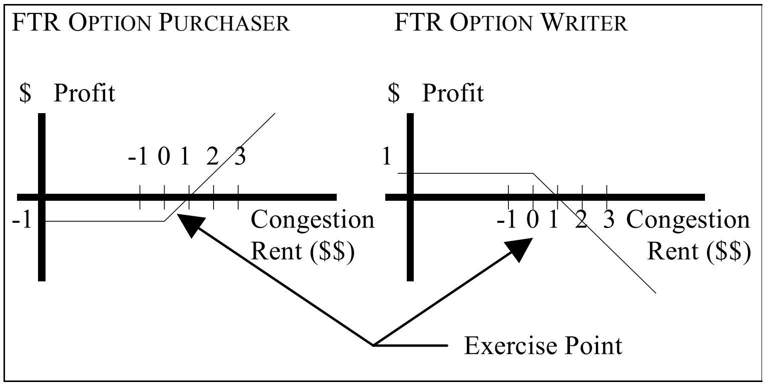


FIGURE 3 FTRO profitability

The FTRO price should reflect the value of the FTRO to the potential holders. The worth of an FTRO may vary from trader to trader because of risk preferences, makeup of portfolios (collection of assets and contracts), etc., but is largely dependent on the uncertainty in the underlying asset. Large uncertainties in the direction of flow on the FTR path and in the amount of congestion rent to be collected translate to high premiums (FTRO prices).

The method used to value options must consider the idiosyncrasies of electricity. Many permutations of Black-Scholes have emerged. Another approach that can help identify an upper bound on the value of the FTRO is to determine expected monetary value (EMV) and value at risk (VAR) under scenarios of interest (see Figure 3).

FTR VERSUS FTRO

The FTRO is an additional tool for managing risk. Here the only FTR example is augmented with FTRO for comparison. Revisiting the simple system depicted in Figure 1, Table 5 provides an example of hedging with FTRO. The FTRO protects *G* from congestion charges because of counter flows. The profit of *G* if it purchases an FTRO from Node A to Node B is shown in Table 5. The congestion limit is assumed to be 100 MW.

From Table 5, when LMP A is greater than LMP B, then *G*'s profit is $100, which is the profit of selling electricity to *L* minus the cost of purchasing the FTRO. This is also the case when there is no congestion, and when LMP A is greater than LMP B. Hence with an FTRO, *G* is not responsible for any counter flow from Node B to Node A.

TABLE 5 *G*'s Profit with an FTRO

LMP A ($)	LMP B ($)	FTRO (MW)	FTRO ($/MWh)	*G* Pays ISO CP ($)	*G* Gets from ISO CR ($)	*G*'s Profit ($)
10	15	100	1.5	500	500	100
10	10	100	1.5	0	0	100
10	5	100	1.5	0	0	100

G may decide to use both an FTR and an FTRO from Node A to Node B to hedge its financial risks. The question for the market participant becomes not which instrument to use, but rather how much of each should be used to achieve the desired risk to premium ratio.

AGENT-BASED TECHNIQUES

Developing market strategies that work well in many market scenarios can be a difficult task. Theoretical macroeconomic assumptions do not always hold (e.g., rational behavior) in a microeconomic setting. One way to develop rules is to combine experimental economics with adaptive agent techniques. The agent can be (by some people's definitions) autonomous, intelligent, and adaptive (meaning that it's behavior changes over time based on the input-output successes and failures that the agent has encountered in the past). Computerized agents (consisting of evolving or adapting rules that process state information) can represent market participants. Markets are simulated in which the adaptive agents are allowed to buy and sell and test strategies. Even without a complicated model for each individual agent, interesting behaviors can be observed when many of these agents are allowed to react in an environment.

The agents (or the strategies of which they consist) can be developed to model various problems that market participants face in the evolving deregulated market place. For example, agents may learn a set of rules for producing bids or offers for forward electricity. The agents may evolve rules for hedging against electricity network congestion by valuing and bidding on or offering products like FTRs or FTROs. The rules can be combined to include responses for either problem, depending upon which inputs are presented to the agent. In most cases, agents produce a response or an action when presented with a set of inputs or state information. Possible actions for an agent representing an agent playing in the deregulated electricity market place might be as shown in Table 6. Some of the many possible inputs or environment variables that agents may wish to observe as part of their strategy (depending on the experiment) are as follows:

- Forecasted price,
- Forecasted demand,
- Lowest and highest bid and offer observed during previous negotiations,
- Average market price,
- Measure of market power,
- Measure of market depth,
- Scheduled outages,
- Forecasted network congestion,
- Fuel costs, and
- Competitor's historical actions.

TABLE 6 An Example of an Agent Market Participant Strategy Action

Take Action	Yes/No	Time Last Performed
Request load forecast?		
Request price forecast?		
Update pricing model?		
Perform unit commitment?		
Run trade-mix/risk optimizer?		
Negotiate fuel contracts?		
Initiate cent. market purchase?		
Accept cent. market purchase?		
Initiate cent. market sale?		
Accept cent. market sale?		
		With whom?
Initiate bilateral energy purchase?		
Accept bilateral energy purchase?		
Initiate bilateral energy sale?		
Accept bilateral energy sale?		
Initiate FTR purchase?		
Accept FTR purchase?		
Initiate FTR sale?		
Accept FTR sale?		
Etc.		

GP-Automata

One way to process the state information to develop the input-output rules is through GP-Automata. GP-Automata combine finite state automata with genetic programs (GPs). They were first described as such by Ashlock [11] and were used by Ashlock and Richter [12]. The typical finite state automaton specifies an action and "next state" transition for a given input or inputs. With only one or two binary inputs to work with, it can be fairly simple to develop a finite state diagram to cover the possible input/output relations. When the number of inputs is large, the task is much harder. The number of transitions needed to cover all possible combinations of inputs grows exponentially (e.g., 10 inputs each having 5 possible values would require 5^{10} transitions). This is where genetic programming comes in. The GP-trees are bandwidth compressors. GP-Automata uses them to select which inputs to consider and to perform computations involving these inputs. Table 7 gives an example of a GP-Automaton.

TABLE 7 A Four-State GP-Automaton

	IF ODD		IF EVEN		
State	Action	Next State	Action	Next State	GP (Decider)
1	14.5	1	U	1	lte (mul(10, abs (hbb))
2	*	1	37	3	ite(max(10, asb), hbb, lbb)
3	12	2	5	1	avg (5, abb)
4	U	3	*	2	47
Initial Action	24	Initial State	2		

Reading the rule encoded in the GP-Automaton in Table 7 is fairly simple. The automaton begins by bidding the number in the "initial action" field. Following the initial action, the "initial state" indicates which state is used next (in this case, 2). The GP-Automaton in the table has four states. Coupled with each of these states is a GP-tree termed a decider. When executed, the decider returns a value between 0 and 100. On the basis of that returned value, one of the following two things will happen: (1) if that value is even after truncation, the action listed under "IF EVEN" is taken and we move to the next state listed under "IF EVEN"; (2) if the returned value is odd after truncation, then we use the action and next state listed under "IF ODD." The "action" is the number listed in the action field of the automaton, with two exceptions. The first exception is the "U," which indicates that the value returned by the decider should be taken directly as the action. The second exception is a " *," which indicates that further computation is necessary, and, hence, the GP-Automaton refrains from acting immediately. Instead, it immediately moves to the next state. This gives rise to the possibility of complex (multistate) computation as well as infinite loops. To prevent infinite loops, one can specify a maximum number of " * "s to be honored, after which a valid action is selected at random.

The GP-Automata population evolves as in any GA. After selecting parents, as described in Part 4, offspring are produced using crossover and mutation. Crossover for the GP-Automata involves selecting (with a uniform probability) a crossover point ranging from zero to the number of states. We then copy parent1's states from zero to the crossover point to child1, and parent2's states to child2. Following the crossover point, child1 gets parent2's state information and child2 gets parent1's state information (including the associated decider). Before replacing less fit members of the population, each child is subjected to various types of mutation.

SUMMARY

Electricity derivative instruments are an essential part of strategies that enable market participants to maximize/increase profit and manage risk in the deregulated electricity industry. Accurate modeling of the transmission network and the effects of congestion are becoming increasingly important as the transmission systems become stressed by the added transactions under competition. A financial hedge against congestion is the FTR; for a premium, the FTR's downside risk can be eliminated through options. This paper presents the FTRO in an easy to understand manner. It suggests that an agent-based framework is a useful tool in developing comprehensive market participant strategies, and for populating simulations with behaviors that mimic realistic market behavior. The proposed framework can be applied to forward as well as future markets. Among the challenges in operational implementation of the presented model are predicting spot market prices and demand, measuring and quantifying other state information for the agents' use, and designing easy-to-understand adaptive data structures that can encapsulate good market rules.

REFERENCES AND BIBLIOGRAPHY

[1] Hull, J.C. *Options, Futures, and Other Derivatives*, USA: Prentice Hall, 1997.

[2] Marshall, J. *Futures and Option Contracting: Theory and Practice*, South Western Publishing, 1989.

[3] Kumar, J., and G. Sheblé, "Framework for Energy Brokerage System with Reserve Margin and Transmission Losses," 1996 IEEE/PES Winter Meeting, 96 WM 190-9 PWRS.

[4] Sheblé, G., "Electric Energy in a Fully Evolved Marketplace," Presented at the 1994 North American Power Symposium, Kansas State University, KS, 1994.

[5] Sheblé, G., "Priced Based Operation in an Auction Market Structure," Presented at the 1996 IEEE/PES Winter Meeting, Baltimore, MD, 1996.

[6] Caramanis, M.C., R.E. Bohn, and F.C. Schweppe, "Optimal Spot Pricing: Practice and Theory," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-101, No. 9, 1982, pp. 3234-3244.

[7] Howard, R.A. "Decision Analysis: Practice and Promise," *Management Science*, 1988:679-675.

[8] MIDAS, "Multi Objective Integrated Decision Analysis System," *EPRI Journal*, Vol. 12, 1987:57-58.

[9] Kumar, J. and G.B. Sheble, "A Framework for Transaction Selection Using Decision Analysis Based upon Risk and Cost of Insurance," Proc. of the 29th North American Power Symposium, Kansas State University, KS, 1994:548-557.

[10] Kumar, J., and G. Sheble, "A Decision Analysis Approach to Transaction Selection Problem in a Competitive Market," *Electric Power Systems Research Journal,* 1996.

[11] Ashlock, D., "GP-Automata for Dividing the Dollar," Mathematics Department, Iowa State University, Ames, IA, 1995.

[12] Ashlock, D., and C. Richter, "The Effect of Splitting Populations on Bidding Strategies," GP-97, Proceedings of the Second Annual Conference, pp. 27-34, San Francisco, CA, Morgan Kaufmann Publishers.

[13] Richter, C., A. Jayantilal, and J. Kumar, "Improving Market Participant Strategies with FTR Options," forthcoming in Proceedings of 2001 PICA Conference, Sydney, Australia.

AN ABS INVESTIGATION OF GENERATOR MARKET POWER IN THE ENGLAND AND WALES ELECTRICITY MARKET USING AN EXCEL/VBA PLATFORM

J. BOWER and D.W. BUNN, London Business School[*]

EXTENDED ABSTRACT

In November 2000, the UK electricity market regulator (OFGEM) proposes to replace the existing trading arrangements in the England & Wales Electricity Pool ("the Pool") with a bilateral market. This action has been taken in response to persistently high and volatile prices that are symptomatic of the extent to which generating firms have been able to exercise market power. In addition to changing the trading arrangements OFGEM has also forced the industry to restructure and as a result the largest incumbent generators divested plant to new entrant firms from the United States during 1999.

Under these New Electricity Trading Arrangements (NETA) consumers will contract directly with generators for supplies of physical power rather than a centralised market place. Self despatch of generating plant is envisaged as the main mechanism of delivery, and an optional Balancing (spot) Market will be used to maintain system security in which generators and consumers make firm bids and offers for increments and decrements of power, in real-time, for each half hour of the day. In contrast to the current Pool, in which all generators get paid the bid price of the marginal generating plant in each period (SMP) plus a capacity payment for making plant available to the system, generators would only be paid their own bid price. Underlying the proposed reform is the strong belief that the wholesale electricity market should operate more like other competitive commodity markets and that paying generators SMP only serves to increase the potential for gaming and exploitation of market power, in particular, that of the mid-merit generators. The rationale is simple; pay-as-bid pricing should reduce market power and, hence, wholesale electricity market prices will fall as a result.

Since the beginning of 2000 it is clear that Pool prices have been significantly lower than in previous years. OFGEM has publicly stated that it believes this fall in prices is due to the market's expectation of the imminent introduction of NETA in November 2000. However, this conclusion ignores the potential impact that industry restructuring during 1999 may also have had on market prices. To examine whether the recent price falls are due to proposed changes in the trading arrangements, industry restructuring, or both, we have built an agent-based simulation (ABS) model of the Pool and NETA. In this modelling environment, each generator is represented as an autonomous adaptive agent that submits a separate daily bid price, from each of its plants, at which it is prepared to supply electricity to the market. Equipped with a rudimentary learning ability, each agent is capable of developing its own bidding strategies in response to changes in trading arrangements, and the bidding strategies of other agents. Despite the simplicity of the agents' behavioural rules, complex bidding strategies emerge spontaneously. These are both consistent with behaviour seen in the real electricity markets and with economic theory. Additionally, as the model relies on the micro-simulation of pricing

---

[*] *Corresponding author address*: John Bower, London Business School, Sussex Place, Regent's Park, London NW1 4SA, United Kingdom; e-mail: jbower@london.edu.

decisions for individual power plants, it has been possible to investigate the impact of plant divestment as an alternative means of reducing generator market power.

The ABS model is a discrete event simulation platform that replicates the daily bidding activity, market clearing, and settlement processes in the Pool and NETA. Generating firms (agents) are represented by data arrays containing plant capacity, operating costs, plant availability, and bidding algorithms held on Excel spreadsheets that are manipulated with Visual Basic commands. As Excel is primarily designed to handle large arrays of data and formulas it is possible to simulate four years worth of trading activity, representing over 2 million separate bids, in approximately 30 minutes. Excel appears to offer a higher level of performance, at least for this specific purpose, than could be achieved using alternative agent-based tools such as Swarm or commercial simulation packages such as Mathematica or Powersim. Scaling or amending the ABS model, for example to increase the number of agents or represent an entirely different national different market, only requires the use of 'cut and paste' commands rather than reprogramming. Built in graphical interfaces, analytical and charting tools, data export facilities, and easy portability offer additional advantages.

The simulated results show that, far from increasing competition, the NETA proposal would actually magnify market power by allowing generating firms to segment the market on a half hour-by-half hour basis. In contrast, when recent forced divestments, and other changes to industry structure are included in the model there is a significant reduction in the ability of firms to coordinate their actions, resulting in a significant fall in simulated Pool prices. Our conclusion is that, adopting the NETA reforms proposed by OFGEM could therefore have a detrimental effect and that the industry structure reforms already carried out should be sufficient to curtail the market power of generators.

REFERENCES

Bower, J., Bunn, D. W. (2001) "Experimental analysis of the efficiency of uniform-price versus discriminatory auctions in the England and Wales Electricity market," *Journal of Economic Dynamics and Control*, Vol. 24, No. 3/4, pp. 561-592.

Bower J., Wattendrup, C., Bunn, D. W. (2000) "Implications of German Electricity industry consolidation on competition: An agent-based analysis," *London Business School Working Paper* (forthcoming in *Energy Policy*, 2001)

Bower, J., Bunn, D. W. (2000) "Model-Based Comparisons of Pool and Bilateral Markets for Electricity," *The Energy Journal,* Vol. 21, No. 3, pp. 1-29.

ACKNOWLEDGMENT

This research was partly funded by ESRC Postgraduate Training Award No. R00429724400 and a London Business School Financial Award.

London Business School

An ABS investigation of generator market power in the England & Wales Electricity market using an Excel/VBA platform

Agent2000, Chicago

5-7 October, 2000

John Bower & Derek Bunn

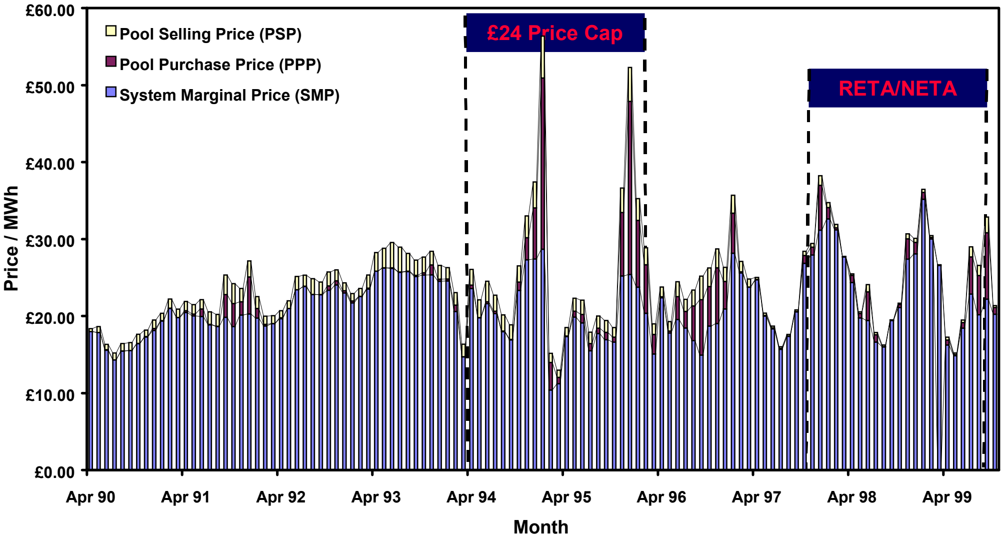
Overview

- Introduction and summary
- An agent-based model of strategic rivalry
- Results
- Conclusion

Introduction and summary

UK electricity prices went up after deregulation as generators exercised market power….

Time Weighted Average Monthly Pool Prices



Introduction and summary

…. and the 1998 RETA (Review of Electricity Trading Arrangements) produced NETA....

Proposed Solutions to Market Power

Market reform: Uniform Price (Pay SMP) Pool day-ahead auction replaced with Discriminatory (Pay Bid) auction in Bilateral Model

Plant Divestment: National Power and PowerGen each to divest 4,000 MW (~25%) of their plant

Gas moratorium: Ban on new gas fired generating plant until Pool reform and divestment has taken effect

Introduction and summary

…. (New Electricity Trading Arrangements) to replace the Pool with the Bilateral Model….

Reform Objectives

1. **Pay-as-Bid (WYBIWYG)**
2. **Firm Bidding**
3. **Simple Bidding**
4. **Transparent Bidding**
5. **Demand Side Bidding**
6. **Security of Supply**

Reform Proposals

1. **Forward / futures / options**
2. **Short-term (24 - 4 hr)**
3. **Balancing (4hr - despatch)**
4. **Settlement Process**

Introduction and summary

…. at a cost of over £500 million but with no hard evidence that it would actually work!

Research Questions

What is the impact of the following factors on wholesale price:

Market mechanism (i.e. wholesale trading arrangements)?

Industry structure (i.e. size and number of generating firms)?

Plant technology (i.e. type and distribution of plant)?

An agent-based model of strategic rivalry

Urgent need to test but few insights from empirical observation, economics, game theory

Operations Research Response

An agent-based simulation approach offered a potential solution:

Discrete event simulation: agents replicate repeated nature of daily trading

Artificial Intelligence: agents build own strategies with reinforcement learning

Behavioural modelling: agent strategies emerge not imposed by modeller

Level of aggregation: micro ‘bottom up’ not aggregate ‘top down’

Experimental method: an economic laboratory with perfect controls

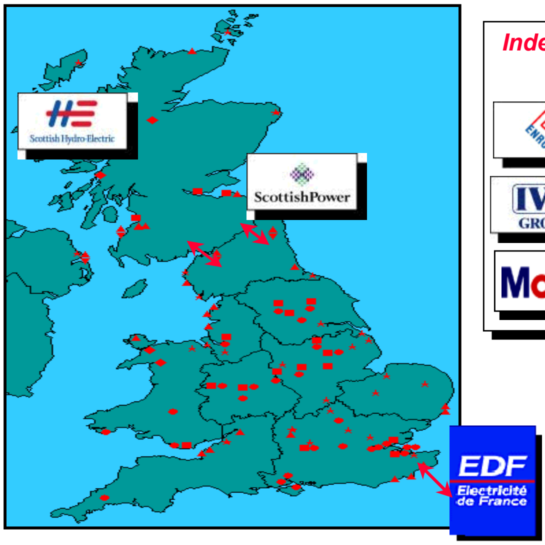
An agent-based model of strategic rivalry

We use a “bottom up” agent-based method to analysing market power in generation ….

Firms and plants operating in England & Wales Pool

Privatised Firms England & Wales

National Power
POWERGEN
EASTERN A TXU Company
BNFL Magnox Generation
British Energy
EDISON MISSION ENERGY



Independent Power Producers

ENRON
INTERGEN
IVO GROUP
Entergy
Mobil
AES

An agent-based model of strategic rivalry

…. which consists of four components and allows us to …..

Model components

Economic environment:	A series of interchangeable auction market types through which electricity is traded
Agents (Supply):	Each of the generating firms operating in the Pool is individually represented at the level of its plants
Agents (Demand) :	Consumers are represented as an aggregate demand curve estimated from empirical data
P&L Archive:	Calculator and data base where results of daily trading evaluated and stored

---

An agent-based model of strategic rivalry

…. capture crucial features of the micro-structure of the supply side of the market

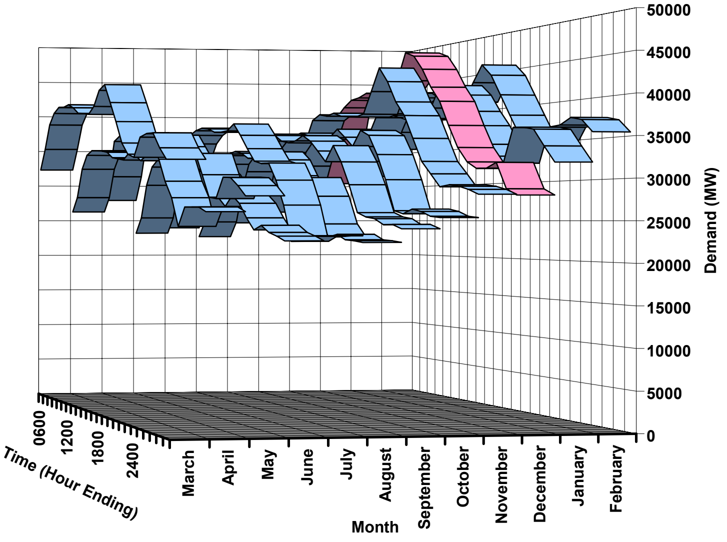
Model Inputs

- (National Grid Company SYS reports)
- (genset outage rates from private industry sources)
- (heat rates from industry / environmental reports)
- (coal, gas, oil prices from Reuters data feeds)
- (OFGEM / industry reports)

An agent-based model of strategic rivalry

…. but demand side agents are aggregated as assumption is they have no market power

Seasonal Aggregate Demand (Working Days)



An agent-based model of strategic rivalry

Generating agents submit a bid for each hour of the day, for each plant, for four years ….

ABS Model Daily Bid Cycle

Step 2. Auction

Step 1. Plant Bids

Step 3. Results

Step 4. Archive

An agent-based model of strategic rivalry

…. so built custom VBA/Excel tool to run this large scale simulation in λ 30 mins on PC

Simulation Bidding Statistics

90 PLANTS

x

4 YEARS

x

248 DAYS

x

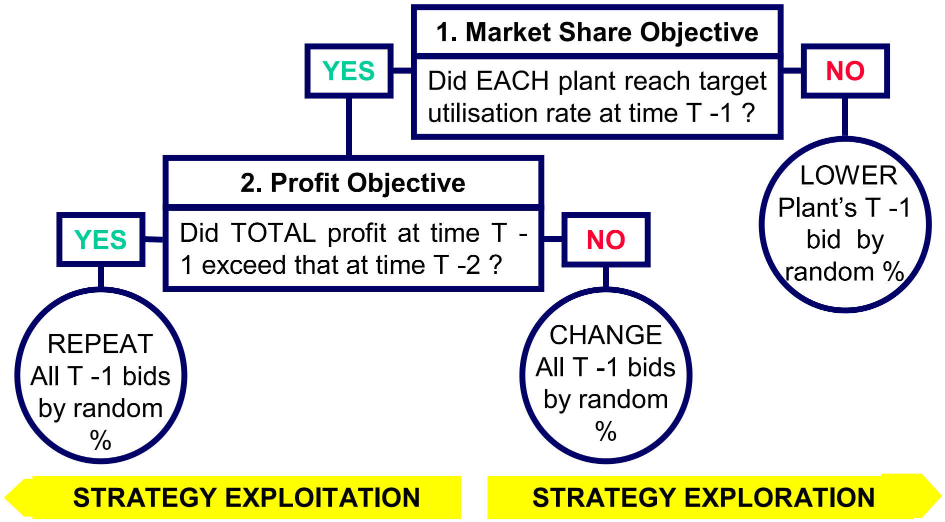
24 HOURS

2,140,000 BIDS

An agent-based model of strategic rivalry

Using naive reinforcement learning each agent develops it bidding strategies through time

Agent behavioural rules



1. Market Share Objective

Did EACH plant reach target utilisation rate at time T -1 ?

YES

NO

LOWER Plant's T -1 bid by random %

2. Profit Objective

Did TOTAL profit at time T -1 exceed that at time T -2 ?

YES

NO

REPEAT All T -1 bids by random %

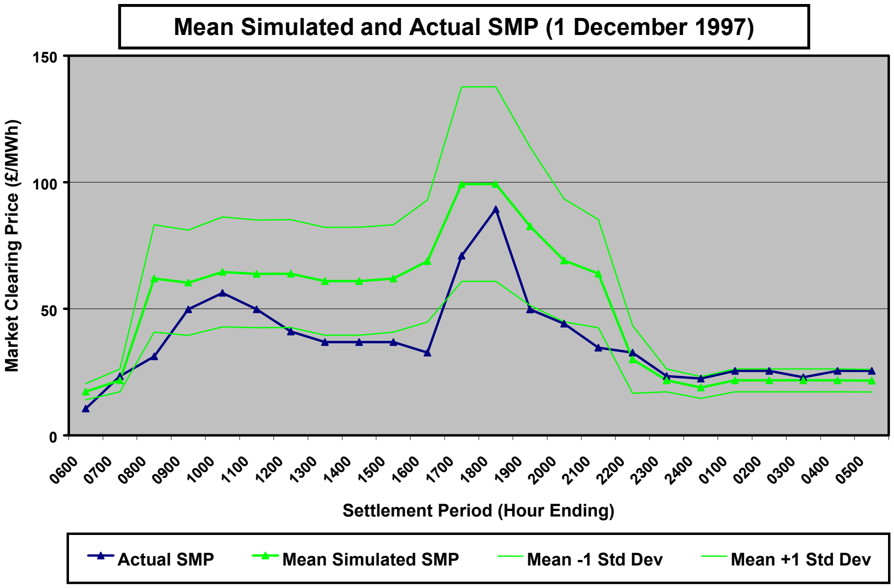
CHANGE All T -1 bids by random %

STRATEGY EXPLOITATION

STRATEGY EXPLORATION

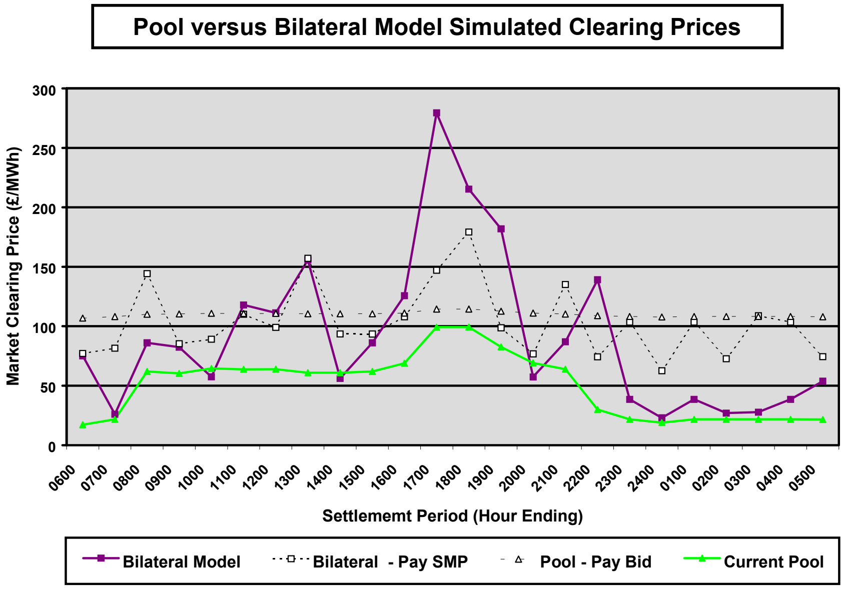
Results

Calibrated model by adjusting Target Utilisation rate and replicate one day of trading….



Results

… and then compared prices from simulated trading in the Pool and Bilateral Model



Results

Agent-based simulation has produced results contrary to the regulators expectations

Summary Results

Bilateral Model would make market less competitive than in current Pool

1. Large generators gain an information advantage over small generators
2. Large generators can raise operational risk for small new entrants
3. Large generators segment market more easily and raise price in peak hours

Ofgem believes that the market abuse licence condition is necessary …. and that there is scope for future abuse of wholesale electricity market under both the present electricity trading arrangements and under NETA.

Source: Introduction of the market abuse condition into the licences of certain generators. Ofgem's second submission to the Competition Commission, OFGEM June 2000, http://www. ofgem.gov.uk/public/pub.htm

Conclusion

This is the only model-based analysis of NETA and it is widely used by industry/regulators

Model Application

- Results cited in industry's representations to UK regulator (OFGEM)
- OFGEM forced to respond with own real-time trading experiment
- Results put before the UK DTI Select Committee on Energy
- Requested to present evidence to UK Competition Commission
- Model used to test merger proposals in German electricity market
- Future application to European cross-border electricity market

Conclusion

By combining a number of OR methodologies we addressed a critical economic issue

Impact and Contribution

- **Government Policy:** Regulator now agrees NETA will not control market power and Govt. is putting in place a tighter regulatory framework
- **Auction Theory:** Showed multi-unit, multi-period auction mechanisms are not *Revenue Equivalent* where bidders exercise market power
- **ABS Methodology:** One of the first applications of a large scale ABS model to solve a real problem with a substantial economic impact
- **OR in Practice:** Produced insights into strategic behaviour that have eluded conventional economic and game theoretic analysis

Speaker

John Bower is a Doctoral student at London Business School and is a member of the Energy Markets Group within the Decision Technology Centre. His research interest is in the study of market power and evolution of cross-border commodity trading in electricity. In particular, he is using agent-based simulation approaches to study the emergence of complex strategic behaviour between firms in these new markets.

John's previous career was in the commodity industry and his experience ranges from energy trading, at Marc Rich & Co, to risk management consultancy, with Coopers & Lybrand, advising commodity traders, producers and processors in base metal, precious metal, 'softs' and energy markets. Before joining the PhD programme he was Global Controller Metals/Commodities at Deutsche Morgan Grenfell.

John also has an MBA from London Business School and an MA in Biochemistry from Oxford University.

London Business School
Sussex Place, Regent's Park
London NW1 4SA
United Kingdom

Telephone: +44 (0)20 7834 6736
Facsimile: +44 (0)20 7834 6736
e-mail: jbower@london.edu

AGENT-BASED MODELING OF COMPLEX INFRASTRUCTURES

M. NORTH, Argonne National Laboratory*

ABSTRACT

Complex Adaptive Systems (CAS) can be applied to investigate complex infrastructures and infrastructure interdependencies. The CAS model agents within the Spot Market Agent Research Tool (SMART) and Flexible Agent Simulation Toolkit (FAST) allow investigation of the electric power infrastructure, the natural gas infrastructure and their interdependencies.

KEYWORDS

Complex adaptive systems (CAS), Agent-based modeling (ABM), Electric power system modeling, Natural gas system modeling, Infrastructure interdependency, Swarm, RePast, CAVE, Virtual reality (VR)

INTRODUCTION

Many insights can be gained by viewing energy analysis from a Complex Adaptive Systems (CAS) agent-based modeling perspective. Argonne has taken such a perspective to produce integrated models of the electric power and natural gas markets. The agents within the present Spot Market Agent Research Tool (SMART) and the future Flexible Agent Simulation Toolkit (FAST) allow investigation of the electric power infrastructure, the natural gas infrastructure and their interdependency.

THE PRESENT AND FUTURE

Several tools presently exist:

- SMART Version 2.0 (SMART II) is a Swarm model with an integrated set of agents and interconnections representing the electric power marketing and transmission infrastructure.
- SMART II VR is a virtual reality (VR) interface for SMART II.

---

* *Corresponding author address*: Michael J. North, Decision and Information Sciences Division, Complex Adaptive Systems Section, Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439; e-mail: north@anl.gov.

- SMART II+ is an extension to SMART II that includes an integrated set of agents and interconnections representing the electric power infrastructure, the natural gas infrastructure and connections between them in the form of natural gas-fired electric generators.

FAST is currently under construction. FAST is a complete redesign of SMART II+ that includes improvements in the modeling environment, model detail and representational fidelity.

SMART II

SMART II is a Swarm-based [1] model that uses a set of agents and interconnections to represent electric power systems. SMART II is the Swarm Development Group 2000 Conference (SwarmFest 2000) Best Presentation winner. SMART II itself builds on several other models [2-3]. The SMART II interface is shown in Figure 1. SMART II includes three different kinds of components as follows:

- Generation agents produce electric power.

- Consumer agents use electric power.

- Interconnections represent the transmission grid.

SMART II considers important economic issues such as production costs, investment capital, demand growth for successful consumers, new generation capacity for profitable producers, and bankruptcy for noncompetitive organizations.

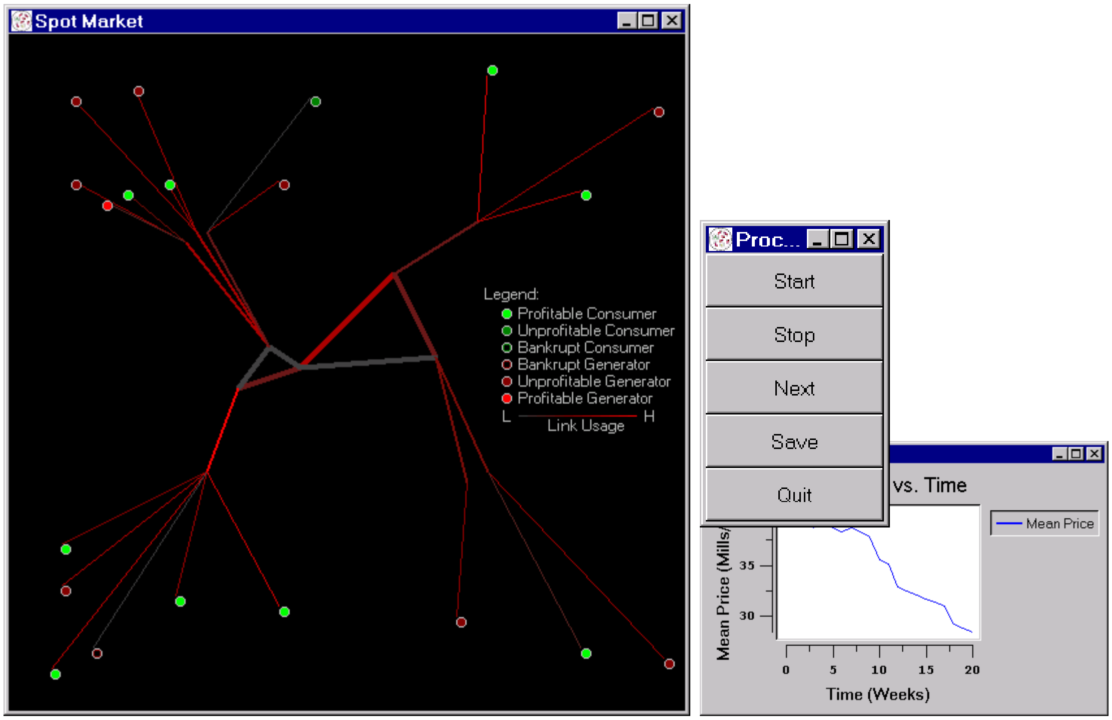


FIGURE 1 The SMART II Interface

SMART II has undergone initial qualitative validation by matching its outputs to the following basic analytic predictions:

- Markets with a single superior producer among a large number of higher cost competitors have been tested.

- Markets with many identical participants have been tested.

Much more work is clearly required to quantitatively validate and calibrate SMART II therefore only limited qualitative insights are currently being derived.

As originally presented at SwarmFest 2000, qualitative insights from SMART II indicate that certain transmission line configurations may encourage price spikes. Soon after SwarmFest 2000, this insight was borne out.

As was specifically noted at SwarmFest 2000, the California electrical grid has a configuration of a type that may cause price spikes. Substantial price spikes of the kind predicted by SMARTII where recently observed in this market.

Further qualitative insights suggest that greater electrical market price stability may be gained by consciously avoiding specific configurations that encourage instabilities. In other words, qualitative insights from SMART II can help us make things better by not making things worse.

SMART II VR

SMART II VR is a prototype agent visualization tool. SMART II VR is intended to explore the use of advanced interactive three-dimensional visualization in agent-based modeling.

SMART II is a CAVE Automatic Virtual Environment (CAVE)-based virtual reality interface for SMART II. The CAVE is a virtual reality library co-developed by the University of Illinois at Chicago and Argonne. From the CAVE User's Guide [4]:

> The CAVE is a projection-based VR system that surrounds the viewer with four screens. The screens are arranged in a cube made up of three rear-projection screens for walls and a down-projection screen for the floor; that is, a projector overhead points to a mirror, which reflects the images onto the floor. A viewer wears stereo shutter glasses and a six-degrees-of-freedom head-tracking device. As the viewer moves inside the CAVE, the correct stereoscopic perspective projections are calculated for each wall. A second sensor and buttons in a wand held by the viewer provide interaction with the virtual environment.

SMART II VR includes an interactive multifunction wand and two rendering modes.

Detail rendering mode focuses on rendering quality. Directional lighting is included. Agents are rendered as lighted spheres. A texture-mapped floor with shadows and first order reflections is included. This mode allows SMART II VR to take advantage of computers with high graphics performance. An example is shown in Figure 2.

Speed rendering mode focuses on rendering time. Agents are rendered as flat shaded cubes. This mode allows SMART II VR to be used on low performance personal computers. An example is shown in Figure 3.

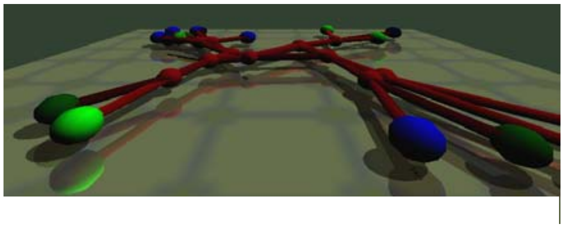


FIGURE 2 SMART II VR Detail Mode

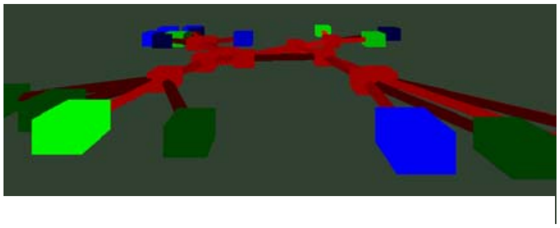


FIGURE 3 SMART II VR Speed Mode

In SMART II VR, generation agents are shown as green spheres or cubes. Spheres are shown in detailed rendering mode and cubes are shown in speed rendering mode. The size of each object represents its total normalized investment capital level. Size can be interactively changed with the CAVE wand. Each object's color intensity represents its hourly profit level.

In SMART II VR, consumer agents are shown as blue spheres or cubes. Spheres are shown in detailed rendering mode and cubes are shown in speed rendering mode. The size of each object represents its total normalized investment capital level. Sizes can be interactively changed with the CAVE wand. Each object's color intensity represents its hourly profit level.

In SMART II VR, interconnections are displayed as red tubes. The size of each tube represents its normalized transmission capacity level. Sizes can be interactively changed with the CAVE wand. Each tube's color intensity represents its hourly utilization level.

SMART II+

SMART II+ is a Swarm-based [1] extension to SMART II. SMART II is the Swarm Development Group 2000 Conference (SwarmFest 2000) Best Presentation winner. SMART II itself builds on several other models [2-3].

SMART II+ includes an integrated set of agents and interconnections representing each of the following:

- The electric power marketing and transmission infrastructure.
- The natural gas marketing and distribution infrastructure.
- The interconnections between the two infrastructures in the form of natural gas fired electric generators.

Both of the infrastructures modeled in SMART II+ include many features:

- Two different kinds of agents, producers and consumers, represent the market participants.

- Interconnections represent transmission or distribution systems with capacities on each line or pipe and complex routing.

- Important economic issues are considered such as investment capital, demand growth for successful consumers, new generation capacity for profitable producers, and bankruptcy for noncompetitive organizations.

- Components can be disabled in real time to simulate failures.

The electric power infrastructure includes the added feature of natural gas fired electric generators. These generators buy fuel from the natural gas market. The resulting electricity is then sold in the electric power market.

SMART II+ PRODUCER AGENTS

SMART II+ producers determine their production level based on the potential profit to be made. Each producer has investment capital that is increased by profits and reduced by losses. If a producer reaches a predetermined level of investment capital it can purchase additional production capacity in the form of new electric generators or new natural gas sources. New producers are similar to their owner and can connect to the distribution network in either the same location or a new one. Producers that run out of investment capital go bankrupt and no longer participate in the market. Producers choose whether or not to sell energy based on either their cost curves or natural gas prices. Standard producers derive their costs and capacities from cost curves with maximum generation limits as shown in Figure 4. Both costs and capacities are exogenous.

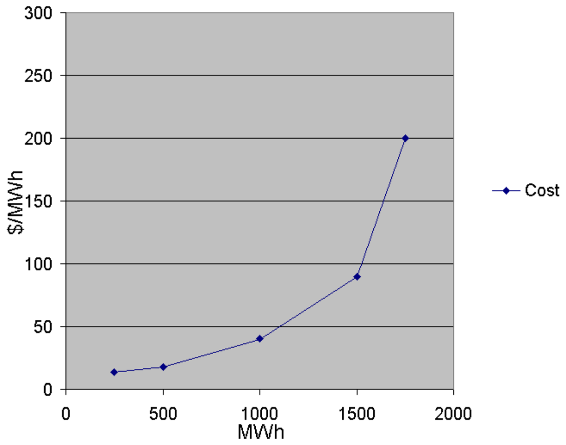


FIGURE 4 An example producer cost curve

Natural gas fired electric generators derive their costs and capacities from the endogenous natural gas market. These generators are consumers in the natural gas marketplace. Their costs are based on the price they pay for natural gas. Their capacities are based on both the amount of natural gas they can purchase and their design limits.

Producer simulation display appearance depends on current profit levels (Figure 5).

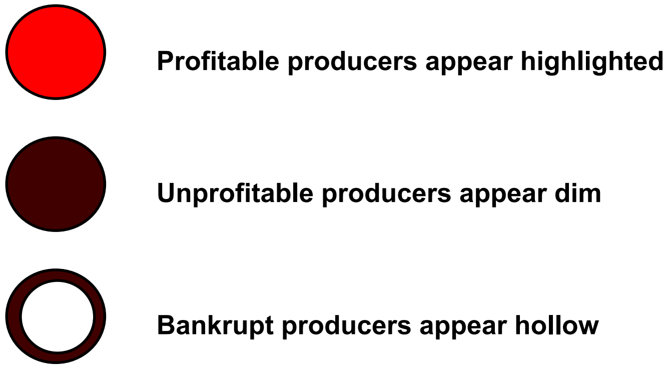


FIGURE 5 Producer Appearance

SMART II+ CONSUMER AGENTS

SMART II+ consumer agents buy energy for their own use. Businesses buy fixed amounts of energy to remain in business. Populations buy fixed amounts of energy to live their lives. Natural gas fired electric generators buy natural gas to produce salable electric power.

Each consumer has investment capital that is increased by profits and reduced by losses. If a consumer reaches a predetermined level of investment capital it can grow in the form of new consumers. Consumers that run out of investment capital go bankrupt and no longer participate in the market.

Consumer simulation display appearance depends on current profit levels (Figure 6).

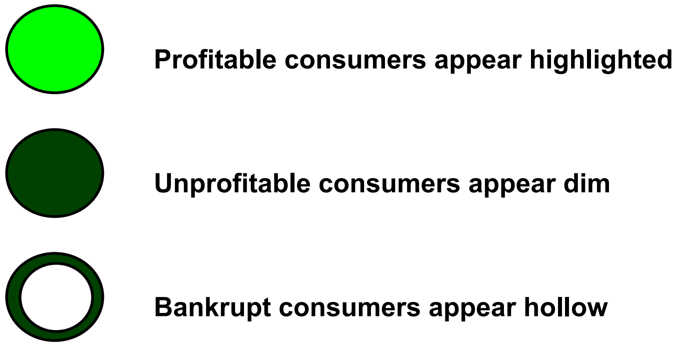


FIGURE 6 Consumer Appearance

Investment capital represents several things. For industrial users it is their total financial capital. For individuals it is the employment and personal opportunities that keep them in an area or encourage them to leave.

SMART II+ INTERCONNECTIONS

Interconnections represent transmission lines or distribution pipes each with an individual capacity limit. Individual capacity limits vary by interconnection type. Central transmission lines or main distribution pipes have high capacity limits and are drawn with thick marks. Outlying transmission lines or secondary distribution pipes have moderate capacity limits and are drawn with medium marks. Feeder lines or pipes have low capacity limits and are drawn with thin marks. Interconnection color represents contents and usage as shown in Figure 7.

FIGURE 7 Interconnection Appearance

SMART II+ MARKET INDICATORS

The key SMART II+ market indicators are market prices, unserved energy and natural gas fired electrical generator market share. All key SMART II+ indicators are represented by graphs updated in real time.

Market price is the per unit purchase price of the given energy resource. Electric power prices are given in tenths of a cent per kilowatt-hour (Mills/kWh). Natural gas prices are given in dollars per thousand cubic feet ($/1,000 cubic feet). The SMART II+ price graphs are shown in Figure 8.

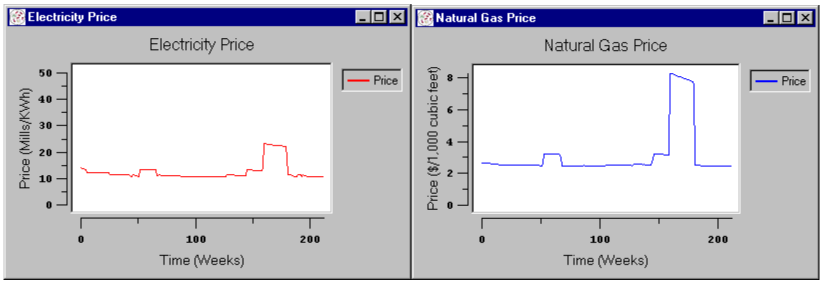


FIGURE 8 Price Graphs

Unserved energy (UE) is the energy demand that was not met by the market. UE represents a form of market failure. UE is given as a percentage of total energy demand. The SMART II+ UE graph is shown in Figure 9.

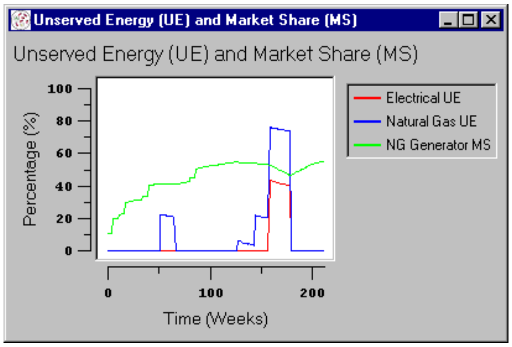


FIGURE 9 UE and NG Generator MS Graph

Natural gas fired electric generator market share (NG Generator MS) is a measure of the electric generation capacity that is supplied by natural gas units. NG Generator MS is key to infrastructure interdependency. NG Generator MS is given as a percentage of total capacity. The SMART II+ NG Generator MS graph is also shown in Figure 9.

SMART II+ NETWORK DISPLAY

The geographical SMART II+ display is based on an equivalenced network. An example notional SMART II+ network is shown in Figure 10.

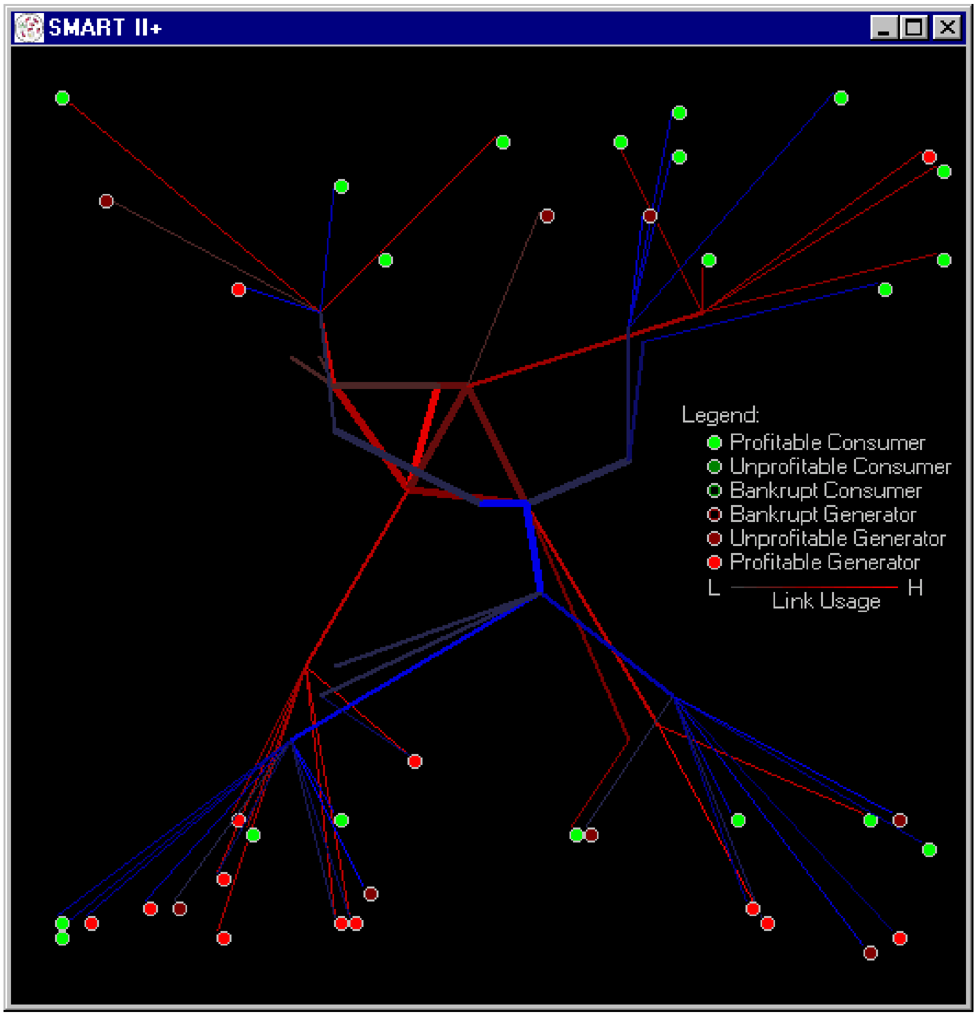


FIGURE 10 Example Notional SMART II+ Network

SMART II+ VALIDATION AND CALIBRATION

As with SMART II, SMART II+ has undergone initial qualitative validation by matching its outputs to basic analytic predictions:

- Markets with a single superior producer among a large number of higher cost competitors have been tested.
- Markets with many identical participants have been tested.

SMART II+ has undergone initial qualitative calibration by comparing the model's natural gas-fired electric generator market share trends to those found in real systems.

Much more work is clearly required to quantitatively validate and quantitatively calibrate SMART II+. Therefore, only limited qualitative insights are currently being derived.

SMART II+ INSIGHTS

As originally presented to our research sponsors in May 2000, preliminary insights from SMART II+ indicate that:

- Rising natural gas-fired electrical generator market share radically increases market interdependence.

- Increasing market interdependence can pit the electric power and natural gas markets against one another during simultaneous failures since both markets are fighting for the same underlying resource, natural gas.

This interdependency insight was borne out in the aftermath of the recent El Paso natural gas pipeline explosion.

What is the state of the world today? Nationwide natural gas-fired electrical generator market share is roughly 15% to 20%. Nationwide natural gas-fired electrical generator market share is expected to radically increase over the next five years. J.P. Morgan analysts predict that there is expected to be a 31% increase in generation capacity [5]. These analysts predict that roughly 95% of new electrical generation capacity will come from natural gas-fired units [5]. An example is the midwestern region dominated by Commonwealth Edison (ComEd).

ComEd presently gets less than 10% of its current 20,000 MW generation capacity from natural gas-fired generators. Permits are being issued for the construction of 8,000 MW of new capacity, over 95% of which will be natural gas-fired.

The interdependency between the electric power and natural gas markets implies that when natural gas-fired electrical generator market share becomes high enough a single energy resource, "virtual natural gas," is being traded in both markets. Viewing energy systems from the perspective of virtual natural gas suggests that future electrical system capacity expansion planning should explicitly feature the natural gas distribution infrastructure as a key component. Power Systems Engineers should note that electrical models might be substantially incomplete without explicitly including the natural gas infrastructure. Highly distributed electrical generation plans including local load servicing schemes may especially benefit from this view since they rely heavily on the existence of other energy sources such as natural gas.

FAST

FAST is an integrated infrastructure model based on SMART II+. FAST includes many of the features of SMART II+ along with improvements in modeling infrastructure, detail and fidelity. FAST is currently under construction. FAST has three components:

- FAST:Run is the runtime infrastructure that will be merged with RePast [6].

- FAST:E is the electric power system model.

- FAST:G is the natural gas system model.

FAST:Run is designed to be a lightweight large-scale system with the following major features:

- FAST:Run is written entirely in Java.
- FAST:Run is fully distributed.
- FAST:Run has a multithreaded scheduler that focuses on maximizing parallel execution.

The underlying design paradigm of FAST is that of a time continuum ranging from decades to seconds:

- On the scale of decades the focus is long term human decisions constrained by economics.
- On the scale of years the focus is short-term human economic decisions constrained by economics.
- On the scale of months, days and hours the focus is short-term human economic decisions constrained by economics and physical laws.
- On the scale of minutes or less the focus is on physical laws that govern energy distribution systems.

Modeling over the full range of time scales is necessary to understand the complex infrastructure interdependency between the electric power and natural gas markets.

FAST includes a large number of different agents to model the full range of time scales. The focus of agent rules in FAST varies to match the time continuum. Over longer time scales human economic decisions are emphasized. Over shorter time scales physical laws dominate.

Many FAST agents are relatively “thick” compared to typical agents. FAST agents are highly specialized to perform diverse tasks ranging from acting as Independent System Operators to being transmission lines. To support specialization, FAST agents include large numbers of highly specific rules.

The FAST system and its component agents will be subjected to rigorous quantitative validation and calibration.

CONCLUSION

Developing the initial capability to create CAS models requires substantial organizational investment. Once this initial investment has been made tools can be created that allow many insights.

ACKNOWLEDGEMENT

This work is sponsored by the U.S. Department of Energy under contract number W-31-109-ENG-38.

REFERENCES

1. Burkhart, R., M. Askenazi, and N. Minar. *Swarm Release Documentation*. Available as http://www.santafe.edu/projects/swarm//swarmdocs/set/set.html.

2. VanKuiken, J.C., T.D. Veselka, K.A. Guziel, D.W. Blodgett, S. Hamilton, J.A. Kavicky, V.S. Koritarov, M.J. North, A.A. Novickas, K.R. Paprockas, E.C. Portante, and D.L. Willing, *APEX User's Guide (Argonne Production, Expansion, and Exchange Model for Electrical Systems) Version 3.0*, ANL/DIS/TM-21, Argonne National Laboratory, Argonne, Ill., 485 pages (November 1994).

3. Veselka, T.D., E.C. Portante, V.S. Koritarov, S. Hamilton, J.C. VanKuiken, K.R. Paprockas, M.J. North, J.A. Kavicky, K.A. Guziel, L.A. Poch, S. Folga, M.M. Tompkins, and A.A. Novickas, *Impacts of Western Area Power Administration's Power Marketing Alternatives on Electric Utility Systems*, ANL/DIS/TM-10, Argonne National Laboratory, Argonne, Ill., 216 pages (January 1994).

4. Pape, D., C. Cruz-Neira, M. Czernuszenko *CAVE User's Guide*. Available as http://www.evl.uic.edu/pape/CAVE/prog/CAVEGuide.html (1997).

5. Anderson, J.A., "What's Electrifying Natural Gas," *Business Week,* (September 19, 2000).

6. Collier, N. and D. Sallach. *RePast*. Available as http://repast.sourceforge.net/.

DISCUSSION: ENERGY AND INFRASTRUCTURES

R. CIRILLO, Argonne National Laboratory, Moderator

[Presentation by Bower]

Tony Andrews: If they bid zero, are you saying that they are not bidding?

John Bower: No, when they bid zero they're not charging for their product, they're giving it away for nothing. The trading rule is they get paid the marginal plant bid. They want to make sure they run, though, which is why they're bidding zero.

Gale Boyd: Gale Boyd from Argonne National Laboratory. Have you done any simulations looking at changes in market structure? If so, do bilateral market prices still come out higher? In other words, does the more competitive supply side affect the bilateral solution?

Bower: Yes. At the limit, if you have a very competitive market structure, say with 50 firms, it really doesn't matter what kind of market you have; you're going to end up with marginal pricing. If you have a monopoly, it also doesn't matter what kind of trading arrangements you have, you're going to get the monopoly price. In between, when you have a duopoly or an oligopoly of firms, the types of trading arrangements *do* make a difference. As for changing the market structure, we have simulated this in some detail using our model, and the way we have changed the structure has been sufficient, we think, to curtail the market price. And in fact we've seen the market price begin to fall. So we hope it's worked.

Boyd: I have another comment. I found a great deal of similarity between your presentation and yesterday's paper by Jing Yang ["Price Efficiency and Risk Sharing in Two Inter-Dealer Markets"]. Her paper talked about dealer-to-dealer as a bilateral trade, the broker as essentially a pool, and some of the circumstances where there's a mixture of these elements. But you dealt with these elements as either/or. Have you considered a mixed market, where people can make both bilateral trades and trades in the pool?

Bower: We have thought about modeling some of the questions that Jing Yang talked about, because at one point there was a proposal to have both an on-the-telephone kind of bilateral market and an on-screen market, and we wondered whether that would make any difference. In fact, the regulators canned the idea because they said they didn't want to prescribe how people should trade. So they're not going to have a screen-based market. But what we found was really crucial here is the information available to different agents. Big agents, big firms, gain an informational advantage in the bilateral market over small firms, and that allows them to drive the price up. So information — what's on the screen, what isn't on the screen — it's absolutely crucial.

Prakash Thimmapuram: I'm Prakash Thimmapuram from Argonne. You said they can withdraw capacity by the generator. At what time can they withdraw the capacity?

Bower: They can make their bid in the pool, and they can withdraw their capacity one second before it's dispatched.

Thimmapuram: Is there any penalty for withdrawing the capacity?

Bower: No.

Thimmapuram: Another question: You show that the payer's bid is higher than the marginal price. Can you explain why this is so?

Bower: Under marginal price, the large, low-cost generators always bid zero — so they're always undercutting the coal plants, which can't bid zero. That's in the pool. But in a bilateral market, everyone has to bid around the marginal price. Occasionally, a low-cost operator will overbid a coal plant, and that means the competition against the coal plants is reduced. Another reason is that bidding is now by the hour rather than the whole day, and that means the large generators can segment the market. They can bid much higher prices for peak hours than they bid for the base-load periods. In fact, base-load prices are virtually the same in the bilateral market and the pool. As you know, segmenting the market is a great way of discriminating against certain customers. There's not much elasticity in this market.

Randal Picker: What is the source of the information advantage for the big traders in the bilateral market? Is it that they observe more trades?

Bower: Yes. If you've got 10 plants you're putting in a lot more bids than the one-plant operator. Also, if you're a small, marginal-cost operator, you're really quite risk averse because you've got large amounts of debt. You want to run your plant all the time. Often we find the small generators are having to shave their bids dramatically. So really the small generators are nowhere near the marginal price some of the time, and they haven't many plants, so they just don't have the same capacity to learn.

Richard Cirillo: The issue was raised yesterday about whether agent-based simulation techniques, at least in some domains, were ready to be used to help shape and formulate policy. What's your experience with the agent-based simulation technique in this domain — is it ready for prime time?

Bower: In fact, this model is being used by the government in the U.K. already, by the Competition Commission, and it's being widely cited by the industry. The regulators dislike the result so much that they set up their own experiment with real people. They had 50 students trade the new market. One of the students found a way of getting the price up, but they only allowed the students to trade for 12 days, so, in my opinion, the students had no chance to learn. Their results showed no difference between the pool trading and the bilateral trading. So they just carried on and implemented the marketplace.

[Presentation by North]

Picker: What happens to the generating capacity if the agents are destroyed?

Michael North: The generating capacity is reused. It's available for some other potential entrepreneur, or even an existing large company, to take it over. Given that this is a very competitive market, it's usually acquired very quickly.

Picker: As you know, FERC [Federal Energy Regulatory Commission] has proposed moving from ISOs [independent system operators] to regional transmission grid operators. How easy is it for you to capture that kind of institutional change in your agents?

North: It depends on the amount of change you want to make. It's actually relatively easy to replace agents in the model. Building the model of the agent that you're working with, though, can be a very difficult task. It's not really an infrastructure issue; it's a matter of thoroughly identifying and characterizing a different agent. We hope it's going to be easier with the new infrastructure because one of the things we're focusing on is changing market roles. It's just one of our interests.

Blake LeBaron: Blake LeBaron, Brandeis University. You've got a couple of complicated infrastructure investment problems underneath, both in transmission capacity and generating capacity. How do your agents do forecasting? This is a difficult problem, especially the transmission one. I'm curious about how the grid forms based on forecasts.

North: Infrastructure investment *is* a difficult problem, and I don't think there is any one ideal solution. Our agents use a relatively simple learning process, very similar to the other learning that you've seen presented today. They look at the past, they keep a record of what they did, and they find out what the market did. They don't know about one another's bids or involvement. They know about costs; they know how much it costs to create one thing or change another. They use a relatively simple multivariable linear model, essentially, to calculate what they're going to do next.

We think that it's a weakness in the model, that right now we have relatively thin agents, and we're working on changing that. We've talked about embedding neural network systems in the new "thick" agents, for instance. We've also been considering genetic algorithms, so we'll probably use a combination of those.

Andrews: I'm Tony Andrews, with the Navy. Many generators now are looking at futures and options to minimize their risk. How might you be able to incorporate that? I think the real test is going to be whether it will drive the price down, as they anticipate.

North: It's an open question. People like to think that those things drive prices down, but it's not really guaranteed. In the SMART models, we don't look at those issues, because we're dealing with the spot market, which is essentially short term. For FAST, though, we've created data structures for looking at different packages or bids that people can make, not only for the next hour, but also for the next year. So that's the next step.

Bower: Where does the market price come from in your model, and how do the agents influence it?

North: We have basically an ISO who's taking a series of bids into the marketplace. The agents look at what they've bid in the past, what the market price was, and whether their bids were accepted, and they use that information to adjust their current bidding. Then the ISO basically does the merit ordering to allocate people.

We deal with failures that are user-imposed in our model. When I select something from the screen and make it unavailable, for example, that could be a unit that was allocated but didn't actually run. But beyond that, we're not having explicit failures outside of user input.

[Presentation by Richter]

Cirillo: When you run the simulation and the agents are adapting to the change in the market price, have you experimented with the rate of adaption or the amount of adaption that the agents are allowed to utilize, and does that have a significant effect on the results?

Charles Richter: I've experimented with that some, yes. And it does have a significant effect.

Hong Lei: I'm Hong Lei from Warwick Business School. This is a question I've been forming over these two days about the overall topic of agent-based simulation. I believe it is, of course, useful to learn what these models do; however, it seems they are presented as closed systems. I would like to learn more about this useful methodology by knowing how the models are organized.

Cirillo: I believe the question as stated is what can we do to reveal more of the structure and inner workings of the models so that others can learn how they were put together. Anyone want to comment on that?

North: Michael North from Argonne National Laboratory. I think open source is a good step in that direction. We can't open up a source code for everything all the time — there are proprietary and competitive concerns — but when we can, I think it's the most helpful teaching tool of all. And obviously, publications where you describe a model's organization have a value, too.

Jonathan Bendor: Jon Bendor, Stanford University. Two quick comments on the issue of making the models comprehensible. I think that's a very important issue. First, we could all try to keep our models simple. Not only will the models be more comprehensible to other people, they'll be more comprehensible to us. And going back to the earlier discussion, when things go wrong, as they always will, it will be easier to fix them. So start simple and work incrementally. Resist the temptation — and here I'm seconding Ian Lustick's point enthusiastically — resist the temptation to throw in everything you know about the real world that pertains to your question.

Secondly — and this is related to the model that Diermeier and Ting and I worked on, and why we think that deduction and computation complement each other so well — what one can do, both to gain insight into your own model as well as to communicate it, is deductive work on small aspects of the model. It's probably utopian to try and solve any of these complex models completely analytically; that's why we're simulating in the first place. But what you can do is to freeze an aspect of the model and then solve little parts of it. You're triangulating via deduction and gaining insight into the properties and characteristics of the model. At the same time, you can generalize the unfrozen aspects of the model and say, "Well, this doesn't just hold for Bush-Mosteller, it holds for a whole class of adaptive rules."

So I think a useful strategy is start simply, work incrementally, and marry deduction to computation.

Lars-Erik Cederman: Another simplification would be to have standardized models to use as starting points. This was one of the goals in the Swarm project. Then you wouldn’t have to start from scratch in explaining the inner workings of your components. For instance, game theorists talk about “prisoner’s dilemma,” and you don’t have to redescribe the prisoner’s dilemma, because everybody knows exactly what is meant by that. I think we haven’t got very far along these lines because there is very little overlap. Apart from using the same basic package that relieves us from reinventing the wheel when it comes to showing graphs and setting up schedules, we have much more work to do in trying to build up clusters of modeling components or libraries. I hope that when we get together in maybe five years we will have made serious progress along these lines.

Closing Panel

CLOSING PANEL

C. MACAL, Argonne National Laboratory
J. PADGETT, University of Chicago
R. PICKER, University of Chicago

Charles Macal: At this point in the program we will wrap up with closing comments. We also invite the audience to bring up any points for discussion. It's certainly not meant to be anything more than a brief wrap-up, possibly reviewing main points of the workshop or offering general comments, perhaps raising some questions that somehow didn't get answered during the workshop. John, do you have some comments?

John Padgett: It's well known in this area that the comparative advantages of the agent-based simulation approach fall on two dimensions. The first is heterogeneity: the degree to which you care about heterogeneity of agents is the degree to which this approach looks appealing. The second is topological issues: the advantages depend on the degree to which there is a structure to the interaction rather than an averaging sort of behavior. So topology and heterogeneity are the standard arguments I hear for why this approach is useful.

Where do the models that we see here fit on those two standard dimensions? First, on the topology issue, it's interesting what Peyton Young said, that an agent-based model is instantly recognizable, qualitatively. I think it's amazing, the degree to which that statement fundamentally comes down to two-dimensional spatial topologies. You see a two-dimensional graph with some nearest-neighbor topology and entities that move around on the graph and change colors. And that's how you recognize an agent-based model.

Now, regarding a more general topology, everybody in the field understands that it's very useful. It's particularly useful when you're talking about cases of real space, and we've had some of those, such as the ecological and housing segregation models. There are many domains for which topology is terribly important, and these tend to be the domains in which you most often see success in these models — when you have some real spatial framework that makes this two-dimensional fit really work.

Miles Parker: I guess I tend to think of the 2D lattice as just a feature, a way in which the field is developing, and a natural representation that people use. So in that sense it's merely a historical accident. There's nothing that particularly reifies the 2D lattice, and there's nothing that says the 2D lattice is any worse in general than other kinds of graph configurations for these kinds of models. I'm sure that there are arguments that could be made; Rob Axtell has a recent paper on different topologies of interactions, so now a lot of work is beginning to develop in that area. I do think it's a natural place to go, but we've certainly done work in other kinds of structures, as well. I see it as simply another development that makes some of the analysis, and certainly some of the methodology, a bit simpler.

Padgett: I understand that as well. I'm just trying to point out that this historical accident has potentially self-limiting qualities that might blind us to broader application of these tools.

Christopher Langton: Yes, but as one of the speakers said, architecture matters. And because we have to study many of these systems as wholes, it's often the case that there are multiple architectures at play here. In ecology, for instance, there may be something you would like to model as a smooth flow for hydrology, but at the same time there are discrete animals moving around that you might be able to model on a lattice — and those animals might be humans with cell phones talking to each other over a telecommunications network. I think it's critical not to try to force all of the processes into a single architecture.

Padgett: I agree completely that in fact architecture does matter, but I'm trying to point out that for whatever historical reason, we are self-limiting ourselves away from that idea, which is the promise of the whole approach.

But returning to the issues of topology and heterogeneity, it's tempting to say that since both are benefits of agent modeling, then more of both is better. And that runs us right into what I'll call the "Bendor critique," that you get ever more complicated, ever more unmanageable models that way. And so my exhortation to greater richness, while I deeply believe it, is very dangerous, for the reasons that Jon has mentioned.

And I think the solution really comes down to what Lars-Erik [Cederman] said earlier. We have to identify in this huge space of possible agents a few areas that our models can actually come to grips with and say something new about. It's not enough to just populate the model world with all sorts of disparate things, which then go off centrifugally in 50,000 different directions. You have to concentrate on some sort of puzzle, which brings us back to the question: What is the theoretical framework? It's good to be focused in on the powers of these methods, but you do have to ultimately concern yourself with the theoretical framework, such as evolution, within which you are seeking to illustrate phenomena.

Two candidates have been expressed here: the evolution framework and the learning framework. We didn't have as many evolution papers as I would have liked, but certainly Blake [LeBaron's] paper fits very nicely in that framework. The learning framework is also very much lurking in the background. And I say the task at hand is to be a bit more self-conscious, not just about the tools, but about evolution and learning. What are the puzzles? What are the dynamical questions? Without that, we're just doomed to go off in all these different directions and not really congeal as a field.