

A Mathematical Analysis of Collective Cognitive Convergence

H. Van Dyke Parunak
NewVectors division of TTGSI
3520 Green Court, Suite 250
Ann Arbor, MI 48105 USA
+1 734 302 4684

van.parunak@newvectors.net

ABSTRACT

Multi-agent systems are an attractive approach to modeling systems of interacting entities, but in some cases mathematical models of these systems can offer complementary benefits. We report a case study of how the two modeling methods can profitably engage one another. The system we study [12] is an agent-based simulation of how groups of interacting entities can come to think alike. Though formal analysis of most of the models in that paper is intractable, a mean field analysis can be performed for the simplest case. On the one hand, while the formal analysis captures some of the basic features of that model, other features remain analytically elusive, reinforcing the benefits of agent-based over equation-based modeling. On the other hand, the mathematical analysis draws our attention to certain interesting features of the model that we might not have considered if we had not performed it. Responsible modeling of a domain should include both approaches.

Categories and Subject Descriptors

I.2.11 [Distributed Artificial Intelligence]: Multiagent systems;
I.6.1 [Simulation Theory]: Types of Simulation

General Terms

Measurement, Experimentation, Theory.

Keywords

Agent-based simulation, Equation-based modeling, Collective Cognitive Convergence, Mean field analysis.

1. INTRODUCTION

Formal, closed-form models of dynamical systems are attractive for a number of reasons.

- Their formulation requires the researcher to explicate the underlying processes in a way that construction of an agent-based model does not. They thus provide an account of the system that can be extrapolated more reliably than one or a few runs of an agent-based simulation.
- They provide a rigorous structure that can be related to a large body of mathematical theory (notably in nonlinear systems [9] and probability theory [6]), allowing generalizations that would be difficult to justify working solely from a simulation.

Cite as: A Mathematical Analysis of Collective Cognitive Convergence, H. Van Dyke Parunak, *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, Decker, Sichman, Sierra and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary, pp. 473–480
Copyright © 2009, International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org), All rights reserved.

- Evaluation of a set of equations is usually much more efficient computationally than running an agent-based model. When the objective of a model is to deliver results, equations can often deliver them faster.
- Perhaps most importantly, development of such a model draws attention to features of the system that might otherwise be overlooked, and that repay attention even in the agent-based form

But equation-based models are useful only if they indeed describe the same system as the simulation. Inevitably, formal models must make simplifications that are naturally avoided in the agent-based case. These simplifications often take the form of working with aggregate or average quantities instead of individual ones. In such cases the analysis is described as a “mean-field theory.” Sometimes the aggregation is over the agent population, while at other times it may be over the environment in which the agents interact. Because of the nonlinearity of agent interaction, the individual agent differences can lead to qualitatively different performance in an agent-based model from what is seen in a mean-field analysis.

The purpose of this paper is to demonstrate the interplay of analysis and simulation in studying a simple system, and thus to argue for the importance of pursuing both approaches concurrently. Our moral is that while a multi-agent model is more likely to be right in detail, an analytical model is likely to lend deeper insight to what is going on, and by developing both together, researchers can realize benefits that neither approach by itself can deliver.

We focus our attention on a system we presented at AAMAS 2008 [12]. People who work together frequently tend to adopt similar beliefs, and sometimes become blind to positions not represented in their group. This tendency has long been recognized empirically (e.g., [15]). The previous paper used a simple multi-agent model to study this dynamic, which we termed “Collective Cognitive Convergence” (CCC), but offered no formal analysis of the system. We argued, based on other research in the field, that such an analysis was likely to be intractable. In general, that claim is correct, but we can develop a formal model of a very simple form of that system. Even that simple version can demonstrate the interaction between the two models.

Section 2 reports previous work in comparing agent-based and equation-based system modeling, and describes the original CCC system. Section 3 develops our model. Section 4 compares its behavior with that of the CCC system. Section 5 concludes.

2. PREVIOUS WORK

Relevant previous work falls in three categories: the general debate concerning the relative virtues of agent-oriented and equation-based models, previous comparisons of the two kinds of models of the same system, and models of cognitive convergence, including the CCC model.

2.1 Equations vs. Agents

Equations and Agents have often been viewed as competing technologies for modeling systems. Epstein [5] argues against the widespread use of equation-based analyses of equilibrium conditions in the social sciences with the observation that a model must show how an equilibrium is reached, not just that it would be stable if reached. His motto is, “If you didn’t grow it, you didn’t explain it.” Agent-based models are his tool of choice for generating social phenomena, though the formal model we propose in this paper is generative as well. For those who insist on a mathematical model, Epstein points out that any computer system corresponds to a unique equivalent recursive (partial) function, but such a function is unlikely to offer the clarity of understanding that usually motivates the equation-based modeler.

2.2 Equations and Agents

The point of this study is that it can be helpful to study the same system both analytically and using agent-based modeling. This is not the first such effort. Axtell *et al.* [2] emphasize the importance of applying multiple modeling disciplines to the same problem, a technique they describe as model “docking.”

There have been other examples of docking equation-based and agent-based models. Parunak *et al.* [13] discuss the relation between an agent-based and an equation-based analysis of an industrial supply chain. Wilson [18] compares two models of the same ecosystem and shows how successive refinements of the equation-based model are needed to match it with the agent-based model. Shnerb [14] presents an abstract predator-prey system that exhibits qualitatively different behavior when modeled using differential equations and agents. In each of these cases, the comparison identifies unwarranted assumptions made in the equation-based model that lead to differences in behavior. Shnerb’s model is particularly relevant to our effort here, because it highlights the distinction between the use of aggregate or average quantities in an equation-based model and the idiosyncratic behavior of individually divergent agents in an agent-based model.

Our current effort is of interest because the problem is much more complex than those that have been handled previously, and as a result the simplifications needed in the mathematical model become more salient.

2.3 Collective Cognitive Convergence

This section relies heavily on material in [12].

2.3.1 The General Field

The question of how different agents converge on a common cognitive state is one of the foundational issues of MAS research, and has been studied under a number of names, including multi-agent agreement, convergence, collective agreement, convention, consensus, and game-theoretic equilibrium, and a variety of mathematical techniques have been employed to formalize them. For example, for agents with predefined (static or dynamic) connections to one another, spectral analysis of the Laplacian matrix describing their connectivity is a natural way to study how

local agreements spread to form global consensus, and Lyapunov stability analysis has been applied successfully to study the convergence of such systems [11].¹ Such analysis is particularly natural in exploring the dynamics of artificial agents, whose connectivity is usually engineered explicitly as part of the systems, and thus available for graph-theoretic analysis.

Our approach is motivated more by the need to understand the dynamics of populations of humans, whose connectivity is not readily accessible for analysis, and emerges from the consensus process rather than driving it. Empirically, groups of people who interact regularly with one another tend to converge cognitively. For more than 50 years [7], computational social science has been preoccupied with the dynamics of consensus formation [8]. Some studies are analytic, while others use simulation. They differ in the belief model and three characteristics of agent interaction (topology, arity, and preference). The CCC model, which we follow, represents a unique combination of these characteristics. In particular,

- CCC considers a vector V of m topics, rather than a single one. This model lets an agent participate in different interest groups, but greatly complicates the dynamics. With one topic, individuals move along a line, and measures such as the mean and variance of their position summarize the system’s state. In CCC, they live on the Boolean lattice $\{0,1\}^m$ of interests, and our measures must reflect the structure of this lattice.
- CCC allows many agents to interact concurrently. This model captures group interaction more accurately than does pairwise interaction, but also means that agents interact with a distribution over belief vectors rather than a single selection from such a distribution.
- CCC allows agents to modulate the likelihood of interaction based on how similar they are to their interaction partners. This kind of interest-based selection is critical to the dynamics of interest to us, but makes the system much more complex.

CCC thus takes the more complicated options along these dimensions. Even among previous work with simpler assumptions (e.g., [1, 3, 4, 10, 16]), analytic results are not always offered, and the original CCC work makes no attempt to analyze the more complicated system they present.

2.3.2 The CCC Model [12]

The original model represents the topics in which each agent can have an interest as $V \in \{0,1\}^m$. A ‘1’ at a position indicates interest in that topic, while a ‘0’ indicates a lack of interest. At each step, a randomly chosen agent (the “active agent”)

- identifies a neighborhood of other agents based on some criteria (e.g., proximity between interest vectors, geographical proximity, or proximity in a social network), in the original work, using the Jaccard distance,
- either learns from this neighborhood (by picking interest j at random, and if it is 0, setting it to 1 with probability $p_j =$ proportion of neighbors with $j = 1$), or with equal probability,
- forgets (by turning off an interest j currently at 1 to 0 with probability $1 - p_j$).

To measure a society’s convergence, the original model performs single-linkage hierarchical clustering of the population based on

¹ I am indebted to an anonymous reviewer for emphasizing this useful perspective.

the Jaccard distance between interest vectors, and measures each node's *diameter* d , the distance at which it forms in the cladogram. In a random population, d of lower-level nodes is not much less than the root's d ; in highly converged populations, lower-level nodes have d much less than the root's. The ratio of a node's d to the root's is the node's "min-max ratio" (M^2R). The median of this ratio (M^3R) measures system convergence.

2.3.3 Results

Our dominant conclusion in [12] is that collapse of the community to a single set of interests is robust under a wide range of neighborhood formation rules. These include neighborhoods of fixed size made up of the closest neighbors to the active agent, neighborhoods of variable size made up of all agents within a threshold distance to the active agent, the complete population of agents, and randomly formed neighborhoods. Three mechanisms were shown to avoid collapse: random mutations in individual agents' interest strings, agents that regularly adopt and forget interests that are the opposite of their neighbors' tendencies, and pre-formed groups that tend to converge individually to different points, but continually cross-pollinate other groups through bridge agents that belong to multiple communities.

3. AN ANALYTICAL MODEL

In this section, we introduce the structure of our model and derive two non-obvious features: conservation of interests and convergence behavior. Then we refine the model and examine the effect of the refinement on our conclusions.

3.1 Model Structure

Let there be n agents, each with an m bit string representing its interest in m possible topics. We analyze the case where all agents interact as a group (the neighborhood consists of the entire population). As a consequence, we never need to reason about the distance between agents, so the choice of the Jaccard distance in the original paper has no effect. We note in passing, however, that the normalized Hamming distance is a much more appropriate way to estimate the separation of agents. The Hamming distance is the number of positions in two agents' bit strings at which they differ, and the normalized Hamming distance divides this number by the length of the string. This measure, unlike the Jaccard, gives equal weight to agreement between two agents at a given bit position, whether they agree on 1 or on 0.

Let α_i be the i th agent, and α_{ij} the j th position in the i th agent. Thus α_{ij} can be viewed as a matrix, whose rows are agents and whose columns are topics.

Let $\pi_i(t)$ be the density of 1 bits for agent i , $\pi_i = \frac{1}{m} \sum_{k=1}^m \alpha_{ik}$. Let $\pi(t)$ be the average of this value over all agents. For clarity, we omit the time when we are discussing the situation at a fixed time. To keep our model tractable, we do not track the individual evolution of π_i , but work with π . This simplification is an example of the "mean-field" approach. The probability that an agent has a given interest is $p_1 = \pi$ and the probability that it does not have that interest is $p_0 = (1 - \pi)$.

3.2 Conservation of Interests

The critical decision variable that determines whether an agent flips a bit or not when it is activated is the probability of that bit's status in its neighborhood (the set of other agents with which the active agent is comparing itself). This quantity is attractive as the focal point for a formal theory, because it can be analyzed with a large existing body of probability theory. Thus the formal

modeling process draws our attention to a feature of the original model that affords analytical traction. The dynamics of this feature were not discussed in the original study. One contribution of the formal model will be to draw our attention to this feature and examine its behavior in the agent-based model. At the same time, we recognize that we are choosing our focal variable based on analytical tractability, not necessarily relevance to the domain under study, and this is a weakness of the analytic approach.

Consider the probability that the bit is on in the active agent. The state of the bit in each agent in the neighborhood is a Bernoulli trial of this event with probability π , so the expected number of agents with the bit on is just $n\pi$, and the probability that the bit is on is just π . Similarly, the probability of finding the bit off in the neighborhood is $(1 - \pi)$. In other words, in the α_{ij} matrix, the expected probabilities of 1 or 0 in each column are the same as the probabilities in each row.²

When an agent selects a bit, it will reverse its state in two cases. In one case, the agent finds the bit on with probability π , and turns it off with probability equal to the probability that the neighborhood has it off, which is $(1 - \pi)$. Thus the probability of a shift from 1 to 0 is $p_{10} = \pi(1 - \pi)$. In the other case, the agent finds the bit off with probability $(1 - \pi)$, and turns it on with probability that the neighborhood has it on, which is just π , so $p_{01} = (1 - \pi)\pi$. The probability of a flip is the same in each direction, so π , the agent's proportion of 1's, on average, remains unchanged. There is a "conservation of interests": the number of topics in which an agent is interested is constant, and the convergence process simply reallocates that interest to different topics.

This result is a directly testable and non-obvious prediction of our theory. One way that a community could converge would be by developing interest vectors of all 1's or all 0's. Or convergence could lead to a vector with a balance of 1's and 0's, whatever the original proportion. The original paper does not engage this question, but a "conservation of interests" is an interesting characteristic of cognitive convergence, particularly since the system has no explicit representation of cognitive capacity. It is the theory that draws our attention to this emergent behavior.

3.3 Convergence Time

Each flip brings the population closer to unanimity of interest. How many trials does this require? The overall likelihood of a flip is $p_{10} + p_{01} = 2\pi(1 - \pi)$. This quadratic assumes its maximum value for $\pi = 0.5$, suggesting that the system should converge most rapidly at this value, and less rapidly at the extremes. But this tendency must be balanced by the observation that the community's initial level of agreement will be much greater when π is far from 0.5. In the extreme case, when π is 0 or 1, though the probability of a flip is 0, the entire community is already in agreement, so the time to convergence is 0.

Can we estimate this net convergence time analytically? The measure of group similarity in [12] is derived from the distances at which agents cluster in a hierarchical clustering. This measure is costly to compute and does not lend itself to analytical treatment. We suggest an alternative: the mutual information

² The reader may want to reflect on the oversimplification in this reasoning, which we will correct later. It is instructive to see how far we can go even with this oversimplification.

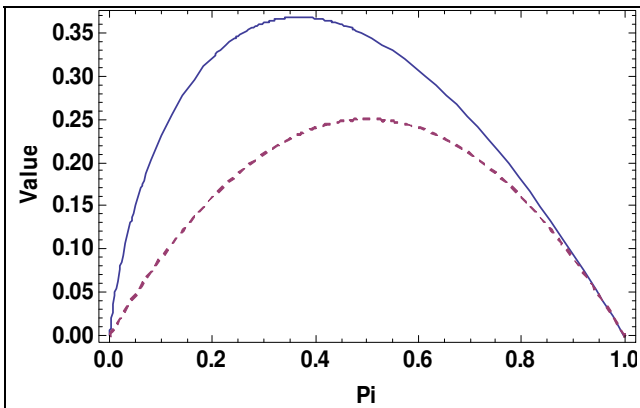


Figure 1: Mutual Information (solid line) and probability of bit change (dashed line) as function of π . *MI* is plotted for $m = n = 1$ to achieve scale comparable with probability. Higher values of m and n simply multiply *MI* by a factor of mn .

between topics and agents. The mutual information between two features (a, b) over a set of data is

Equation 1

$$MI = \sum_{b,a} p(b,a) \log \left(\frac{p(b,a)}{p(b)p(a)} \right)$$

In our case, b indexes over topics and a indexes over agents. $p(b,a)$ is the probability that the a th agent is interested in the b th topic, while $p(b)$ is the sum of this probability over all agents, and $p(a)$ is the sum over all bit positions. Consider two cases.

First, if all agents have the same interests, $p(a)$ is independent of $p(b)$. Then $p(b,a) = p(a)p(b)$, and the logarithm (and thus *MI*) vanishes.

Second, if each agent has a distinctive set of interests, $p(b,a)$ will differ from $p(a)p(b)$. Sometimes it will be less, sometimes more. When it is less, the logarithm will be negative, but will be weighted by a relatively small value of $p(b,a)$. When it is more, the logarithm will be positive, and will be weighted by the larger $p(b,a)$. The resulting *MI* will be greater than zero, with an upper bound equal to the lesser of $\log(m)$ and $\log(n)$.

Thus *MI* is an easily computed measure of the degree of diversity in an agent population. 0 indicates that all agents have identical interests, while a higher value indicates divergence.

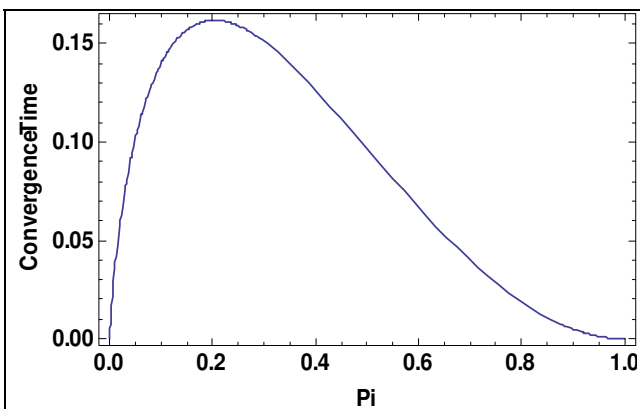


Figure 2: Nominal convergence time as function of π

We can estimate *MI* for a random matrix. There are $m\pi$ 1's in each agent, or a total of $nm\pi$ over all n agents, so the probability of finding a 1 in any one cell out of mn is just π . Summing these values over rows gives $p(b) = n\pi/n = \pi$ and summing over columns gives $p(a) = m\pi/m = \pi$. Neither is a function of a or b , so we can replace summation by multiplication, yielding $MI = mn\pi \log(1/\pi)$. This value is maximum when $\pi = 1/e \sim 0.37$, to the left of 0.5, where convergence proceeds most rapidly (Figure 1).

It is not straightforward to combine these two curves to show the net rate of convergence. They are not even in the same units. However, we can argue intuitively as follows. The slope of the *MI* curve reflects the rate at which diversity increases with π , while the slope of the other curve reflects the rate at which bit change frequency increases with π . Thus we might differentiate the two curves, subtract the second of the derivatives from the first, and integrate to get a general idea of convergence time (or equivalently, subtract the quadratic curve from that for *MI*). Figure 2 shows the result of this computation. It has two features.

First, convergence should be slower at intermediate values of π than at the extremes.

Second, it suggests that convergence will be most rapid for extreme values of π , reaching a maximum somewhat to the left of $\pi = 0.5$, with a long tail toward higher probabilities.

This latter prediction is anomalous. Why should convergence time be asymmetric around $\pi = 0.5$? Our algorithm treats agreement on 0 the same as agreement on 1. We would expect convergence at 0.1 to be the same as at 0.9.

The asymmetry in Figure 2 is an artifact of the definition of *MI*, which does not treat a 1 in the matrix the same as a 0. All three of its component probabilities ($p(b, a)$, $p(a)$, and $p(b)$) are the probability of a 1 in the matrix, not of a 0. This asymmetry leads to the asymmetry of the peak in *MI* in Figure 1, which we use to estimate the *initial* degree of diversity in a random population, and which thus leads to the asymmetry in Figure 2.

The original paper did not study the dependence of the system's dynamics on different values of π . The theory focuses our attention on this quantity because of its tractability for existing mathematical (specifically, probabilistic) tools, and that in turn encourages us to look back at the multi-agent model to see whether the convergence does indeed follow this prediction.

3.4 Column vs. Row Probabilities

The careful reader will have noticed a weakness in the preceding development. The column (topic) probabilities in the α_j matrix are the same as the row (agent) probability π only in the initial random matrix. As individual agents update bits to bring them more in line with the values of those positions (columns) in other agents (rows), the column probabilities will tend to diverge. The average column probability will continue to be π , but this average will not be representative of any individual column.

Let π continue to denote the probability of 1 in an agent (a row of the matrix), which we will assume to be constant across agents as outlined above. Denote the probability of 1 in a predominantly 1 column (topic) by π_1 , and in a predominantly 0 column by π_0 . We must now use $p_{10} = \pi(1-\pi_1)$, and $p_{01} = (1-\pi)\pi_0$. Each time an agent flips a bit, it will increase the number of 0's or 1's in the appropriate column by 1, and thus increase the relevant probability in that column by $1/n$. On average, each flip will either

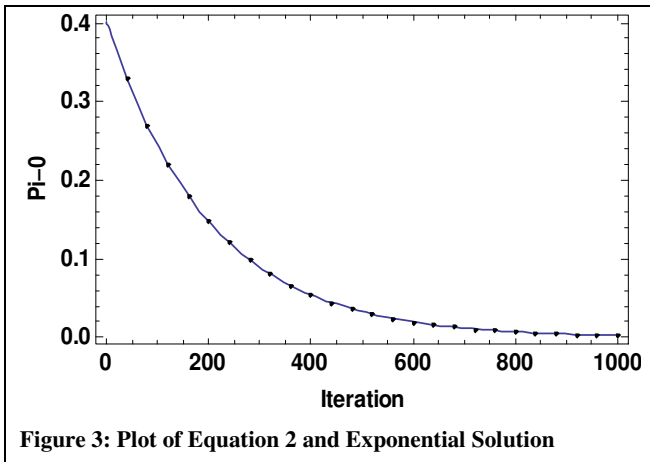


Figure 3: Plot of Equation 2 and Exponential Solution

increase the probability of 1 in a predominantly 1 column by p_{01}/n , or decrease the probability of a 1 in a predominantly 0 column by p_{10}/n .

These changes affect only a single column. To estimate their impact without tracking all columns individually, we need to know how many columns are majority 0 and how many are majority 1. We focus on majority 1 columns; the analysis for the majority 0 columns is precisely parallel. As a run evolves, to maintain the constancy of π , the proportion of majority 1 columns will tend toward π , and in the spirit of ignoring initial transients (which equation-based models are notoriously bad at capturing), we assume this is constant. Then there are πm such columns. The addition of another 1 in one of them, which happens with probability p_{01} , amounts to an average increase of probability of $1/n$ in a single column, or of $1/\pi mn$ across all of the majority 1 columns. Thus π_1 begins at π and increases on average by $p_{01}/\pi mn$ each time an agent is activated. Similarly, π_0 begins at $1 - \pi$ and decreases by $p_{10}/(1-\pi)mn$ with each activation.. The resulting difference equations are

Equation 2

$$\pi_1(t + 1) = \pi_1(t) + \frac{\pi_0(t)(1 - \pi)}{\pi mn}$$

and

Equation 3

$$\pi_0(t + 1) = \pi_0(t) - \frac{\pi(1 - \pi_1(t))}{(1 - \pi)mn}$$

Equation 4 and Equation 5 are solutions to these equations, as the reader can show by evaluating the solutions at $t + 1$ and making use of the Taylor series approximation $e^x \approx 1 + x$ for $x \ll 1$. The approximation is excellent; Figure 3 and Figure 4 show the evolution of π_0 and π_1 , respectively, for $\pi = 0.4$, $m = 10$, and $n = 20$, using Equation 2 and Equation 3. The dots are every fortieth point resulting from the iteration of the difference equations. The lines are the exponential solutions.

Equation 4

$$\pi_0(t) = \pi e^{-\frac{t}{mn}}$$

Equation 5

$$\pi_1(t) = 1 - (1 - \pi) e^{-\frac{t}{mn}}$$

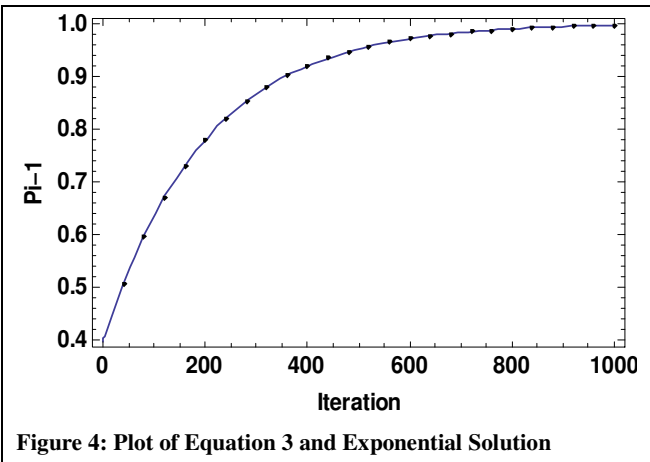


Figure 4: Plot of Equation 3 and Exponential Solution

Now we reevaluate our previous conclusions.

First, is π still constant over time (as this development has assumed)? Consider $p_{10} - p_{01}$, which was 0 under our previous assumptions. Now it is (using the exponential forms for π_0 and π_1)

$$\pi \left(1 - 1 + (1 - \pi)e^{-\frac{t}{mn}} \right) - (1 - \pi)\pi e^{-\frac{t}{mn}} = 0.$$

As in our simpler model, the difference vanishes.

Second, how does the rate of convergence depend on π ? Now we consider $p_{10} + p_{01}$, which is

$$\begin{aligned} \pi \left(1 - 1 + (1 - \pi)e^{-\frac{t}{mn}} \right) + (1 - \pi)\pi e^{-\frac{t}{mn}} \\ = \pi(1 - \pi)e^{-\frac{t}{mn}} \end{aligned}$$

At any time epoch, this quantity is quadratic in π , as before, and we expect the same qualitative convergence shown in Figure 2.

4. COMPARISON OF BEHAVIORS

We have modified our original system [12] in two ways.

1. Instead of measuring convergence by clustering the agents and reporting distances in the cladogram, we use the *MI* measure described in the previous section.
2. Instead of the Jaccard distance, we use the normalized Hamming distance to estimate the separation of agents. (However, when we are considering the neighborhood that consists of the entire agent population, and using *MI* rather than cladogram separation to measure convergence, neither

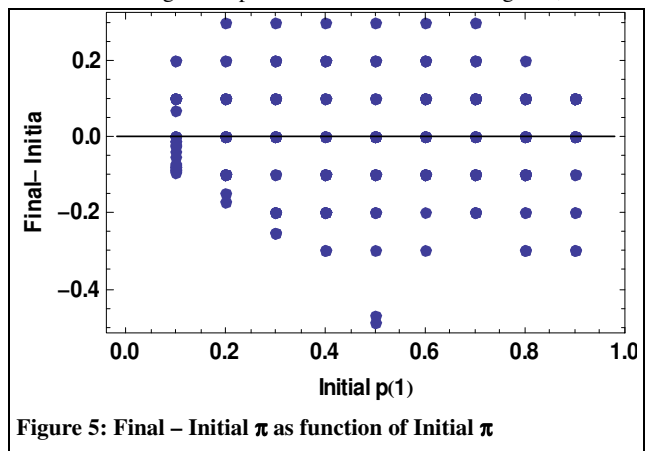


Figure 5: Final - Initial π as function of Initial π

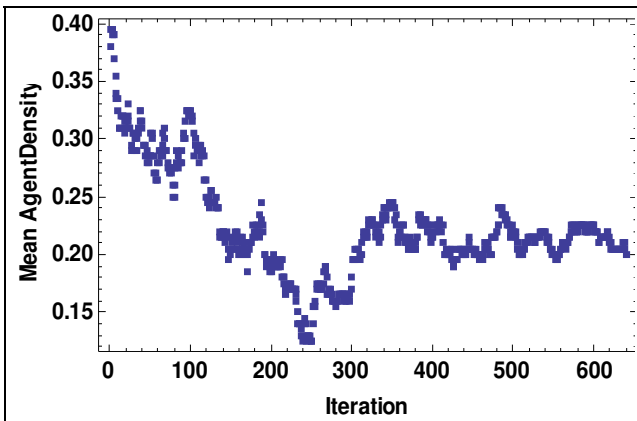


Figure 6: $\pi(t)$, single run starting at 0.4

measure enters into the computation.)

Our implementation gives qualitatively the same results over the full range of configurations as the original report. The experiments below use the same basic parameters as [12] (a population of 20 agents, each with an interest vector of length 10).

4.1 Behavior of π

Our computational model predicts that π remains constant as the system evolved. Figure 5 shows the difference between final and initial π as a function of initial π for 30 runs at each value of π in $\{.1, .2, .3, \dots, .9\}$. The line shows equality. While there is considerable dispersion around the line, the variation is evenly distributed. Of the 270 runs in this set, 79 had the same initial and final π . In 93 cases the final π was larger than the initial one (by an average of 0.14), while in 98 cases it was smaller (by an average of 0.15).

These results are consistent with the overall stability of π predicted by the formal model. However, that model does not account for the (considerable) variation in the agent-based model, and the structure in that variation evident in Figure 5 (for example, the much narrower variation at lower values of π , or the exceptionally high variation at $\pi = 0.5$).

4.2 Topic Proportions

Our theoretical development assumed, with acknowledged oversimplification, that the proportion of Majority 1 and Majority 0 topics would remain constant at π and $(1 - \pi)$, respectively. To test this, we consider a single run with initial $\pi = 0.4$.

Figure 6 shows the evolution of π in this run. It quickly drops from 0.4 to 0.3, where it stabilizes briefly, then plunges to 0.1 before finally settling at 0.2. As noted in the previous section, this variation is not inconsistent with the average tendency over many runs for π to remain constant, but this detailed behavior is inaccessible to our theory.

Figure 7 shows the proportion of majority 1 topics at each iteration. (With rare exceptions when the two are equal, the proportion of majority 0 topics in this run is just $1 -$ the proportion of majority 1 topics.) After the system stabilizes, this proportion is equal to π , and it sometimes approaches it earlier (note 0.3 between iterations 80 and 100, and 0.2 around 150), but the approximate nature of the assumption is obvious. This may be due to the wide variation of π in this particular run, which in turn is due to the specifics of the starting configuration and the

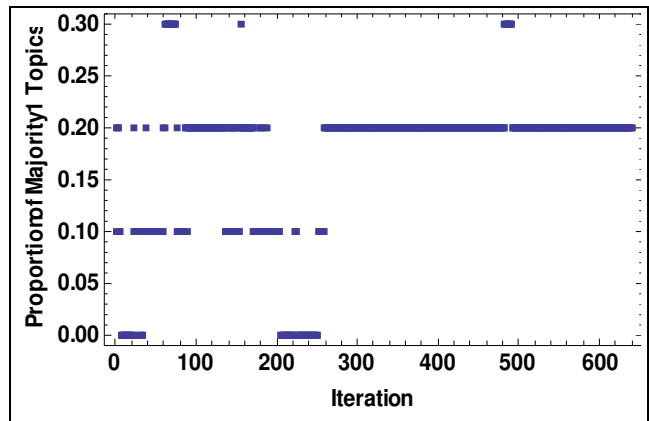


Figure 7: Proportion of Majority 1 Topics

stochastic decisions made by the agents in the course of the system's evolution.

4.3 Topic Divergence

The theory also predicted that while π_i for individual agents would grow closer together, the percentage of 1's in the majority 1 and majority 0 columns would diverge. Figure 8 confirms this behavior.

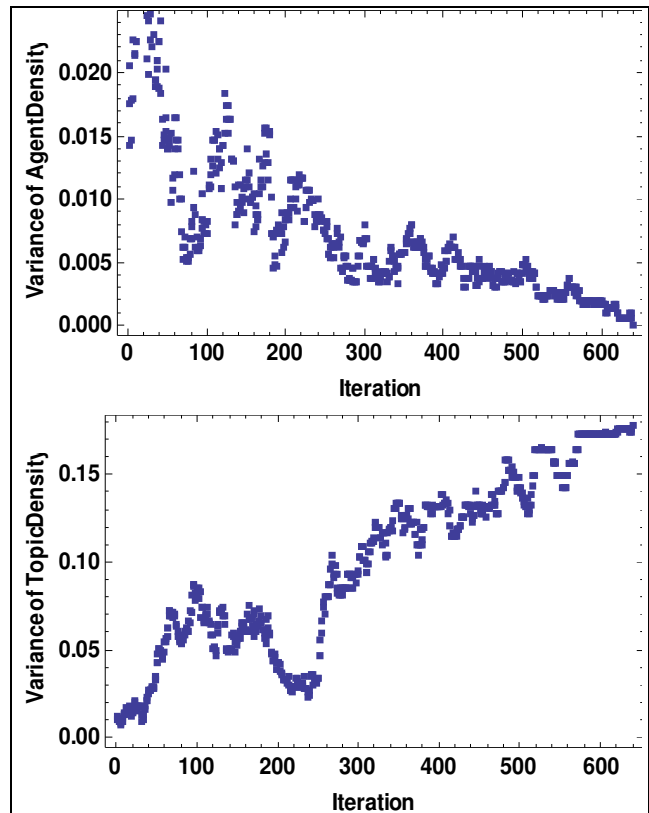


Figure 8: Variance of agent density (top) and topic density (bottom)

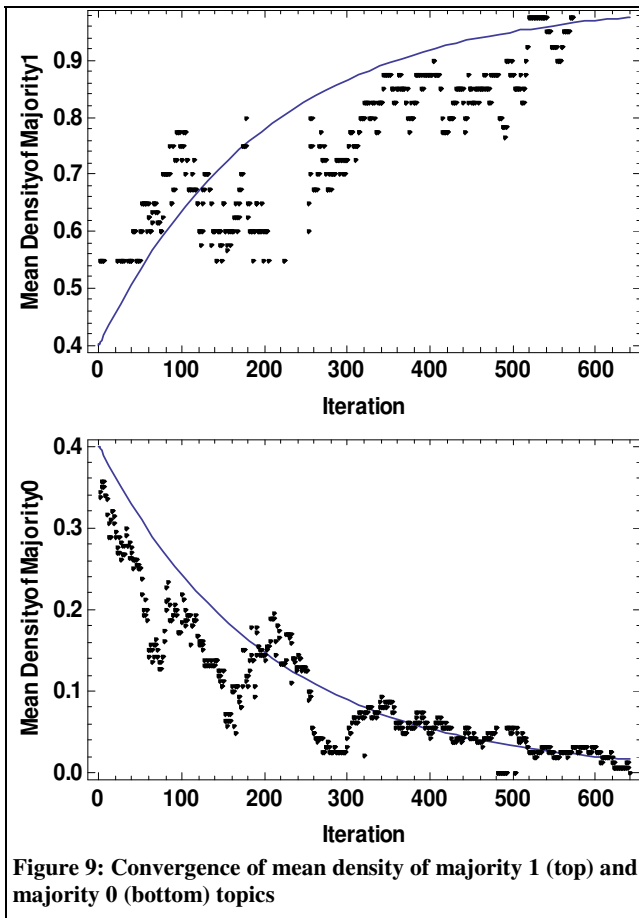


Figure 9: Convergence of mean density of majority 1 (top) and majority 0 (bottom) topics

4.4 Rate of Convergence

Equation 4 and Equation 5 are solutions to these equations, as the reader can show by evaluating the solutions at $t + 1$ and making use of the Taylor series approximation $e^x \approx 1 + x$ for $x \ll 1$. The approximation is excellent; Figure 3 and Figure 4 show the evolution of π_0 and π_1 , respectively, for $\pi = 0.4$, $m = 10$, and $n = 20$, using Equation 2 and Equation 3. The dots are every fortieth point resulting from the iteration of the difference equations. The lines are the exponential solutions.

Equation 4 and Equation 5 provide a concise description of how quickly a CCC system will converge. Figure 9 shows the convergence of the majority 1 and majority 0 bit positions over time in a single run. The exponential model is not unreasonable in the case of the majority 1 data, and quite good for majority 0, especially considering that this data is from a single run.

This convergence model in turn gives us a way to calculate the appropriate amount of noise needed to prevent collapse.

4.5 Shape of Convergence

Our model predicts that convergence speed should depend on π roughly as

shown in Figure 2. Figure 10 (top) shows the distribution of the iteration at which $MI = 0$ for 30 runs at each value of π . The distributions are plotted as box and whisker plots [17]. The box extends between the upper and lower quartiles of the data, and the whiskers span the entire range of data.

This figure shows not only the concave shape suggested by Figure 2, but also its asymmetry, which we explained as an artifact of the definition of MI . In Figure 10, MI is used to assess, not the *initial* diversity of the population (as in Figure 2), but the *final* convergence of a run (when it reaches 0). $MI = 0$ exactly when all agents have the same sets of interests. The intrinsic asymmetry of MI does not affect this condition. Yet populations that begin with $\pi = 0.1$ do converge much faster than those that begin with $\pi = 0.9$, and the slowest convergence does appear to be around 0.3.

As we have emphasized repeatedly, simulation results are subject to the wide variability of individual runs, and confirmation by independent runs is wise. Figure 10 (bottom) shows the distribution of convergence times for another set of 30 runs. Again, the general shape of the dependency on π is concave, but the appearance of asymmetry is much weaker.

Clearly, the dependence of convergence time on π merits further study.

5. CONCLUSION

Multi-agent systems and closed-form mathematical analysis are both valuable tools for modeling systems of interacting entities. Practitioners of both approaches should be conscious of the benefits of “docking” them [2] by applying them both to a single domain. Each class of model informs the other. The process is invisible in a final report such as this paper, but at innumerable points, observation of the agent-based model’s behavior suggested directions for refinement of the mathematical model, while development of the mathematical model indicated experiments on the agent-based model to test the alignment between the two.

Where a straightforward mathematical model is feasible, it can give much more insight than an agent-based model, and provide a firmer foundation for generalization than the agent-based system. However, the mean-field techniques required to achieve a concise mathematical formulation inevitably lose individual detail that the agent-based system provides in abundance. We summarize several examples of these two tendencies.

First, the mathematical model gives concise, generalizable

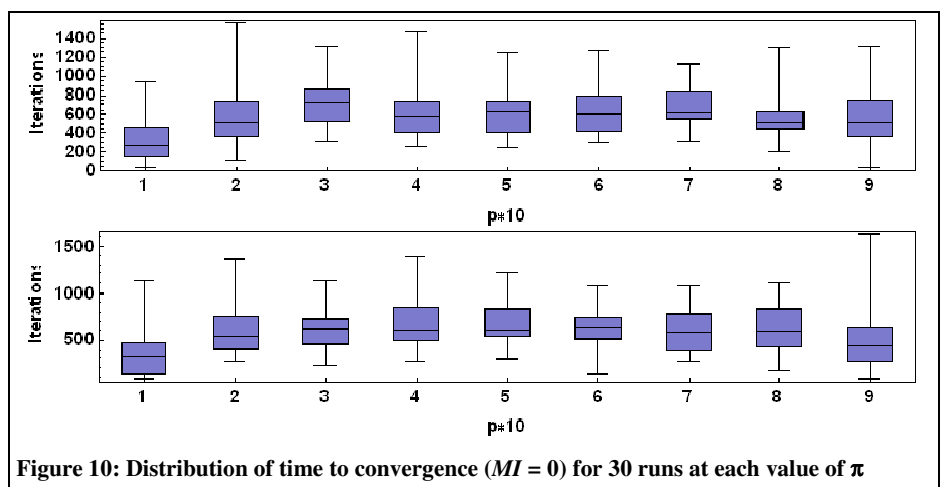


Figure 10: Distribution of time to convergence ($MI = 0$) for 30 runs at each value of π

insights. In the present exercise, these include

- The usefulness of the densities of the rows (agents) and columns (topics) of the interest matrix as focal variables for study;
- The “conservation of topics,” in which agents maintain interest in a relatively fixed number of topics, even in the absence of any explicit constraints on their cognitive capacity;
- An explicit form for the dependence of convergence time on π and the size of the interest matrix (m and n).

Discovery of these from the agent-based model alone would have required extensive sweeps of parameter space, and even then would have yielded only approximate results with limited generalizability. In particular, the elegant form of Equation 4 and Equation 5 are solutions to these equations, as the reader can show by evaluating the solutions at $t + 1$ and making use of the Taylor series approximation $e^x \approx 1 + x$ for $x \ll 1$. The approximation is excellent; Figure 3 and Figure 4 show the evolution of π_0 and π_1 , respectively, for $\pi = 0.4$, $m = 10$, and $n = 20$, using Equation 2 and Equation 3. The dots are every fortieth point resulting from the iteration of the difference equations. The lines are the exponential solutions.

Equation 4 and Equation 5 would be almost impossible to verify apart from formal analysis

Second, the agent-based model shows a richness of behavior that the mean-field formalism is unable to capture. This richness includes

- The dependence of the variability in final π on its initial value evident in Figure 5,
- The behavior of the proportion of majority 1 and majority 0 topics as the system evolves,
- The dependence of system convergence time on initial π (Figure 10). The theory could only hint at the shape of the convergence curve by vague hand-waving.

Agent-based modeling is an invaluable tool for studying complex systems. It will reach its full potential only when it is used as part of a complete toolbox that includes other approaches, such as equation-based models. These different approaches should be considered as complements to one another, and not competitors.

6. Acknowledgments

I am grateful to the coauthors of [12] for stimulating discussions.

7. REFERENCES

- [1] R. Axelrod. The Dissemination of Culture: A Model with Local Convergence and Global Polarization. *Journal of Conflict Resolution*, 41(2 (April)):203-226, 1997. <http://www-personal.umich.edu/~axe/research/Dissemination.pdf>.
- [2] R. Axtell, R. Axelrod, J. M. Epstein, and M. D. Cohen. Aligning Simulation Models: A Case Study and Results. *Computational and Mathematical Organization Theory*, 1(2):123-141, 1996.
- [3] J. Bednar, S. Page, A. Bramson, and A. Jones-Rooy. Conformity, Consistency, and Cultural Heterogeneity. In *Proceedings of Annual Meeting of the American Political Science Association*, Marriott, Loews Philadelphia, and the Pennsylvania Convention Center, Philadelphia, PA, American Political Science Association, 2006. http://www.allacademic.com/meta/p150817_index.html.
- [4] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch. Mixing beliefs among interacting agents. *Advances in Complex Systems*, 3(1-4):87-98, 2000. <http://www.lps.ens.fr/~weisbuch/mixbel.ps>.
- [5] J. M. Epstein. *Generative Social Science*. Princeton, NJ, Princeton University Press, 2006.
- [6] W. Feller. *An Introduction to Probability Theory and Its Applications, Volumes 1 and 2*. Addison-Wesley, 1968, 1971.
- [7] J. R. P. French. A formal theory of social power. *Psychological Review*, 63:181-194, 1956.
- [8] R. Hegselmann and U. Krause. Opinion Dynamics and Bounded Confidence Models, Analysis, and Simulation. *Journal of Artificial Societies and Social Simulation*, 5(3), 2002. <http://jasss.soc.surrey.ac.uk/5/3/2.html>.
- [9] H. Kantz and T. Schreiber. *Nonlinear Time Series Analysis*. Cambridge, UK, Cambridge University Press, 1997.
- [10] U. Krause. A Discrete Nonlinear and Non-autonomous Model of Consensus Formation. In S. Elaydi, G. Ladas, J. Popena, and J. Rakowski, Editors, *Communications in Difference Equations*, Gordon and Breach Science Publ., 2000. <http://www.informatik.uni-bremen.de/~krause/download/ConsensusFormation.pdf>.
- [11] R. Olfati-Saber, J. A. Fax, and R. M. Murray. Consensus and Cooperation in Networked Multi-Agent Systems. *Proceedings of the IEEE*, 95(1 (Jan)):215-233, 2007.
- [12] H. V. D. Parunak, T. C. Belding, R. Hilscher, and S. Brueckner. Modeling and Managing Collective Cognitive Convergence. In *Proceedings of The Seventh International Conference on Autonomous Agents and Multi-Agent Systems*, Estoril, Portugal, pages 1505-1508, International Foundation for Autonomous Agents and Multi-Agent Systems, 2008. <http://www.newvectors.net/staff/parunakv/AAMAS08M2C3.pdf>.
- [13] H. V. D. Parunak, R. Savit, and R. L. Riolo. Agent-Based Modeling vs. Equation-Based Modeling: A Case Study and Users' Guide. In *Proceedings of Multi-agent systems and Agent-based Simulation (MABS'98)*, Paris, FR, pages 10-25, Springer, 1998. <http://www.newvectors.net/staff/parunakv/mabs98.pdf>.
- [14] N. M. Shnerb, Y. Louzoun, E. Bettelheim, and S. Solomon. The importance of being discrete: Life always wins on the surface. *Proc. Natl. Acad. Sci. USA*, 97(19 (September 12)):10322-10324, 2000. <http://www.pnas.org/cgi/reprint/97/19/10322>.
- [15] C. R. Sunstein. *The Law of Group Polarization*. 91, University of Chicago Law School, Chicago, IL, 1999. <http://ssrn.com/paper=199668>
- [16] K. Sznajd-Weron and J. O. Sznajd. Opinion Evolution in Closed Community. *International Journal of Modern Physics C*, 11(6):1157-1165, 2000. http://www.ift.uni.wroc.pl/~kweron/articles/Sznajd-WeronSznajd00_IJMPC.pdf.
- [17] J. W. Tukey. *Exploratory Data Analysis*. Addison-Wesley, 1977.
- [18] W. G. Wilson. Resolving Discrepancies between Deterministic Population Models and Individual-Based Simulations. *American Naturalist*, 151(2):116-134, 1998.