Modeling and Managing Collective Cognitive Convergence (Short Paper)

H.V. Parunak, T.C. Belding, R. Hilscher, S. Brueckner
NewVectors division of TTGSI
3520 Green Court, Suite 250
Ann Arbor, MI 48105 USA
+1 734 302 4684

{van.parunak, ted.belding, sven.brueckner, rainer.hilscher}@newvectors.net

ABSTRACT

When the same people interact frequently, they come to think alike, a phenomenon we call "collective cognitive convergence" (C^3). We discuss instances and practical consequences of this phenomenon; review previous work in sociology, computational social science, and evolutionary biology that sheds light on C^3; define a computational model and metrics for the convergence process; report on experiments with this model and metrics; and suggest how insights from the model can help manage C^3.

Categories and Subject Descriptors

J.4 [**Computer Applications**]: Social and Behavioral Sciences – *Sociology*; I.6.3 [**Simulation and Modeling**]: Applications

General Terms

Measurement, Experimentation, Human Factors, Theory.

Keywords

Groupthink, cognitive convergence, modeling, social simulation.

1. INTRODUCTION

When the same people interact frequently, they come to think alike. We call this phenomenon "collective cognitive convergence" (C^3), since the dynamics of the *collective* lead to a *convergence* in *cognitive* orientation.

C^3 is seen in many contexts, including research subdisciplines, political and religious associations, and even persistent adversarial configurations such as the cold war. Tools that support collaboration, such as blogging, wikis, and communal tagging, make it easier for people to find and interact with others who share their views, and thus may accelerate C^3. This efficiency is sometimes desirable, since it enables a group to reach consensus more quickly. For instance, in the academy, it enables coordinated research efforts that accelerate the growth of knowledge.

But *convergence* can go too far, and lead to *collapse*, reducing the diversity of concepts to which the group is exposed and thus leaving the group vulnerable to unexpected changes. E.g.,

Cite as: Modeling and Managing Collective Cognitive Convergence (Short Paper), H.V. Parunak, T.C. Belding, R. Hilscher, and S. Brueckner, *Proc. of 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, Padgham, Parkes, Müller and Parsons (eds.), May, 12-16., 2008, Estoril, Portugal, pp. 1505-1508..

- Academic specializations can become unintelligible to non-specialists, and papers that do not fit neatly face difficulty being published. The subdiscipline is sustained more by its own interests than by the contributions it can make to the broader research community or to society at large.[1]
- The force-on-force orientation of the Cold War left both the former Soviet Union and NATO ill-prepared to deal with insurgencies and asymmetric warfare.

We became interested in this phenomenon by observing increasing balkanization in MAS research. Since we do multi-agent simulation, we thought that a multi-agent model might illuminate the phenomenon, and show how it can be managed.

Section 2 discusses previous work related to our effort. Section 3 describes our model and metrics. Section 4 outlines experiments that exhibit C^3 and explore techniques for managing it. Section 5 suggests directions for further research, and Section 6 concludes. A long version of this paper is available.[2].

2. PREVIOUS WORK

Our research on C^3 builds on and extends previous work in sociology (both empirical and theoretical) and evolutionary biology.

Empirically, groups of people who interact regularly with one another tend to exhibit C^3. One version of this phenomenon [10] is "group polarization": a group with a slight tendency toward one position will become more extreme through interaction.

For more than 50 years [4], computational social science has long been preoccupied with the dynamics of consensus formation [6]. Some studies are analytic, while others use simulation. They differ in the belief model and three characteristics of agent interaction (topology, arity, and preference). Our work represents a unique combination of these characteristics. In particular,

- We consider a vector V of m beliefs, rather than a single belief. This model lets an agent participate in different interest groups, but greatly complicates the dynamics. With one belief, individuals move along a linear continuum, and measures such as the mean and variance of their position summarize the system's

[1] This paper was motivated by frustration voiced in the industry track at AAMAS07 about how some subdisciplines of agent research were becoming so intellectually ingrown that it was difficult or impossible to apply them to real problems.

[2] www.newvectors.net/staff/parunakv/AAMAS2008M2C3.pdf

state. In our case, they live on the Boolean lattice $\{0,1\}^m$ of interests, and our measures must reflect the structure of this lattice.

- We allow 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.
- We allow 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.

Selecting the more complicated options along these dimensions makes analytic results, accessible with some (but not all) simpler models, elusive. We focus on simulation, developing intuitions for future analytical exploration.

The subgroups that form and cease to interact when convergence turns to collapse resemble biological species, which do not interbreed. So we look for insight to research in the field of biological speciation (see [2, 5] for reviews). Our model resembles runaway sexual selection speciation with mutual mate choice. We assume a homogenous environment, no physical barriers for the exchange of ideas and a symmetric "mating system." In our model, a preference for extreme traits is modeled as the probability of adopting an interest based on its prevalence in a given neighborhood. A successful runaway process in our model corresponds to the development of specializations with little practical relevance.

3. A MODEL AND METRICS

We represent each agent's interests as $V \in \{0,1\}^m$. A '1' at a position means that the participant is interested in that topic, while a '0' indicates a lack of interest. At each step, each 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),
- 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 $1 - p_j$).

We view interests as fundamentally social constructs, persisting only when maintained. Thus an isolated agent will eventually lose interest in everything. V tends to 0^m, and p_j to 0 $\forall j$. Alternative assumptions are possible.

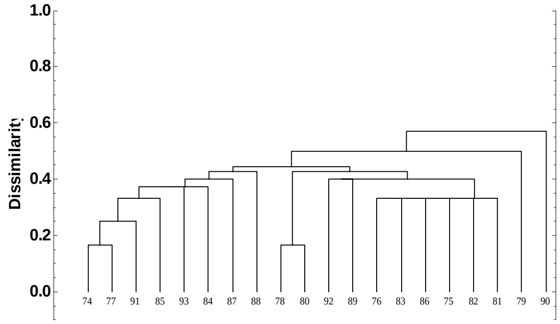

Figure 1 Random interest vectors

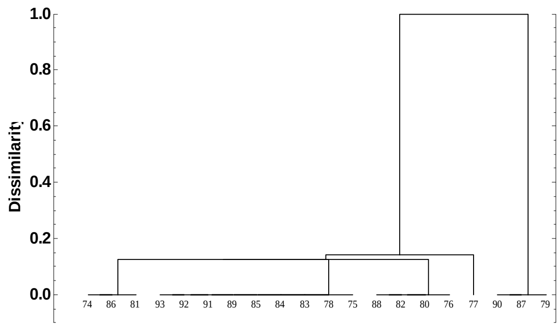

Figure 2 A highly converged population

Sophisticated statistics can estimating a group's consensus, based on questionnaires [9]. Our purposes require a mechanically computable measure of agent convergence. We perform single-linkage hierarchical clustering of the population based on the Jaccard distance between interest vectors, and measure 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* (Figure 1); in highly converged populations, lower-level nodes have *d* much less the root's (Figure 2; agents grouped at $d = 0$ have identical interest vectors). 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. $M^3R = 0$ (Figure 2) means that more than half of the agents belong to groups with identical interest vectors. We also record the maximum diameter D of the clustered population at each generation. Convergence can lead to a D = 1 (when the population fragments into groups with orthogonal interest vectors that collectively span the interest space), a low value, asymptotically 0 (when all agents collapse toward a single point in the interest space), or intermediate values (when groups have overlapping interests but no way to communicate about them to drive further convergence).

Figure 3 shows M^3R over a run with 20 agents and interest vectors of length 10, where the probability of learning and forgetting is equal, and where agents are considered to be in the same group if the similarity between their interest vectors is greater than $\theta = 0.5$. It takes only about 80 steps for M^3R to reach 0. Figure 2 shows this system at generation 300. By generation 370 it has collapsed into two groups of completely homogeneous agents of sizes 3 and 17 respectively. Their interest vectors are orthogonal (D = 1), so the agents still cover the entire interest space, but because they interact only with the agents nearest themselves in that space, they form separate islands.

4. EXPERIMENTS

Forming neighborhoods based on similar interests leads to collapse. Surprisingly, so do other sorts of neighborhoods.

4.1 Things that Don't Work

Perhaps highly tolerant agents might be robust to convergence. Let two agents consider one another neighbors if their similarity is greater than 0 (they have at least one bit position in common). This configuration might model a conference with only plenary sessions. The population still collapses—this time, toward D = 0, with the entire population at a single point in interest space.

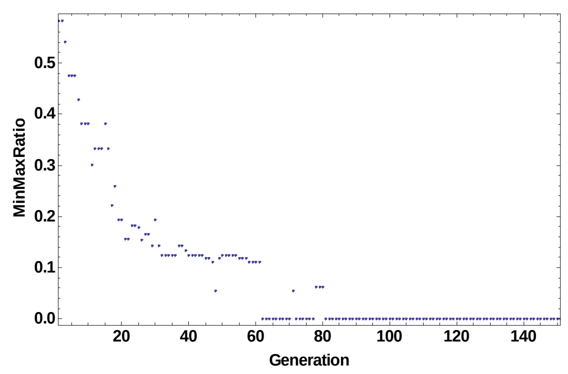

Figure 3 Evolution of 20 agents with length-10 interest vectors, neighborhoods defined by similarity > 0.5

Perhaps the problem is that as agents converge, their neighborhoods grow. So we define an agent's neighborhood at each turn as the four closest agents. This configuration models a conference with separate topical tracks. It corresponds to sympatric speciation: the assortative component is provided by the preference for partners with similar interests, while the limit on group size encourages diversity. Though agents base their adaptation at each turn on only 20% of the other agents, M^{3}R still goes to zero, as agents form subgroups within which interests collapse. D freezes at an intermediate value (here, 0.6). The population has lost some but not all variation, but the selection of partners by interest proximity means that agents never interact with those who differ with themselves.

Even more radically, let an agent's neighbors at each step be four randomly chosen agents. Imagine a conference at which papers are assigned to tracks, not by topic, but randomly. In spite of the resulting mixing, the population again collapses to D = 0.

These examples differ in how long it takes to converge to M^{3}R = 0. The time to convergence is highly variable, even within a single configuration. The one constant across runs is that the system does converge, in fewer than 500 generations (often far fewer).

4.2 Adding Variation

So far, we have no mechanism for introducing variation. Once agents grow similar, they cannot diverge. We have explored three mechanisms for adding variation to the population: random mutation, curmudgeons, and interacting subpopulations.

The simplest approach is mutation. At each generation, with probability p_m, after learning or forgetting, the active agent flips a randomly selected bit, modeling spontaneous agent curiosity. Figure 4a has the same parameters as Figure 3 (neighborhoods defined by θ = 0.5), but with p_m = 0.03. Mutation reintroduces variation, but the level is critical. If mutation is too low (say, 1%), it is unable to keep up with the pressure to convergence, while if it is too high (10%), the community does not exhibit any convergence at all (and in effect ceases to be a community). The nature of its contribution follows a clear pattern. When it is in the critical range, the system occasionally collapses to M^{3}R = 0, but then discovers new ideas. D converges to 1, since even when mutation is too low to avoid collapse within groups, it can introduce new interest vectors orthogonal to the converged groups.

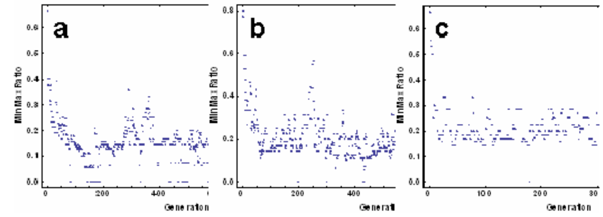

Figure 4 Variation from mutation (a), curmudgeons (b), fixed groups (c)

A curmudgeon is someone who regularly questions the group's norms and assumptions. Sunstein [10] observes that "group members with extreme positions generally change little as a result of discussion," restraining the polarization of the group. Ordinary agents learn by flipping a 0 bit to 1 with probability p_j, the proportion of neighbors with bit j on, and forget by flipping a 1 bit with probability equal to $1 - p_j$. To model curmudgeons, when an agent decides to learn or forget, with probability p_{cur}, it reverses these probabilities. Its probability of forgetting when it is curmudgeonly is p_j (instead of $1 - p_j$ in the non-curmudgeonly state), and its probability of learning is $1 - p_j$.

Figure 4b has 10% curmudgeons, again with the configuration of Figure 3. The system converges, but seldom to M^{3}R = 0. p_{cur} achieves this balancing effect over a wider range than p_m. D tends to 1. As much as researchers may resent reviewers and discussants who "just don't get it," curmudgeons can effectively and robustly save a community from collapse.

The third source of variation is even more robust, and endogenous rather than exogenous. So far, our agents choose new neighbors at every step. What if we assign each agent to a *fixed* group of initially nearby neighbors?

The behavior depends on the structure of the graph induced by a given θ. The number of components in populations of 20 agents with 10 interests each shifts suddenly from many components at θ = 0.6 to a few at θ = 0.55, the well-known phase transition in random graphs in which a giant connected component emerges as the number of links increases [3]. Consider four cases.

- High θ yields 20 components, one per agent. As discussed above, all agents independently approach $V = 0^m$.
- For low θ, the agents form one large group and collapse.
- At intermediate θ above the phase shift, the agents clump into small disjoint components. For example, one run at θ = 0.7 yielded two groups of size 3, three of size two, and eight of size one. Each of these groups evolves independently, yielding high diversity among groups (D = 1) but collapse within groups (M^{3}R = 0). This model corresponds to allopatric speciation, where physical separation allows separate evolution.
- For intermediate θ below the phase shift, the agents form neighborhoods, but some agents (e.g., 20 in Figure 5) belong to more than one neighborhood. Because neighborhoods are fixed, each can converge relatively independently of the others, but bridging agents repeatedly displace each neighborhood's equilibrium with the emerging equilibrium of another group.

This interplay of separate but linked groups yields convergence without collapse (Figure 4c). Like curmudgeons and unlike mutation, it is robust to intermittent collapse. It reflects subdisciplines that recognize the value of members who bridge with other subdisciplines and exchange ideas between them.

Such members are likely to be tolerated better than curmudgeons by subgroups, because the source of the variation introduced by the bridging individuals is perceived as resulting from multidisciplinary orientation rather than orneriness. Under fixed groups, D tends to 1.

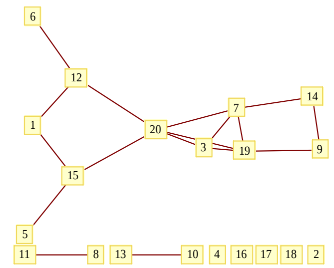

Figure 5 Agent graph, θ = 0.5

5. FUTURE WORK

Our simple model has shown a surprisingly rich space of behaviors. A number of directions for further work suggest themselves.

- An analytical model of C^3 (cf. [8]) would suggest additional mechanisms for monitoring and avoiding collapse. Existing work on the mathematics of biological speciation offers a promising foundation for this analysis.
- How can convergence be monitored? Our metric, while effective for simulation, is impractical for actual groups of people. Explicit questionnaires [9] are appropriate for experimental setting but cumbersome in monitoring groups "in the wild." One might monitor the amount of jargon that a group uses, or lack of innovation, as indicators of convergence. A promising example of initial work in this area is Schemer [1].
- What is the ideal degree of convergence, to allow the production of specialist knowledge without fostering collapse?
- How does convergence vary with group size? Recent work [7] suggests that convergence requires specialized knowledge in small groups but more general knowledge in large ones.
- We have assumed homogeneous tendencies to learn, forget, mutate, or behave curmudgeonly over all agents. How does the system respond if agents vary on these parameters? In particular, what is the impact of these parameters for bridging individuals in comparison with non-bridging individuals?

6. CONCLUSION

People naturally converge cognitively. This convergence facilitates mutual understanding and coordination, but can lead the group to collapse cognitively, becoming blind to viewpoints other than their own. Experiments with a simple agent-based model show that seemingly obvious mechanisms do not check this tendency. In the domain of academic conferences, these well-intended mechanisms include plenary sessions, special tracks, or even random mixing. A source of variation must counteract the natural tendency to converge. Mutation is effective if just the right amount is applied, but allows intermittentl collapse. Curmudgeons are more robust, but socially distasteful. Perhaps the most desirable mechanism consists of bridge individuals who provide interaction between individually converging subpopulations. These individuals arise when groups are well-defined, but have thresholds for participation low enough that some individuals can participate in multiple groups.

Insights from this simple model can give guidance in monitoring and managing collaboration. For example, consider the problem of academic overspecialization. Topical conference tracks can contribute to collapse. Their narrow focus is enhanced by conventional practice in selecting reviewers for each paper who are experts in the domain of the paper. Papers must be well aligned with the subdiscipline to rank high with such experts, and bridging papers are at a disadvantage. One might require one reviewer for each paper to be a senior researcher (thus capable of discerning high quality in problem formulation and execution) but *not* a member of the paper's main topic (and thus more tolerant of cross disciplinary results). Such a scheme might encourage acceptance of quality papers that would otherwise fall in the cracks between subspecialties. The presence of these papers in topically-organized conference tracks would then provide the bridging function that avoids collapse in our experiments.

7. REFERENCES

[1] C. Behrens, H.-S. Shim, and D. Bassu. Schemer: Consensus-Based Knowledge Validation and Collaboration Services for Virtual Teams of Intelligence Experts. In R. L. Popp and J. Yen, Editors, *Emergent information technologies and enabling policies for counter-terrorism*, pages 209-230. John Wiley & Sons, Hoboken, NJ, 2006.

[2] J. A. Coyne and H. A. Orr. *Speciation*. Sunderland, MA, Sinauer Associates, 2004.

[3] P. Erdös and A. Rényi. On the Evolution of Random Graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17-61, 1960.

[4] J. R. P. French. A formal theory of social power. *Psychological Review*, 63:181–194, 1956.

[5] D. J. Futuyma. *Evolutionary Biology*. 3 ed. Sundarlund, MA, Sinauer, 1998.

[6] R. Hegselmann and U. Krause. Opinion Dynamics and Bounded Confidence Models, Analysis, and Simulation. *Journal of Artifical Societies and Social Simulation*, 5(3), 2002.

[7] G. Palla, A.-L. Barabási, and T. Vicsek. Quantifying social group evolution. *Nature*, 446:664-667, 2006.

[8] H. V. D. Parunak, S. A. Brueckner, J. A. Sauter, and R. Matthews. Global Convergence of Local Agent Behaviors. In *Proceedings of Fourth International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS05)*, pages 305-312, ACM, 2005.

[9] A. K. Romney, S. C. Weller, and W. H. Batchelder. Culture as Consensus: A Theory of Culture and Informant Accuracy. *American Anthropologist*, 88(2 (June)):313-338 1986.

[10] C. R. Sunstein. The Law of Group Polarization. 91, University of Chicago Law School, Chicago, IL, 1999. http://ssrn.com/paper=199668