Cognitive Collapse: Recognizing and Addressing the Hidden Threat in Collaborative Technologies

Presentation at AFCEA’08

H. Van Dyke Parunak
Ted Belding
Rainer Hilscher
Sven A. Brueckner

Overview

- The Phenomenon of C³
- Some Relevant Research
- An Agent-Based Model of C³
- Conclusion
The Phenomenon of Collective Cognitive Convergence (C³)

- Repeated interactions among members of the same group makes them “think alike”
- Examples:
  - Research Communities
  - Political/Religious Associations
  - Persistent Adversarial Configurations (e.g., Cold War)
  - Intelligence/Economic Analysts Studying a Topic
- Pro: creates efficiencies through faster consensus
- Con:  
  - Reduced diversity of concepts
  - Resistance to alternative views
  - Increased vulnerability to unexpected change
  - In extreme cases, group can experience cognitive collapse

Web 2.0 (wikis, blogs, tags) and Web 3.0 (content-based VIG identification) enhance ability of like-minded people to find each other and converge.
Overview

• The Phenomenon of C³
• Some Relevant Research
• An Agent-Based Model of C³
• Conclusion

Some Relevant Research

• Sociology
  – C.R. Sunstein “group polarization”: a group with a slight tendency toward one position will tend to become more extreme in its position through interaction
• Computational Social Science
  – Axelrod’s adaptive culture model: emergence of disjoint regions of cultural (cognitive) homogeneity as agents interact with those who are adjacent to them spatially
  – Studies of consensus formation (see table, next slide)
  – Bednar et al.: competition of desire for internal consistency and external conformity slows convergence
• Evolutionary Biology
  – Speciation Models: C³ = runaway sexual selection speciation model with mutual mate choice
  – Sewall Wright “Shifting Balance Theory”: subdivided population with intermittent migration could exhibit more rapid evolutionary change than a single cohesive breeding population
### Studies of Consensus Formation

<table>
<thead>
<tr>
<th>Study</th>
<th>Belief</th>
<th>Topology</th>
<th>Arity</th>
<th>Preference?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Krause</td>
<td>Real variable</td>
<td>Random</td>
<td>Many</td>
<td>Yes</td>
</tr>
<tr>
<td>Sznajd-Weron</td>
<td>Binary variable</td>
<td>Lattice</td>
<td>Two</td>
<td>No</td>
</tr>
<tr>
<td>Deffuant</td>
<td>Real variable</td>
<td>Random</td>
<td>Two</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Binary vector</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Axelrod</td>
<td>Nominal vector</td>
<td>Lattice</td>
<td>Two</td>
<td>Yes</td>
</tr>
<tr>
<td>Bednar</td>
<td>Nominal vector</td>
<td>Random</td>
<td>Many</td>
<td>No</td>
</tr>
<tr>
<td>This paper</td>
<td>Binary vector</td>
<td>Random</td>
<td>Many</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Analytical results available only for
- Single real variable \( x(t+1) = Ax(t) \), no preferences (common)
- Six or fewer agents with preferences (Krause)
- Vector belief but binary interaction and no preferences (Bednar)

### Overview

- The Phenomenon of C³
- Some Relevant Research
- An Agent-Based Model of C³
  - Specification of the Model and Metric
  - Baseline Experiments
  - Experiments with Group Characteristics
  - Experiments with Variation
- Conclusion
A Simple Model
(Specification)

a. Agent = vector of binary interests

b. Agent distance = Jaccard distance

\[ |A \cap B| \] \[ |A \cup B| \] \[ 1-(2/5) = 0.6 \]

Community(A) = all agents within \( \theta \) of A
CommunityFocus(j) = % of agents in community with bit \( j = 1 \)

c. Learning
With probability \( p_{\text{Learn}} \)
- Pick random bit \( j \)
- If interest \( j = 0 \), set to 1 with probability CommunityFocus(j)

d. Forgetting
With probability \( p_{\text{Forget}} \)
- Pick random bit \( j \)
- If interest \( j = 1 \), set to 0 with probability 1 - CommunityFocus(j)

A Simple Model
(Metrics)

Cladogram of random interest vectors of length 10
- Median min-max ratio = 0.583
- Diameter = 0.57
Overview

• The Phenomenon of C³
• Some Relevant Research
• An Agent-Based Model of C³
  – Specification of the Model and Metric
  – Baseline Experiments
  – Experiments with Group Characteristics
  – Experiments with Variation
• Conclusion

Common Experimental Parameters

• 20 agents
• Interest vectors of length 10
• \( p_{\text{Learn}} = p_{\text{Forget}} = 0.9 \)
A Simple Model
(Experiments)

Neighborhoods defined by $\theta > 0.5$

Collapse of Groups after 60 Interaction Cycles

A Simple Model
(Metric)

Cladogram of converged interest vectors
- 300 steps
- median min-max ratio = 0.0
- By step 370, collapses into two groups, sizes 3 and 17
Overview

- The Phenomenon of C³
- Some Relevant Research
- An Agent-Based Model of C³
  - Specification of the Model and Metric
  - Baseline Experiments
  - Experiments with Group Characteristics
  - Experiments with Variation
- Conclusion

A Simple Model
(Experiments with Grouping Characteristics)

Neighborhood(s) defined by $\theta = \text{zero}$ (conference with only plenary sessions)

Collapse of Groups after 190 Interaction Cycles
A Simple Model
(Experiments with Grouping Characteristics)

Fixed-size neighborhood (4 closest) (conference with focused tracks)

Collapse of Groups after 55 Interaction Cycles

A Simple Model
(Experiments with Grouping Characteristics)

Neighborhood = 4 randomly chosen agents

Collapse of Groups after 85 Interaction Cycles
Overview

- The Phenomenon of C³
- Some Relevant Research
- An Agent-Based Model of C³
  - Specification of the Model and Metric
  - Baseline Experiments
  - Experiments with Group Characteristics
  - Experiments with Variation
- Conclusion

A Simple Model
(Experiments with Variation)

θ > 0.5, 3% mutation (1% too low, 10% too high)

Collapsed Groups
Repeatedly Break Up
A Simple Model
(Experiments with Variation)

$\theta > 0.5$, 10% curmudgeons

Groups do not Collapse

Static neighborhoods based on threshold on initial similarity

phase transition to Giant Connected Component
A Simple Model
(Experiments with Variation)

20 agents with length-10 interest vectors, neighborhoods defined by similarity > 0.5

Groups do not Collapse

0.5 threshold for static groups

Overview

- The Phenomenon of C³
- Some Relevant Research
- An Agent-Based Model of C³
- Conclusion
Conclusion

- Collective Cognitive Convergence is a natural phenomenon
- $C^3$ facilitates understanding and coordination, but may lead to cognitive collapse, creating blind spots
- Naïve mechanisms (e.g., mixing, globalization, controlling group size) do not prevent collapse
- Management requires a source of variation
  - **Mutation** requires tuning and leads to intermittent collapse
  - **Curmudgeons** are more robust, but socially distasteful
  - **Bridging** interdisciplinary individuals seem to work best

Questions?