Computational Social Science and Complex Military Operations

Aaron B. Frank
Department of Computational Social Science
George Mason University

Background

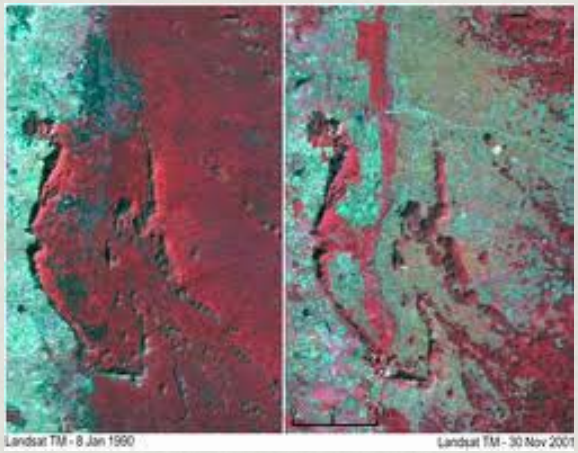
- Complex military operations are those whose execution demands interagency, transnational, multidisciplinary participation and perspectives
  - Interdependence between autonomous organizations
  - Fundamental diversity and disparity of social systems and beliefs about them
- While the challenges posed by complex military operations have been brought to forefront of the analytic community, they are nothing new
  - Maximization of force vs. the continuation of policy by violent means
  - Political and military objectives are not the same and may conflict
- Problems are fundamentally social and interactive
  - Collective and strategic interaction and communication
  - Social processes, memory and learning, identity

What is Computational Social Science?

- CSS is a discipline in process of formation
- “Big Tent” and “Small Tent” perspectives
  - Not directly important to DOD analytic community, but has implications for structure of a CSS “market” of services
  - Know what you are buying!

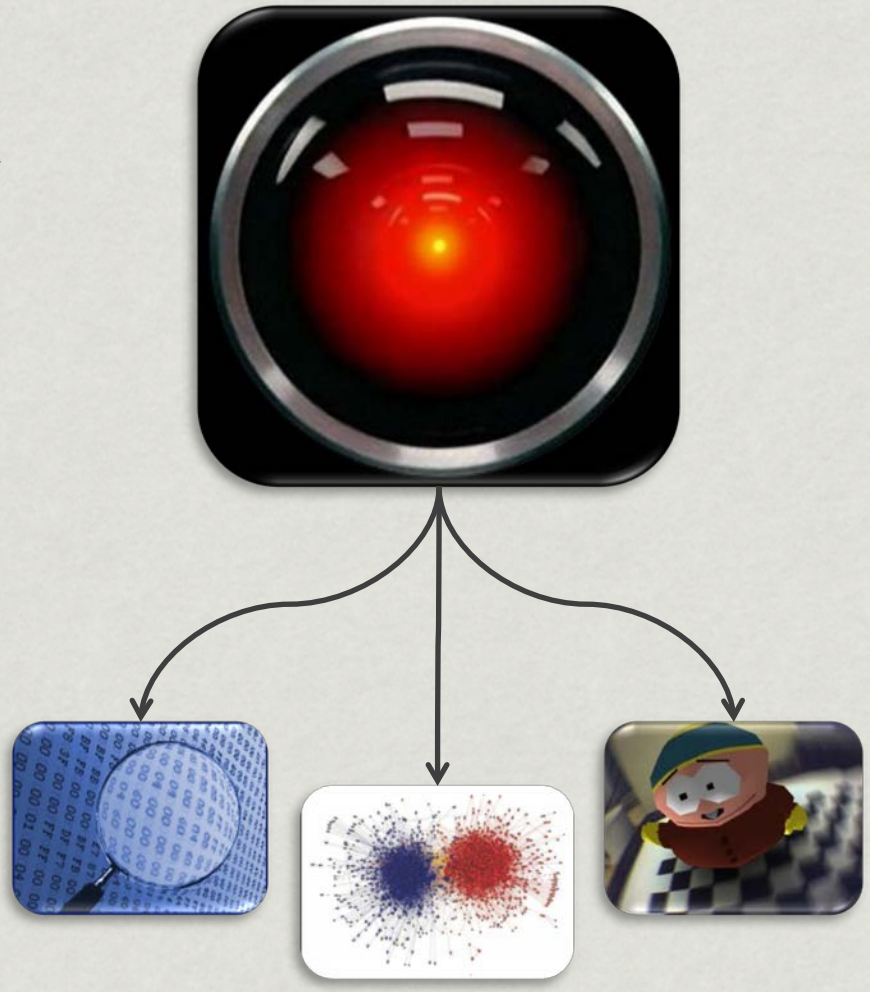
Social or Not: Land Use/Cover Modeling

- Cellular Automata (CA) vs. Agent-Based Modeling (ABM) approaches
- Consider how land changes from rural to urban
  - A CA model will examine how cells on a grid change land use types based on local conditions
    - Each cell asks: If my neighbors are urban than I will become urban with probability *x*
  - An ABM will examine how parcels of land will be converted to different types based on their owning agent's choices
    - Land owners make decisions based on market conditions, social norms, law, etc.
- Both models are of a social subjects, but are both social models?



The ‘Big Tent’ Approach

- Three research approaches may be considered CSS
  - Data mining and analytics
  - Social network analysis
  - Social simulation
- Computers are essential to the conduct of the research
  - Not just tools for data gathering, word processing or email machines



Simulation and the Social Sciences

- Many different kinds of simulation models have been used in the social sciences
  - Manual or seminar simulations
  - Mathematical simulations
  - Computational simulations
- Different treatment of behavior, data, and theory
  - All move a model forward through time

What are Agent-Based Models?

- ABMs are composed of autonomous, interacting units
  - Agents have *attributes* and *behaviors*
  - Attributes and behaviors can be *dynamic* and change as a result of *feedback*
- ABMs have environments that provide resources and mediate interactions
  - Spatial, network positions
  - Sources of wealth, power or advantage in interactions
- The simulation of the individual entities computes two levels of analysis
  - The macroscopic state of the system and its emergent properties
  - The microscopic life history of individual agents and the environment

How it Works

Why use an Agent-Based Model?

- Allows for the observation of systems and their dynamics that result from bottom-up interactions and choices
- Provides an opportunity to explore relationships and behaviors not well represented by traditional mathematical models
  - Relaxation of ‘heroic assumptions’
- Enables the simultaneous incorporation of:
  - Heterogeneity
  - Bounded-rationality
  - Explicit space or interaction topology (i.e. networks)
  - Non-equilibrium dynamics and path dependence
    - Journey vs. destination

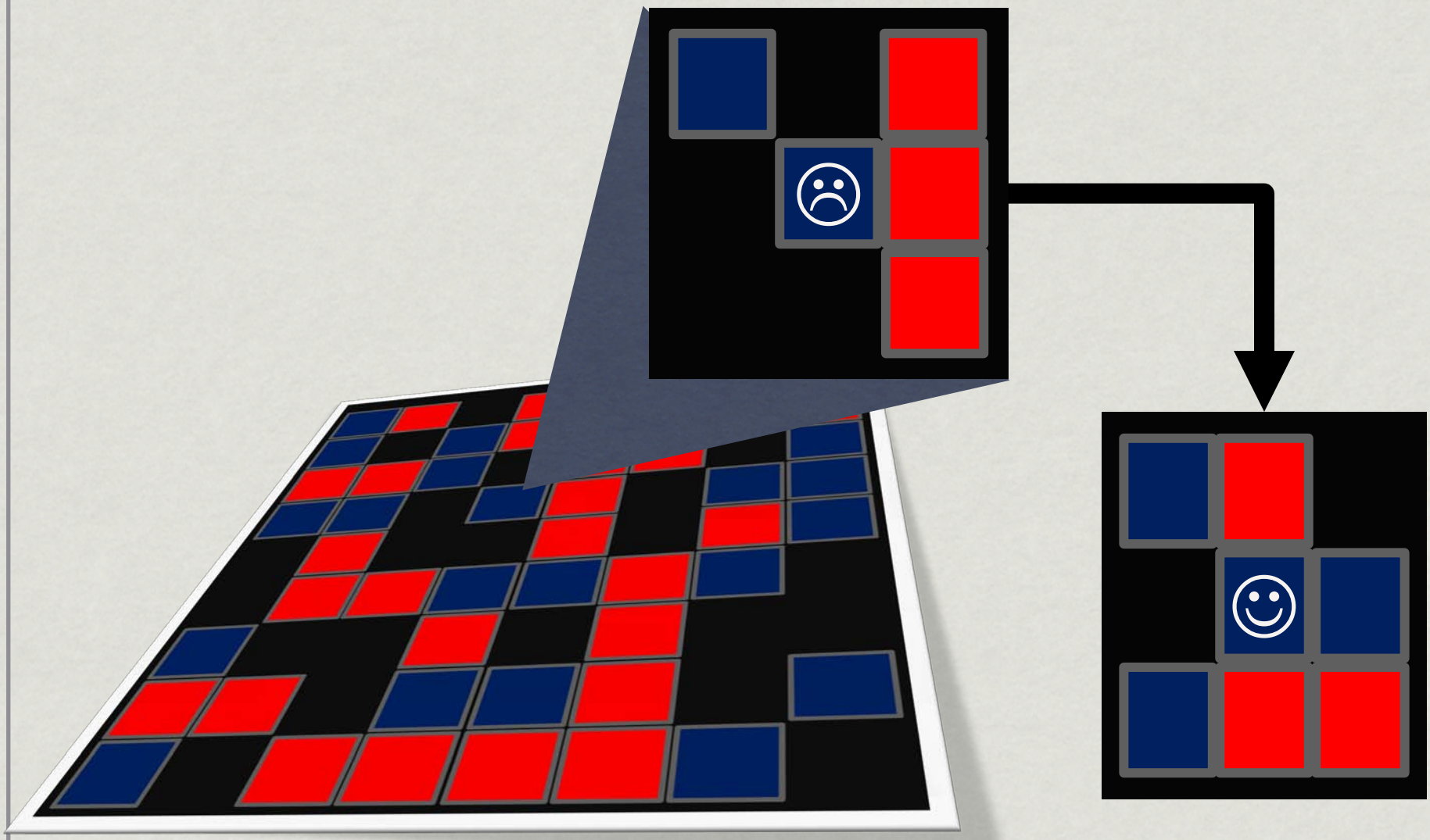
ABM and the Strategic Roots of ORSA

- Early ORSA applications were highly informal, simple models that illuminated key features of military problems
  - Focused on the trade-offs of alternative choices and strategic dynamics
  - First OR study conducted in two hours in 1940
- Formalization of ORSA into a set of methods and a scientific discipline let to split and emergence of Net Assessment
  - Non-linearity
  - Adaptation and evolution
  - Human and organizational factors
  - Context, history, and culture
  - Comprehensive, interdisciplinary sense of strategic balance and competition
- ABM specifically, and CSS techniques generally, provide an opportunity reintegrate ORSA and Net Assessment

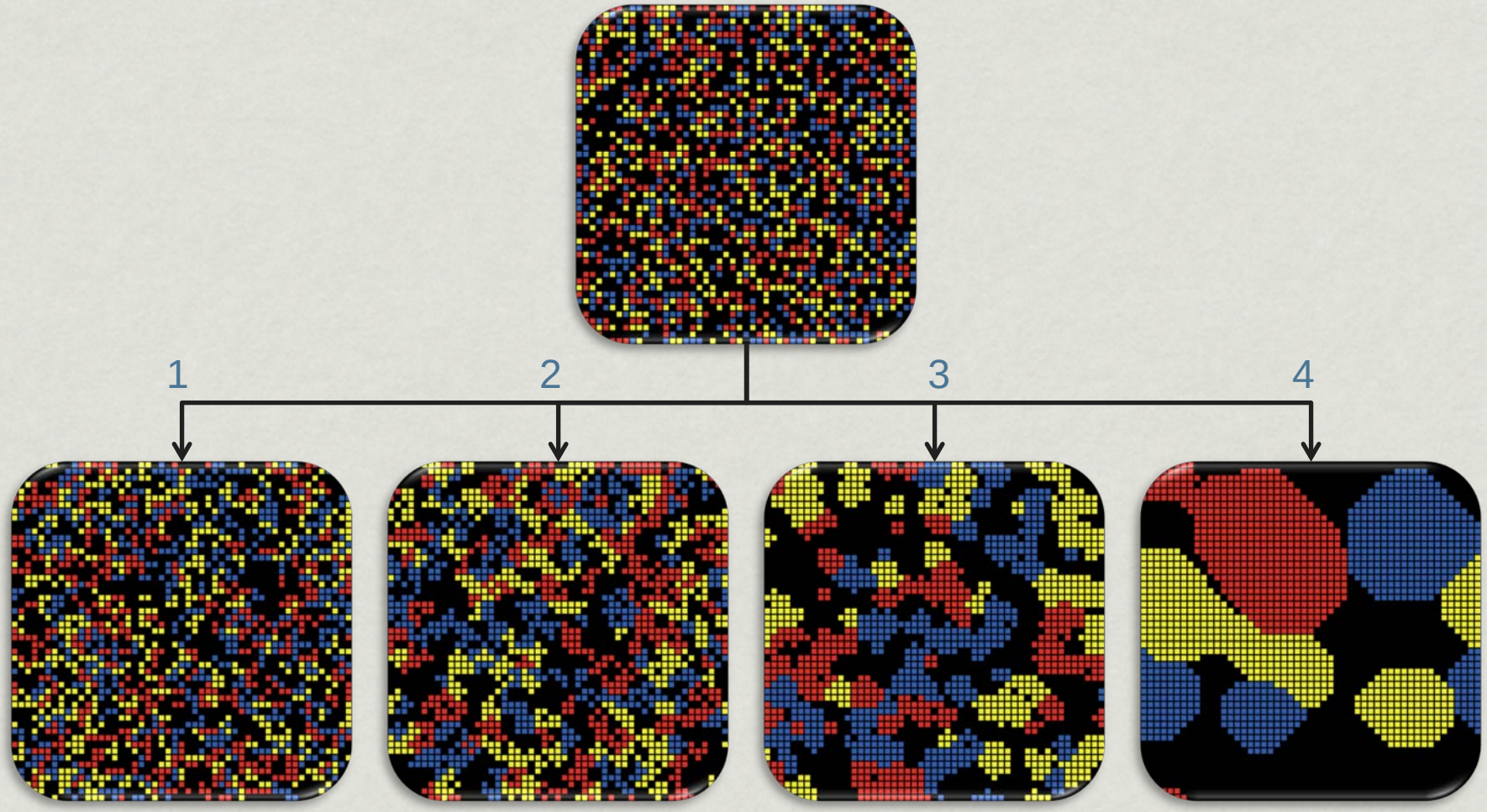
The Schelling Model

- In the late 1960s Thomas Schelling noted that US cities were racially segregated but individuals were not overtly racist
- Developed a simple model of *interacting*, *individual* choices
  - Agents were coins on a chessboard, each having eight neighbors
  - Agents had one simple rule:
    - If fewer than three of agent's neighbors were the same type move to a random empty space on the board
- What happened?

The Schelling Model



Generated Outcomes



ABM and Results

- Identification of outcomes and relationships that emerge from microlevel specifications of agent attributes and behavior
  - Allows for identification of sufficient conditions
  - Patterns of outcomes reveal alternative futures or possibilities resulting from a set of initial conditions
- Emphasis on generative or constructive knowledge
  - Simulation states are computed based on specified processes
  - Differentiates from outcomes that are logical or mathematical solutions but cannot be arrived at by interacting agents

Caveat Emptor!

- ABMs are incredibly flexible, but with a cost
  - “Curse of dimensionality”
    - Overly complex
    - Data overload
  - Computationally expensive
  - Scalability concerns
  - Fragile outcomes
  - Empirical grounding of attributes and behaviors
- ABMs are most effective when used in concert with other research and analytic methods
  - Allows for the generalization of patterns and behaviors characterized by individual case studies
  - Allows for the explorations of dynamics that may explain relations discovered by statistical analysis

What Makes for a Good ABM Project?

- Organization
- Theory
- Data
- Implementation
- Validation
- Experimental Design
- Customer Engagement

What is the Frontier?

- Exploitation of ABM in decision-making and analytic processes
  - Formalization of behavioral theories, counterfactual analysis, and scenario generation
  - Methods for *N of 1* problem
- Empirical specification of ABM populations
  - Microlevel collection of data
  - Microlevel specification of behavioral processes
  - Merging models with live datastreams
- Non-utilitarian logic
  - For example, deontic logic
- Technical implementation
  - Model instrumentation
  - Model parallelization and non-traditional processing
- Data storage, analysis and recovery

Organize for the Customer Needs

- ABMs are multidisciplinary endeavors, get the right group
  - Social theory, computer science, quantitative analysis, domain expertise, customer knowledge
  - Most productive members can do two roles
- Construct the team carefully and know its identity
  - Policy analysis with computers?
  - Software development that runs on social data?
- Ask the customer what the final product looks like
- Put authority in the people closest to the customer needs

Theory

- The most important choice of any model
  - All opportunities and limitations follow from the representation of the system
  - Representational choices have two paths
    - How experts think
    - Scientific literature and theory
  - Both paths have challenges
- Suitable theory should provide parameters that can be used to represent a variety of cases or conditions
- Suitable theory should have parameters that of interest to customers
- Understand the data before settling on a representation

Data

- ABM development should start with an examination of the available data
  - Consistent trap is the assumption that models can be populated and parameterized after construction
- Valuable data will exist in case studies and behavioral descriptions
- Quantitative data generally has three applications
  - Parameterization of agent population attributes and environmental variables
  - Specification of agent rules*
  - Target result of the simulation
- Data analysis and collection plans should co-develop with model
- Don’t confuse data and theory!

Implementation

- ABMs may be implemented in a nearly infinite number of ways
  - Choices navigate gaps between theoretical specification, technical performance
- ABM are complex systems
  - By design they do things that are unexpected
- Start with simple model(s)
  - Capitalize on mathematical models
  - Generate statistical patterns
  - Test ‘corner’ solutions
- Develop in modules and save intermediate versions
- Identify and test within technical trade-off space
  - Areas not specified by theory, but affect model behavior
- Social scientists, experts, and users must be involved and aware of implementation choices and decisions
- **The most valuable insights will be discovered here**

Validation

- The single most difficult issue facing modeling
- Validation is always contextual
  - Valid for *what* purpose?
  - Valid *compared* to what alternatives?
- Validation criteria carries powerful and subtle epistemological subtexts
  - Is the world predictable?
  - Do we possess agency or free-will?
  - How should our observations be coded and interpreted?
- Many models and questions are not subject to empirical validation
  - Decision-making over alternatives
  - Finding an optimal allocation of resources is not the same as identifying the consequences of different choices
- In this context, ABM is nothing new or mystical, no different than other models
- Model validation by tiers and docking

Experimental Design

- The model is two artifacts in one
  - A representation of a theorized world
  - A computational object with its own properties
- Understand the behavior of each
- Search across uncertainties
  - Unknown parameters estimates
  - Alternative model configurations
  - Competing theories
- Search for model sensitivities and expressiveness
  - Know what outcomes the model cannot achieve as a representation!
- Experimental agenda may reveal more about the structure of the problem than any particular model run or set of runs

Customer Engagement

- Know your customer
- At the start of any project, ask how a successful project ends
- Keep them engaged
  - ABM is learning by doing
  - Prevents doing “what you know”
  - Makes technical issues understandable
- Model building is expensive
  - Slower-better-more expensive
- Make this clear to sponsors
  - Only makes sense when the problem merits the resources


That's all Folks!