SCAMP's Stigmergic Model of Social Conflict

H. Van Dyke Parunak · Jason Greanya · Peggy McCarthy · Jonathan A. Morell · Srikanth Nadella · Laura Sappelsa

Abstract SCAMP (Social Causality using Agents with Multiple Perspectives) is one of four social simulators that generated socially realistic data for the Ground Truth program. Unlike the other three simulators, it is based on a computational principle, stigmergy, inspired by social insects. Using this approach, we modeled conflict in a nation-state inspired by the ongoing scenario in Syria. This paper summarizes stigmergy and describes the Conflict World we built in SCAMP.

Keywords Agent-Based Model · Social Simulation · Psychology · Causality · Stigmergy

1 Introduction

The DARPA Ground Truth program used computer simulations (produced by four simulation teams) to generate data from artificial societies. The two research teams used this data to exercise and test their methods for extracting the underlying causality of the society.

H. Van Dyke Parunak · Jason Greanya
Parallax Advanced Research, Beavercreek, OH 45431
E-mail: van.parunak, jason.greanya@parallaxresearch.org

Jonathan A. Morell
4.699... LLC, Ann Arbor, MI
E-mail: jmorell@jamorell.com

Srikanth Nadella
Kairos Research, Dayton, OH 45458
E-mail: srikanth.nadella@gmail.com

Laura Sappelsa · Peggy McCarthy
ANSER, Falls Church, VA 22041
E-mail: laura.sappelsa, margaret.mccarthy@anser.org

To avoid leakage of causal information from the simulation teams to the research teams, the two groups were unknown to one another, and all data transfer was by means of a test and evaluation (T&E) team [20]. SCAMP (Social Causality using Agents with Multiple Perspectives) is one of the four simulators. [1] "Multiple Perspectives" in the acronym reflects SCAMP's ability to simulate different dimensions of causality.

The scientific contribution of this paper, as of the other simulator descriptions in this special issue, is to document how the features of the simulation support the experiments in the overall program. Other papers [25, 27, 26] and an ODD protocol [14] for SCAMP [24] provide technical details on SCAMP itself, though Section 2 provides a brief overview of SCAMP's distinctive features, particularly in comparison with the other simulators in the program.

The program ran through three challenges, with successively more complex models in each challenge. Most of the data in this paper comes from the third challenge. Challenges 1 and 2 supported the event (Section 5) and goal (Section 6) perpectives, while Challenge 3 added the geospatial (Section 7) and social (Section 8) perspectives. In each challenge, the research teams faced three tests: *explain* the underlying ground truth on the basis of an initial data package containing simulation results over an extended period, *predict* how the scenario would evolve in the future based on this ground truth, and under different proposed changes, and *prescribe* changes to achieve certain objectives.

Section 2 outlines SCAMP's distinctive architecture, introducing stigmergy and briefly comparing our approach with other simulation technologies in general

---

and the other simulators in Ground Truth in particular. Section 3 outlines the scenario in which we embedded our simulation. Section 4 describes the different groups involved in the conflict, and provides more detail on architectural features that support group distinctions among agents. The next four sections describe the four perspectives that SCAMP currently supports: the types of events in which these agents can participate (Section 5), the goals that each group pursues (Section 6), spatial constraints (Section 7), and social dynamics (Section 8). Each of these four sections begins with the mechanics of the perspective and the tools used to model it, then presents some sample data to show how that perspective contributes to SCAMP's social and psychological realism. Section 9 discusses the Predict and Prescribe tests that we gave the research teams, and Section 10 concludes.

2 SCAMP and Other Social Simulators

In this section, we briefly review the very large space of technologies used for social simulation. Then we describe the distinctive features of SCAMP, and compare it with the other simulators used in the Ground Truth program.

A wide range of technologies have been used for social simulation [6, 19, 9, 12]. The major distinction among them is between equation-based approaches such as system dynamics [43] and its qualitative foundation of causal loop diagrams, and agent-based approaches. Equation-based simulations track the evolution of measurable variables during a system's operation, typically using differential equations, and their structure makes it convenient for them to focus on population averages over the variables describing individual actors, using what a physicist would term a mean-field approach [23]. Agent-based approaches naturally represent each actor's variables separately, and if the dynamics of the simulation are complex, the values of these variables can diverge from one another, yielding an overall result that is qualitatively different from an equation-based model of the same scenario [35]. All four simulators used in Ground Truth were agent-based. While equation-based simulations run faster than agent-based ones, advances in computational technology have led to increasing use of agent-based models, and several frameworks support their development, including Repast [1], MASON [18], and NetLogo [49].

The fundamental property of an agent is that it continuously monitors its inputs, reasons about them, and takes action. Various agent-based simulators differ in the kind of internal logic used by the agents in reasoning from inputs to outputs. Common approaches include architectures that explicitly model the beliefs, desires, and intentions of individual agents (BDI agents, [38]), Bayesian reasoning (often applied in a BDI framework [5]), and raw conditional logic driven by numerical state variables.

SCAMP is based on an agent architecture [21, 30, 31, 34] known as "stigmergy." Grassé [13] coined this word in 1959 from the Greek στίγμα (sign) and ἔργον (action) to describe insect actions that are mediated by signs in the environment, such as pheromones, rather than by direct inter-agent messages.

Fig. 1 shows schematically the interaction between agent and environment. An agent's state and local environment determine its actions, and are modified by its actions. The environment modifies its own state, and agents interact by sensing signs left by other agents.

Because of our ignorance of linguistic capabilities in non-human organisms, stigmergy is widely used to model animal communities [4, 2, 45]. Agents are localized in a graph-structured environment (for ecological studies, typically a geospatial lattice). Each time step, they "deposit pheromone" by augmenting some variable associated with their current location, and sense the value of this variable in adjacent locations to guide movement decisions.

The relevance of stigmergy to human behavior draws on an observation by Herbert Simon [41] that is known as "Simon's Law." He writes: "An ant, viewed as a behaving system, is quite simple. The apparent complexity of its behavior over time is largely a reflection of the complexity of the environment in which it finds itself." Then he extends this insight to human behavior: "Human beings, viewed as behaving systems, are quite simple. The apparent complexity of our behavior over time is largely a reflection of the complexity of the environment in which we find ourselves." The trick is to encode the agents' complex behavior, not in the agent code, but in the environment in which the agents live. While conventional cognitive agents put the domain model inside the agents, SCAMP puts the agents inside the model. That is, the agents move over external graphs reflecting the choices available to them, including both spatial lattices (Section 7)
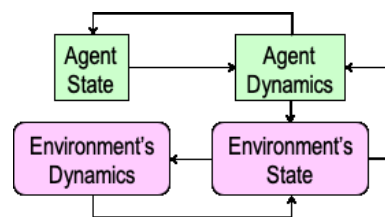

Fig. 1 Basic stigmergic schema

and directed graphs (Section 5)). Different kinds of agents represent different geopolitically active groups to which they belong (Section 3). Each node in these graphs has a series of presence variables for each kind of agent. These presence variables support a digital imitation of insect pheromones. Agents augment the presence variables of their own kind for the nodes they visit, and sense the values of all presence variables in adjacent nodes when considering a move. These graphs and the variables that they support constitute the environment invoked by Simon's law, and determine the behavior of the agents.

In a purely stigmergic system, agents interact only with their local environment. SCAMP also allows agents to exchange information directly with agents whom they meet on a node of the environment (Section 8, but this exchange is highly stereotyped. Thus, unlike the other Ground Truth simulators [37,50,36], SCAMP does not support remote messages between agents. However, it does support two forms of information movement among nodes of the environment that can allow one agent to react to the actions of other agents that are remote from it: influence edges in the Causal Event Graph (Section 5), and changes in the urgency of one event based on other remote events by way of the goal system (Section 6).

A SCAMP agent repeatedly chooses among accessible alternatives. The basic alternatives are event types in a directed graph (Section 5) and locations in a geospatial lattice (Section 7). An agent is always participating in one event type, and considering the next. It always has a current location, and some event types require it to move in geospace toward a goal specified in the definition of the event type.

Every alternative open to an agent (an event type, or a geospatial location) is a node in a graph, and is characterized by a vector of scalar features (Section 4). These include the presence variables mentioned above. Each agent has a vector of scalar preferences over this same feature space, based on the groups with which it is affiliated. To choose among alternatives, it computes the cosine (normalized dot product) between its preferences and the features of each alternative, exponentiates each value (to make them positive), and normalizes the set to form a roulette wheel that it spins.

Each entity in the model is represented by a polyagent [29], a single *avatar* that continuously deploys a swarm of *ghosts* to explore the future. The ghosts execute the selection logic described above. As they move, they deposit pheromone by ~~augmenting~~ modify the presence features of the event types or geospatial locations they visit (the arrow in Fig. 1 from *Agent Dynamics* to *Environment's State*.

In each decision cycle, the avatar sends out a limited number of *waves* of ghosts, each consisting of a limited number of ghosts. Each ghost carries its avatar's preference vector. Each ghost explores the future to a limited horizon. Then the avatar follows the crest of the presence pheromones deposited by its ghosts. In psychological terms, the ghosts implement the common human practice of planning by mental simulation [16]. Each ghost represents an actor's mental projection of one possible future, and repeated waves of ghosts reflect an actor's repeated mental simulation informed by the results of a previous cycle ("Hmm, that didn't turn out well. Let's think this through again ...."). The limits on number of waves, number of ghosts, and exploration horizon are important bounds [42] on SCAMP's rationality.

As a ghost from one group runs into the future, it may encounter features modified by ghosts from other groups, also in the future. If its preferences are sensitive to those features, it will respond to them, with a degree of attention modulated by its preference vector (Section 4). At first glance, the notion that the ghosts belonging to one avatar deposit presence features that can be sensed by other avatars seems to violate the concept of ghosts as an individual avatar's mental simulation of possible futures.[2] In fact, this facility gives our avatars a primitive theory of mind that allows them to reason recursively about the reasoning of other groups. A ghost cannot see other ghosts, only the local value of their pheromones (their presence features). This pheromone field corresponds to a probabilistic estimate of where those ghosts actually were and where their avatar may move. While our implementation generates this estimate based on the actions of other agents, the nature of the information available to an avatar's ghosts is consistent with that produced by a local theory of mind informed by observations of other agents.

This decision process, based on the mathematical manipulation of the features of decision alternatives and the preferences of each agent, is mechanical. Stigmergic behavior by people is well documented [22], but stigmergy might seem incapable of replicating the kind of higher-level biases that dominate human thought. In fact, the processing provided by each of SCAMP's perspectives opens the door to behaviors that can be aligned with many recognized cognitive processes and biases [25].

While not commonly used for social modeling, stigmergy offers a number of benefits over more conventional architectures.

---

[2] We are grateful to a referee for requesting this clarification

- *Speed of processing:* The computations involved (dot products and roulette selection) are extremely efficient on modern processors.
- *Extensibility:* Additional perspectives can easily be added to the model, as long as they meet one of three conditions. 1) They can be traversable graphs, that is, graphs over which it makes sense for agents to move by comparing agent preferences with node features (like our current event graph (Section 5) and geospatial (Section 7) perspectives). 2) They can take as inputs the features of nodes on an existing traversable graph and deliver their output as changes to such features, as does our current goals perspective (Section 6). 3) They can modulate the preferences of agents, as does our current social perspective (Section 8).
- *Modeling alternative futures concurrently:* The multiple ghosts sent out by an avatar explore alternative possible futures for the avatar in a single run, and the pheromone field that ghosts develop over a traversable graph is, up to a normalizing constant, isomorphic to a probability field giving the probability that the corresponding avatar will be at each location at a given time. Currently, we use this field only to guide the avatar, but it could be analyzed to provide information on the likelihood of different possible outcomes, without the need for costly repetitive simulations.
- *Model accessibility:* The logic of SCAMP's mode is external to the individual agents, encoded not in computer code, but in artifacts accessible to analysts who are not programmers. Modelers construct these artifacts using CMapTools [15] for network structures, GIMP [46] for geospatial information, and Excel for detailed parameters. This inversion of the relation between agent and domain model makes the behavior of SCAMP much more accessible to domain experts who are not programmers than traditional approaches.

One disadvantage of a purely stigmergic model is that agents cannot exchange messages directly with one another. SCAMP relaxes this constraint in the social perspective (Section 8): if two agents participate in the same event type or visit the same geospatial tile concurrently, they exchange their preference vectors and use them (according to individual parameters) to modify their own preferences, modeling the influence of our associates on our own attitudes. Though highly stereotyped, this extension demonstrates that a stigmergic code base can be extended to provide more conventional features, such as a rich interagent language.

It is illuminating to compare SCAMP with the other three social simulations in the Ground Truth program,

each using a different agent-based social modeling technology. Each of these systems is discussed in further detail elsewhere in this special issue. We highlight in this comparison the ability of non-programmers to create and modify the models [27].

George Mason University, Tulane University, and the University of Buffalo produced a model of Urban Life [50] in the MASON modeling toolkit [18] and its GeoMASON extension [44]. Numerous aspects of the model are generated algorithmically, including the agent population, the geospatial map, and the social network among the agents. The system provides a predefined set of triggers (sensitive to both internal and external factors), behaviors to which they lead, actions that make up behaviors, and goals that determine when actions stop. Defining new triggers, behaviors, actions, and goals requires programming, but a drag-and-drop interface allows modelers to assemble and parameterize these components to define a scenario.

Raytheon BBN produced ACCESS [37] in the Repast framework. The model highlights the interactions among individual agents, groups to which they belong, and the overall population or "world." ACCESS models space as a list of locations, but without orientation or distances, so there is no "map" for a user to enter, and the individual behaviors are determined by equations embedded in the code.

USC ISI produced a disaster world [36] in their PsychSim social simulation framework. Agents are driven by partially observable Markov decision processes (POMDPs) and can reason recursively about one another. PsychSim provides one interface that allows social scientists to create simulation models directly, and another allowing them to manipulate the parameters governing the simulation. However, both interfaces abstract over the full complexity of a PsychSim model (e.g., limiting the types of probability distributions and reward functions), so specifying arbitrary probability and utility models requires sufficient programming ability to use the PsychSim API.

GMU and ISI both support non-programmers who wish to modify a scenario, but still require programmers to modify details. In addition, they introduce proprietary interfaces. SCAMP allows non-programmers to define new groups, actions, and goals and their relations, using tools with which they may already be familiar.

3 Scenario

The model we built in SCAMP reflects multipartisan conflict in a country inspired by (but not a replica

of) Syria, with an authoritarian government seeking to cling to power in the face of a democratic armed opposition and an ideological extremist force seeking to establish a religious state. Relief agencies seek to relieve the condition of the general population, who just want to get on with their lives. Section 4 defines these groups in more detail. Over the course of the program we modeled two imaginary countries, Donglap and Tharum.

A particular challenge in the Ground Truth program was enabling the research teams to interact with the simulation teams as though they were interacting with a real social scenario, rather than with computer scientists seated at terminals. To help us structure a coherent interaction with the research teams, the SCAMP team adopted the persona of the Fourth Marine Expeditionary Battalion (MEB), tasked with a variety of oversight and stabilization assignments in Donglap and later in Tharum. We cast the research teams as advisors to our civil affairs officer. We sought to stay in character with this persona in all interactions with the research teams. In particular, our tasking memos were delivered as though they were prepared on a manual typewriter (we were, after all, deployed to remote locations), subject to appropriate military approvals (Fig. 2). Limitations to research requests were explained in terms of realistic operational constraints on a military unit. Predictions envisioned events that the Fourth MEB suspects might occur or actions that it might reasonably take in line with its mission and general military policy such as rules of engagement and status of forces agreements, and prescriptions were motivated as guidance to the Fourth MEB for future actions.

Here is the tasking memo for the Explain test in Challenges 1 and 2 (which used the same ground truth). We comment on the rationale for the various details and the constraints they pose on our model.

> You are advising the civil affairs officer attached to the 4rd Marine Expeditionary Brigade, recently deployed to Donglap, a new country that

Fig. 2 Example Tasking Letter

formed after the dissolution of an unstable third world country.

This role is a realistic one in view of DARPA's mission, and motivates the interaction of the research teams with us. However, it does make them an extended part of our organization. To avoid any questions about whether their interaction with us might influence the scenario, we did not model the Fourth MEB as a causal player in our ground truth. However, some of the changes that we asked the research teams to explore in the Predict and Prescribe tests were actions that the Fourth MEB could take to modify the ground truth (for example, refuse access to a geospatial region, suppress events of a given type, or modify the number of actors associated with a given group, whether through diplomatic channels or direct military action).

> Donglap is in the midst of a civil conflict not entirely unlike the current Syrian situation. Our CO has tasked you with figuring out how the society works, so that US operators can win the support of the local population rather than alienating them through unintentional gaffs. The locals are understandably nervous about talking with outsiders, but we have found one actor whom you can interview, and from whom you can learn more. We will identify actors with labels like A8263. In this culture, extended interrogations are considered rude, and any one actor will tend to stop cooperating if you query it too much.

SCAMP runs offline, recording its results to logs from which we respond to queries. Early in the program, we were concerned that the research teams might request all our data at the outset, leaving us with no more to give them. A shy populace and cultural taboos against extended interrogations provided a valve to control the flow of information. The lack of data turned out to be a problem, and in later challenges we generated longer logs and were more generous with our data.

> Actors participate in different events over time. We will provide you with succinct descriptions of some of them, but most you can recognize only as distinct events, identified with labels like E32. There are two reasons for this ambiguity: your unfamiliarity with the culture, and the reluctance of the actors to talk to strangers about some of their activities. (This is, after all, a time of war, and one sometimes doesn't know the loyalty of the person to whom one is speaking.) However, over time, as you learn more about the culture, the meanings of some of these events

may also be revealed. We will also provide you with a history of environmental events, that is, events caused by forces other than the actors we simulate.

Our events have very full descriptions (Section 5), and DARPA was concerned that the research teams would become so preoccupied with the causal implications of the descriptions themselves that they would not focus on extracting causality from the behavioral data we produced. This part of the scenario justifies the obfuscation we applied to event names to avoid this problem.

Our persona also supported contextualization of experiments requested by the research teams. Direct experiments with subsets of the population (e.g., focus groups) seemed out of keeping with our persona, but we adopted the following mechanism:

> Unfortunately, there are many conflicts going on in the world today. Many involve similar cultures to the one we are presenting you (so that the causal structure remains the same, though details may differ). As you discover our ground truth, you can request data from another conflict that differs from our base conflict in specified elements of the ground truth. To make such a request, you must learn enough about our ground truth to specify what elements you want changed. It is likely that the other conflict will also differ in other ways beyond your control and knowledge, but it will have the same causal structure as the central conflict you have been studying.

This rationale is not only supports experimentation, but also allows us an excuse for rejecting experimental requests that come too close to asking us to solve Predict and Prescribe tests that we posed to the research teams. We simply replied, "We can't find such a country." Sometimes we justified this with the additional note, "If we were able to find such an example, we wouldn't need to ask you for the prediction" (respectively, prescription).

In Challenge 3, we added two new perspectives to our model: geospace (Section 7), and social dynamics (Section 8). These changes significantly enlarged the underlying causal structure. To help the research teams avoid confounding the new structure with the old one, we redeployed the Fourth MEB from Donglap to Tharum. Here is the tasking memo for Challenge 3.

> The 4th Marine Expeditionary Brigade has been redeployed to Tharum. Like Donglap, Tharum is suffering from internal unrest that is stimulating refugee activity, but it differs from Donglap in three main ways.

> 1. We can now provide you with a geospatial map of the terrain as well as the current location of actors over time. Actually, we provide two maps: elevation data (light is higher), and one with features such as roads, water, locations of cities, and boundaries of countries. Tharum's culture is rooted in ancient Babylonian society, whose mathematics were hexadecimal, and so location is measured on a hex grid, with (0,0) at the upper left corner and (23, 20) near the lower right (the first number indicates horizontal position; the second, vertical). In reporting the history of actors, we will give you not only the sequence of events in which they participate and their satisfaction level, but also their geospatial coordinates. Tharum is small enough that environmental events ("nature") apply to the whole country, and we do not give coordinates for them.

> 2. Most of the event types that we experienced in Donglap also occur in Tharum, but we have identified about 200 more, and there is no guarantee that the relation of the ones that you've seen before to one another is the same as it was in Donglap.

> 3. People in Tharum are more open about their associations than they were in Donglap, and we are now able at each point in time to tell you, for each actor, what other actors it has encountered in its history up to this point, and how close that relationship is.

For computational reasons, we use a hex grid for geospace, and the first item creates a backstory for this unusual reference system, as well as introducing the new geospatial data that is available. The third item supports the new social perspective.

> As before, we are trying to collect data on conflicts elsewhere that share the underlying causal structure of Donglap, and on request, will try to find data for you from other conflicts that differs from our base conflict in specified elements of the ground truth. To make such a request, you must specify what elements you want changed. You already know that it is sometimes possible to find situations that do not include certain event types, or in which the relative populations of different types of actors differ. Some cases in our archive also include situations in which a specific event becomes inaccessible partway through the data. It is likely that the other conflict will also differ in other ways beyond your control and

knowledge, but except for the changes you request, it will have the same causal structure as the conflict in Tharum.

In Section 9, we will show how we extend this persona to posing Predict and Prescribe problems for the research teams.

4 Groups and Feature Space

Feature space is the vector space in which event features and agent preferences are defined. The structure of this space depends on the distinct groups of agents in a model. Every agent is affiliated with one or more groups. Conflict World has six groups representing different agents in the conflict:

1. The *government* (GO)is authoritarian, bent on retaining its own control of the situation, and willing to oppress its people to keep them in line.
2. The *military* (MIL) is initially aligned with the government, but can diverge. (Challenges 1 and 2 did not have a separate MIL group.)
3. The *armed opposition* (AO), inspired by the Syrian opposition, is a movement from within the country that seeks to reform or replace the government with democratic institutions. Our mission as the Fourth MEB, reflected in our Predict and Prescribe tests, is to promote this change.
4. The *violent extremists* (VE), inspired by ISIS, are an ideologically driven foreign faction that seeks to include the local territory (whether Donglap in Challenges 1 and 2 or Tharum in Challenge 3) in a larger religious state.
5. *Relief agencies* (RA) seek to provide humanitarian relief for civilians, largely in the form of refugee camps both within and just outside of Tharum.
6. *People* (PEO) are pro-opposition, anti-government civilians, just trying to get on with their lives.

Each group has a hierarchical goal network (HGN) described in Section 5, and a vector of preferences in [-1, 1] over the same feature space that encodes the features of alternatives that agents may consider. The dimensions of feature space are of three types: exogenous, urgency, and presence.

– *Exogenous* features (each in [-1, 1]) are defined by the modeler, and in Conflict World are three in number: economic, physical, and psychological well-being. As event features, these describe the impact of the event type on the agent, while as preferences, they describe the agent's priorities. A positive feature means that the event type improves that facet of agents who participate in it (for example, an event type "go to work for the day" would have positive impact on a participating agent's economic well-being, reflecting the expectation of income), while a negative feature indicates that an event type will reduce an agent's well-being ("participate in street riot" would have a negative physical feature, reflecting the risk of physical harm). These features correspond to positions in each agent's preference vector. For instance, a preference of 0.1 for physical well-being means that agents tend not to care much about their health and safety, while a preference of 0.9 means that agents are very much concerned about this feature. While the decision mechanisms can support masochistic agents (with well-being preferences < 0), we restricted preferences for exogenous features to values > 0 for the Ground Truth program.

– Each group has one *urgency feature* in feature space, reflecting the current state of that group's HGN. One group's preference for another group's urgency feature reflects its desire to promote (preference > 0) or block (preference < 0) the other group's goals.

– Each group has one *presence feature*. In imitation of insect pheromones, a group's feature on an event type or location is augmented each time a ghost representing the group visits that alternative (the arrow from *Agent Dynamics* to *Environment's State* in Fig. 1), and evaporates over time (*Environment's Dynamics* in Fig. 1). One group's preference for another group's presence feature reflects its attraction to or repulsion from members of that group. Evaporation of the presence feature imposes a temporal limitation on knowledge of past participation of agents in events or locations. It is an important example of the bounded rationality [42] of SCAMP agents, and an instance of recency bias.

As noted in Section 2, a ghost responds to the presence of ghosts from other groups if its preference for the presence features of those groups is non-zero. Since ghosts are exploring possible futures for their respective avatars, the strength of the presence features for a given group reflects the probability that an avatar of that group will be at that location at that time, and when ghosts attend to the presence features of other ghosts, they are behaving recursively [47], reasoning about the futures being considered by the avatars represented by those ghosts. As in other recursive agent formalisms, the recursion is nested: agent A reasons about what agent B thinks about what agent A thinks about what .... The depth of the recursion is the number of waves of ghosts that an avatar sends out, and the impact of the recursion depends on the ghost's preference for the

Table 1 Key group parameters in Conflict World

Group	Population	Preference Variation	Affiliation Threshold
Government	13	0.2	0.98
Military	6	0.2	0.98
Armed Opposition	11	0.3	0.98
Violent Extremists	8	0.3	0.98
Relief Agencies	4	0.3	0.90
People	18	0.4	0.80

presence feature of the other group. Both of these parameters depend on the ghost's group.

The preference vector for each group is a baseline from which the preferences of individual agents are generated. Each group is assigned a number of initial agents and a variation v in $[0, 1]$. Each preference for a new agent is selected uniformly from the baseline value $\pm v$.

Once an agent has preferences, it considers whether to affiliate with other groups, based on the cosine of the angle in feature space between its preference vector and the baseline preference vectors of the other groups. Each group has a threshold that the cosine must exceed for affiliation to take place, and the cosine defines the weight of the affiliation. The preference vector that the agent uses in choosing alternatives is the weighted average of its own preference vector and those of other groups with which it is affiliated, and its ghosts augment presence features of all affiliated groups, again proportional to the affiliation weights. In both cases, the weight for its home group is 1, and the weights for the other groups are its cosine proximity to them. Its preferences for exogenous features vary with its experience, while its preferences for urgency and presence features are fixed unless the social perspective (Section 8) is active.

Table 1 shows each group's initial populations, preference variations, and affiliation thresholds. Note the relative homogeneity of the Government and Military (low variation) compared with other groups, and the willingness of People to affiliate with other groups.

At first glance, the relatively small initial population (60 agents) may seem unrealistic. However, each avatar decision comes from a swarm of 48 ghost agents, so the actual exploration of alternatives involves nearly 3000 agents. In addition, the social dynamics perspective (Section 8) allows new agents to join the simulation as it runs (for example, by influx of foreign fighters). In

our baseline run for Tharum, the simulation ended with 548 avatars, represented by more than 26k ghosts.

Each group also specifies one or more geospatial regions over which its agents are initially distributed.

In some configurations, we add Neutral agents, who have no home group and no baseline preferences. Their preferences are assigned randomly, and then with threshold 0 they affiliate with the closest group. Neutral agents are not necessary if we can estimate in advance the proportion of actors in each defined group, but they allow us to model a population with a large proportion of actors of unknown group affiliation. In the Ground Truth program, they also provide a lever to adjust the complexity of the simulation.

SCAMP also supports an Environment group whose single agent moves over a subgraph of background events to generate natural events (droughts, famine), or events who detailed causality is not modeled in the system (economic downturn, assistance from foreign governments). The Fourth Marine Expeditionary Battalion is not a group in the Conflict World, and participates in the causal dynamics only by modifying the ground truth for Predict or Prescribe tests.

Groups are defined in a worksheet in the model's Excel workbook.

5 Event Types

The fundamental dynamic of SCAMP is agent choice over alternatives, and the main alternatives are recorded in a Causal Event Graph (CEG), a directed graph whose nodes are event types. Every trajectory through this graph is a coherent narrative for an agent's experience.

5.1 Event Types and the Causal Event Graph

There are over 400 event types in Tharum. These include:

- large numbers of people move to urban areas
- public demands democratic reforms
- Government security forces arrest minority leader
- military refuses to carry out government's orders
- govt & opposition leaders commence official talks
- Neighbors leave
- people arm themselves
- relief agencies identify an increase in unplanned need
- Military bombs opposition-controlled neighborhoods
- funders of relief agencies lose interest in conflict

- Protesters share political news on social media
- QOL at IDP camp improves
- government & armed opposition forces cease negotiations
- head of state/government calls for end to violence
- transitional government invites election monitors

We call these "event types" because they can recur over the course of a run. Strictly speaking, an "event" is a consecutive period of time during which there are agents participating in a given event type. However, where there is no risk of confusion, we often refer to the nodes in the CEG as "events."

Each event type has a *feature vector* over the same feature space that defines group parameters. The exogenous features are coded by the modelers, while the urgency features are modulated by the current state of the world via the hierarchical goal networks (Section 5) and the presence features are augmented each time an agent participates in an event, and evaporate exponentially through time. For example, consider a stationary agent that augments its group's presence feature at its current location by d each time step, while the feature evaporates each step by the factor $e \in (0, 1)$. The most recent deposit contributes d, the one from the previous time step de, the one from two time steps back de^2, and so forth. Thus the total presence feature after n steps is $d + de + de^2 + ... + de^{n-1} = d \sum_{i=0}^{n-1} e^i$, which is just the geometric series, and asymptotes to $d/(1 - e)$.

Not every event type makes sense for all groups. For example, only Government agents can meaningfully participate in "Government security forces arrest minority leader." Each event type is scored to indicate which groups have *agency* for it. Some of the directed edges through this graph are *agency edges*, connecting event types for which the same group has agency, and indicating that an agent currently participating in the origin of the edge may choose to traverse the edge to reach its next activity.

The nodes and agency edges together form a narrative space [39]. Any single trajectory through this space is a plausible narrative for agents that follow it. For computational convenience, the narrative space begins with the node $START$ and ends with the node $STOP$, so that every valid narrative is a path from $START$ to $STOP$.[3] To generate trajectories longer than those encoded in the CEG, we send agents that reach STOP back to START. Since their own preferences and the features of the events have changed, they may very well follow a different trajectory on each transit of the CEG.

For example, here is the trajectory of an agent belonging to the People group:

- people desire a government that recognizes civil liberties and legitimizes all sectors of the population
- civilians are dissatisfied with the authoritarian govt
- public demands democratic reforms
- protesters attack government facilities & destroy assets
- government and protesters clash
- Protesters attack police station killing officer and burning it down
- large numbers of people protest throughout the country

Each agent maintains a memory of its entire life history, and each time an agent moves, it logs its movement for later analysis.

An important part of intelligence analysis (or scenario modeling) is anticipating possible patterns of behavior that the actors of interest might exhibit. Analysts commonly describe scenarios in terms of possible narratives, making the CEG a natural representation for capturing complex social situations, and the CEG was originally developed in support of intelligence analysis. One of the benefits of this representation is that it amplifies the creativity of analysts by combining narratives that they explicitly formulate to yield a huge number of other narratives that are consistent with these. Ground Truth is not concerned with analytic creativity, but this same amplification means that the CEG can generate an incredibly large number of different behavioral trajectories as data for the research teams.

A simple example illustrates this amplification. An analyst might consider possible narratives A → B → C and D → E → F, offering agents a total of two possible histories. But if the analyst decides that B could also lead to F and E to C, the number of possible trajectories doubles, without defining any additional events.

The amount of combinatorial amplification of the narratives explicitly defined by the analyst depends on the length of the analyst's individual narratives and the number of interconnections among them, but we can get an idea of the possibilities. Because of the START and STOP nodes, the CEG is an irregular directed lattice.[4] To calibrate our intuitions, consider the paths between diagonally opposite corners in a square directed lattice of side n. Such a lattice contains $(1+n)^2$ nodes. Except for the START and STOP nodes, which number only 2 for any sized lattice, this is the number of event types

---

[3] In the CEG for Ground Truth, each group has its own STOP node, but to simplify analysis we envision an additional STOP node with an incoming edge leading from each of these.

[4] In fact, SCAMP does allow causal cycles, so that the CEG no longer defines a partial order over events, but for the purpose of this illustrative computation we exclude such cycles. They only add to the space of generated narratives.
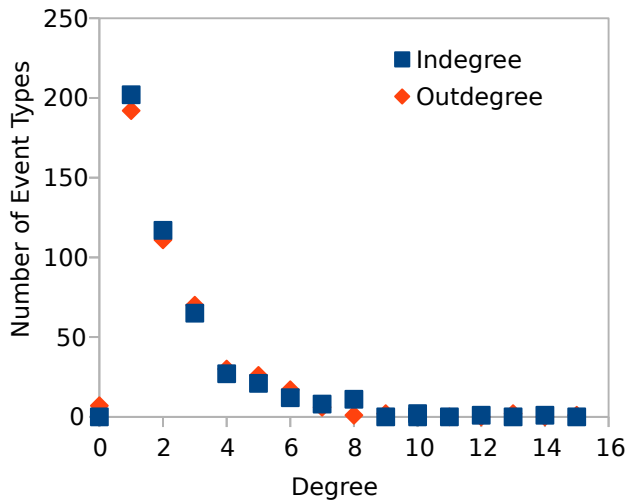
Fig. 3 Degree Distribution in Challenge 3 CEG

that the analyst must define. The average node degree in a square lattice asymptotically approaches 4, and by symmetry in-degree = out-degree = 2.

A simple counting argument [8] shows that the number of simple paths in such a structure is defined by the central binomial coefficients, $_{2n}C_n$. A lattice of 441 nodes (n = 20) thus generates more than 1.3E11 possible trajectories, each of length 2n = 40. The analyst needs to conceptualize only enough narratives to generate the desired number of event types, with enough overlap to link them into a lattice. For a square lattice with indegree = outdegree = 2, each event type needs to appear on average in two narratives. Thus 22 narratives of length 40 (2 ∗ 441/40) covering 441 distinct event types suffice to yield a 441 node lattice, far fewer than the 1.3E11 trajectories such a lattice contains.

Our Challenge 3 model has 467 event nodes (including START and STOP) and an average degree of 4.7. The event indegree and outdegree distributions are highly skewed (Fig. 3), and nearly identical. The events with no outgoing edges are STOP nodes for each group, and one event (START) has no incoming edges. This skewing will reduce the generative power of the CEG, but even so the number of possible paths greatly exceeds those that the analyst constructing the CEG can explicitly consider. SCAMP's swarming ghosts develop a probability field over this massive space, sampling it for data generation (in Ground Truth) or for intelligence analysis.

Most events are restricted to agents of one or a few specific groups. Events record which groups have *agency* for them. Thus the narrative space is partitioned into smaller subgraphs for each group. However, agents can move from one subgraph to another, if they are affiliated with both groups.

Fig. 4 shows the CEG for the Conflict World. Colors reflect the agency of the various events. This CEG generates a huge number of alternative narratives, giving a very rich *event space* within which agents move.

The START node for all groups is at the upper left of Fig. 4, and has edges to all of the group subgraphs. When an agent reaches the END node, or some other node that (because of *prevent* influence edges) offers no next choices, it returns to START, but in most cases will not retrace its previous path, since its own state and the state of the event nodes will have changed. Each time the overall participation of an event type node goes from 0 to non-zero, a new instance of that event type has begun, and ends when the node's participation drops again to zero.

Most events unfold over a period of time, and an agent's participation in an event will generally take some time before it moves to the next. However, the delay imposed by an event on different agents will not necessarily be the same. To model this effect, we assign each event a nominal *transit time* describing how long agents participate in it before moving on. In the absence of more detailed knowledge, we assume that events are Poisson distributed, which means that their inter-arrival time follows an exponential distribution. So each agent samples its individual transit time from an exponential distribution with the event's transit time as the parameter. Each time an agent completes an event, it increments its local time by its transit time, and does not execute again until all agents with lower agent times have executed. In our current model, agents do not take the transit time of an event into account in deciding whether to participate in it, a form of the duration neglect bias.

Event participation has other effects on agents in addition to advancing their clocks.

Fig. 4 Causal Event Graph for Tharum Conflict World

- If the event involves geospatial movement (Section 7), the agent's participation moves it spatially.
- The event's exogenous features modify the agent's overall physical, emotional, and economic wellbeing, which in turn modifies the agent's preferences for those features, a form of learning. For example, an agent that participates in many events with high positive economic wellbeing features will tend to reduce its preference for that feature.
- SCAMP's social dynamics (Section 8) modulate its preferences based on the other agents whom it meets, allowing a form of conformity bias.

In addition to agency edges, the CEG also has *influence edges*. Agents do not move over influence edges. Influence edges can connect events in different subgraphs, and modulate the availability of the destination event and the probability of its selection to agents considering it, based on the degree of participation (the total presence features) on the source event. Influence edges, along with HGNs and geospace (Sections 6, 7), allow different groups to interact in SCAMP.

Modelers construct events and edges among them in CMapTools [15], and record event parameters in a sheet of the Excel workbook.

5.2 Data from the CEG

We originally planned to disclose full event names to the research teams, but DARPA was concerned that these teams might focus their attention on the semantics of the event names, and miss the true causal influences, such as the interaction of features and preferences and the role of influence edges. In addition, many full event names disclose elements of our ground truth, such as the names of groups, the groups having agency for an event, or critical locations. So we developed two levels of reducing the information in the event identifiers, a process we call "cheshiring" (inspired by the successive disappearance of details of the Cheshire Cat in *Alice in Wonderland*). The level with the least semantics is a simple event number, e.g., E24, for each event. The intermediate level is a short event name that obscures the identity of groups, locations, and key resources. For example, "Government security forces arrest minority leader" becomes "Activity 85 takes place," "government & armed opposition forces cease negotiations" becomes "negotiations cease," and "QOL [Quality of Life] at IDP [Internally Displaced Person] camp improves" becomes "QOL at Location 21 improves." (Our persona allowed us to motivate the imprecision in cheshired event names on the grounds that the non-English languages spoken in Donglap and Tharum were unfamiliar and un-

common, and that the Defense Foreign Language Institute was slow in responding to our request for a linguistic consultant.) In Challenge 1, we disclosed only event numbers and gave the research teams access to an event's short name once they had seen information about all of its neighbors in the CEG. By Challenge 3, we released the short event names initially, but with a cautionary note not to rely on their (very sparse) semantic contents.

The main evidence about events available to the research teams is an agent history reporting, for each agent, the event in which it is participating at successive times. For example, Table 2 shows the trajectory for an unaffiliated military agent, that is, the history of events in which it participates. Each agent maintains its history internally, but after each agent movement, we also log the move in a file for subsequent analysis. This trajectory is meaningful with full event names, but with only the short names as requested by DARPA, the trajectory must be combined with other information to discern the underlying causality.

Even though different agents in a group can affiliate differently with other groups, the similarity among them can still be detected by comparing the similarity among their trajectories. One way to do this is with the "string edit distance," which is the minimum number of changes (additions, deletions, or replacements of one element by another) needed to change one string of characters (or list of event types) into the other. This measure, the "Levenshtein distance" [17], can be computed in time proportional to the product of the lengths of the two words by an iterative algorithm due to Wagner and Fischer [48].

Fig 5 shows the result of plotting such similarities from a run in Challenge 1, using multidimensional scaling (MDS). We use Kruskal's nonmetric version [7], as implemented in R's isoMDS function.

The plot clearly reflects the impact of the per-group parameters that define how much each agent's preferences vary from the group baseline, and how tolerant agents in the group are in affiliating with other groups. Table 1 shows the variation and affiliation threshold for each group. Groups with low sampling variation and high thresholds generate very similar trajectories, while those with higher variation and lower threshold generate more diverse trajectories.

6 Goals

The exogenous and presence features on event nodes support tactical decisions by agents, based on their preferences for these features. Hierarchical goal networks

Table 2 Example event trajectory for unaffiliated military agent

Domain Time	Event Number	Event Name	Short Name
1	E521	military perceives threat to peace and security	threat to peace and security is perceived
17	E536	military implements operational plans	operational plans are implemented
28	E351	government sets up checkpoints at official border crossings	Resource 12 set up at Location 50
143	E24	govt & armed opposition forces increase fighting in border regions	Activity 112 increases in Location 50
209	E281	government & armed opposition forces wage fierce battle for control of critical territory	control of critical Resource 110 under contention
409	E281	Agent dies	Agent dies

(HGNs) for each group modulate the urgency features, supporting strategic decisions.

6.1 Computing over HGNs

An HGN is a directed acyclic graph (not necessarily a tree) of goals with labeled edges (*and* and *or*) with a single root goal. Fig. 6 shows the top two levels for the Government HGN; the entire HGN has 27 goals with a maximum depth of four goals. A goal's satisfaction is computed from the satisfaction of its subgoals (those at the origin of edges terminating on it). If the subgoals connect through *or* edges, the satisfaction of the higher-level goal is the maximum of the satisfaction of the subgoals, while an *and* delivers the minimum. In this fragment, the satisfaction of the government's top-level goal 2095 "maintain continuity of rule" is the minimum
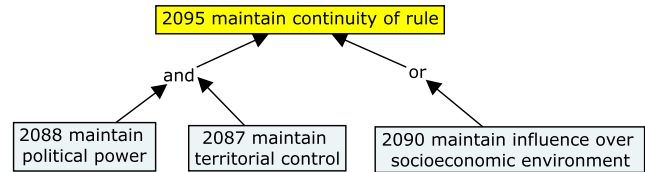

Fig. 6 Top of Government HGN

of the satisfaction of goal 2090 and the maximum of goals 2087 and 2088. The urgency of the root is 1 – satisfaction, and propagates down to subgoals.

Every HGN has leaf subgoals, those with no further subgoals. These derive their satisfaction from the total presence features on events in the CEG to which they are *zipped*. They can be zipped to any events, not just those for which the HGN's group has agency. Thus HGNs provide a second mechanism (in addition to influence edges) for interaction among groups. Zipping can either support (add to) or block (diminish) the satisfaction of the leaf goal. For instance, a leaf subgoal under 2087 "maintain territorial control" is 2034 "maintain military superiority." (Goal 2034 is at the bottom of the government HGN, and so does not appear in Fig. 6). Goal 2034 has six zips from the CEG. Three event types support it, two from the Government subgraph and one from the Environment:

1. govt defeats armed opposition forces on multiple fronts
2. government defeats violent extremists on multiple fronts
3. foreign governments provide military support to government

Three event types block it, two from the military and one from the armed opposition:

1. significant portion of the military deserts their posts
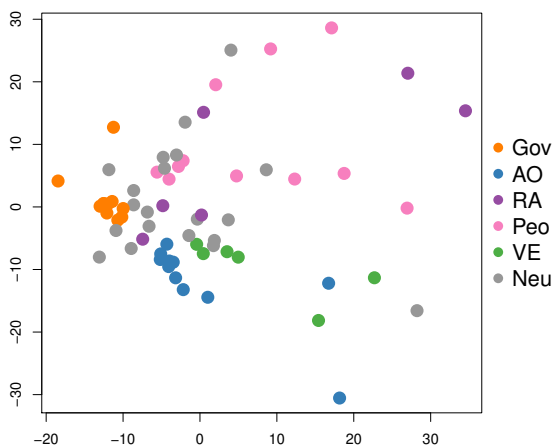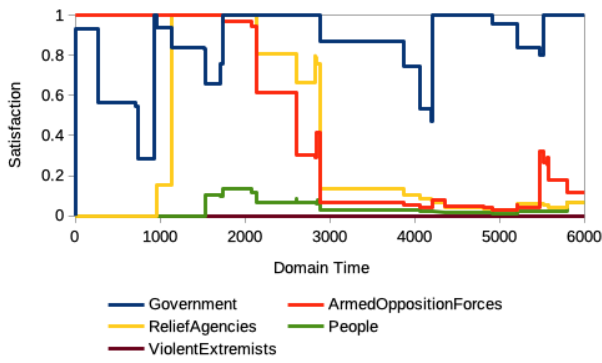

Fig. 5 Trajectory Similarities

Fig. 7 Satisfaction levels in an early experiment

2. military strength weakens
3. armed opposition forces kill large numbers of govt forces

The HGNs thus update an event type's urgency features from its presence features.

Though HGNs are specific to individual groups, they can impact agents in other groups. Each group's preferences can include not only its own urgency, but also the urgency of other groups (so that it can act to support or block them), and the actual preference that an agent shows for a given group's urgency is the weighted average of all groups with which it is affiliated.

Like the CEG, an HGN for each group is constructed by analysts in CMapTools. A tab in the Excel spreadsheet specifies which events are zipped to which goals, and whether they support or block those goals.

6.2 Group Satisfaction over Time

The research teams received goal information in the form of satisfaction levels for each individual agent, which is the average of the root satisfactions of the groups with which the agent is affiliated, weighted by affiliation strength. Some agents in each group had no affiliations, and so presented pure group satisfactions.

Fig. 7 shows an example of how satisfaction levels varied over time for different groups in an early experiment (in a version of the model without a distinct Military group). The Armed Opposition rapidly gains satisfaction, but then becomes more frustrated as Government satisfaction increases. Relief Agencies take longer to achieve their goals, but as long as they do, the People achieve some satisfaction. However, as the satisfaction of Armed Opposition decreases, so does that of Relief Agencies, and then that of People.

7 Geospatial Constraints

An important dimension of causality is the spatial movement of agents, initiated by some events in the CEG.

7.1 Geospatial Events and Movement

Fig. 8 shows a map of Tharum and the surrounding countries. The different colors represent meaningful regions, including cities, borders, water features, roads, and (small circles) camps for displaced persons both inside and outside Tharum, and are defined in the model's Excel spreadsheet.

A hexagonal grid on the map defines tiles between which agents move. Our map is 400x400 km, comparable to Syria, and each tile is 20km across.

Each agent is initialized at a geospatial location dependent on its home group. For example, Violent Extremists start in Muqaa and unofficial border areas (the magenta region just east of Muqaa), while Government agents start in large cities (Sag X).

Some events have specific geospatial destinations. For example, the destination of event type "people go to IDP camp" is a region identifying camps for internally displaced persons (IDPs). We call event types with destinations, "geospatial events." When an agent participates in such an event, it must move through geospace from its current location to the event's destination in order to complete the event. In this case the duration of its participation on the event is determined, not by the event's transit time, but by the time it takes the agent to reach the destination.

Movement through geospace uses the same feature-preference mechanism used in event space, but with a reinterpretation of the three kinds of features.

1. The *exogenous* features reflect difficulty of movement, from a terrain map.
2. SCAMP constructs a gradient field over the entire map for each region or location that can serve as a destination, and when an agent participates in a geospatial event, the gradient for that event's destination serves as the *urgency* features that it follows.
3. As a ghost moves from one tile to another, it deposits *presence* features for its group not only in the tile it currently occupies, but also in surrounding tiles, allowing other ghosts to respond to its presence. Thus agents interact stigmergically in geospace just as they do in event space.

The route the agent takes and the time it requires thus depend on the terrain, its destination, and its at-
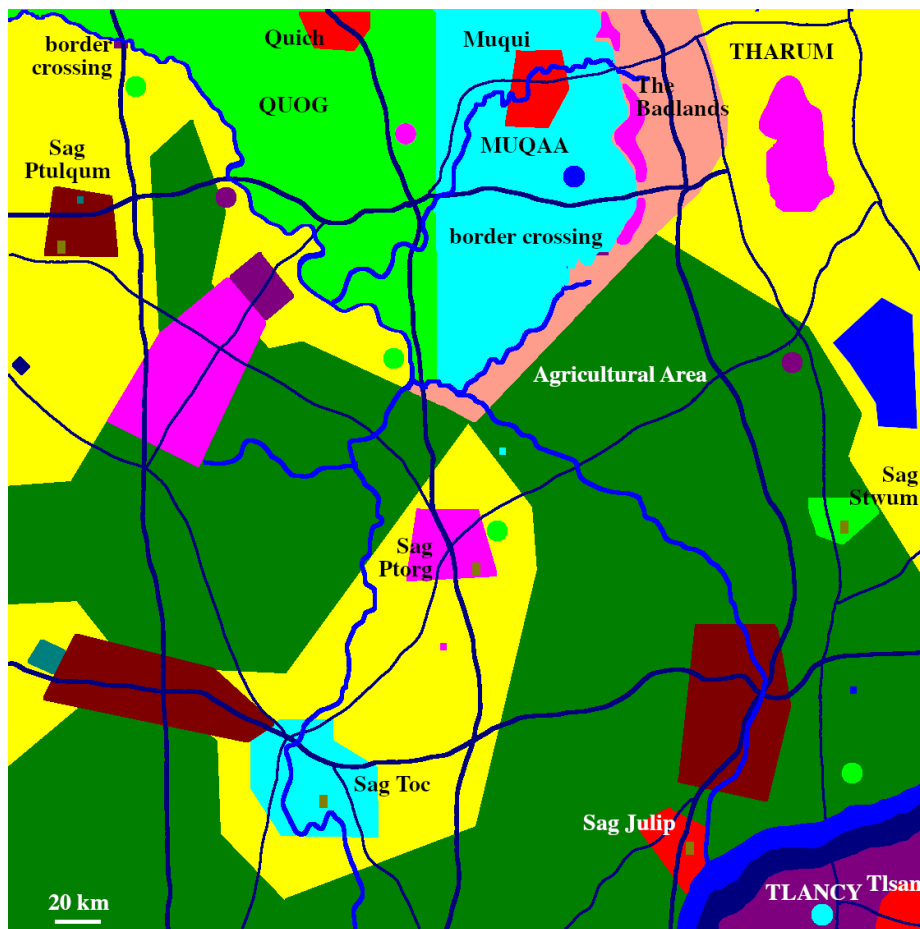
Fig. 8 Map of Tharum and neighbors Quog, Muqaa, and Tlancy

traction to or repulsion from other agents who are also in geospace.

Modelers prepare the geospatial map in GIMP [22], using separate layers for distinctive regions, and save it in the OpenRaster (ora) format. The layers are identified in the Excel spreadsheet for the model.

7.2 Examples of Geospatial Movement

Consider an example of agent movement under these constraints. A10263 is a Neutral agent affiliated with Relief Agencies, and participates in E211, "relief agencies send teams to affected areas to determine immediate needs." The destination for this geospatial event consists of large cities, which include Sag Ptulqum, Sag Ptorg, and Sag Julip. A10263 has a negative preference (-0.29) for Armed Opposition agents (blue), but a positive preference (0.99) for Government agents (orange). Fig. 9 shows the trajectory of A10263 (purple) as it executes this mission. Each purple dot represents the agent's location on a successive day. Its starting location on the west of Sag Toc is closer to Sag Ptulqum

and Sag Ptorg than to Sag Julip, so it moves northward, along a trajectory that commits to neither (left). It senses and moves away from the presence features left by Armed Opposition, keeping it from approaching Sag Ptorg. That repulsion, and attraction for Government agents (center), leads it to the west, and then north to reach Sag Ptulqum (right).

In addition to this overall trajectory, note how narrow the path of A10263 is. This agent is affiliated with Relief Agencies, and its preference for this group's presence features is positive (0.29), so its ghosts, who are planning its path, are attracted to each other, and tend to form a more focused plan for it to follow.

SCAMP allows the modeler to tune the impact of presence features on ghost movement in geospace. (Avatars still follow their ghosts' pheromone field.) Fig. 10 shows another run of the configuration in Fig. 9, with ghost presence preferences set to 0. Now A10263 ignores the presence of both Government and Armed Opposition, and its ghosts have no inclination to bunch together, but instead explore more widely. This difference has two implications.
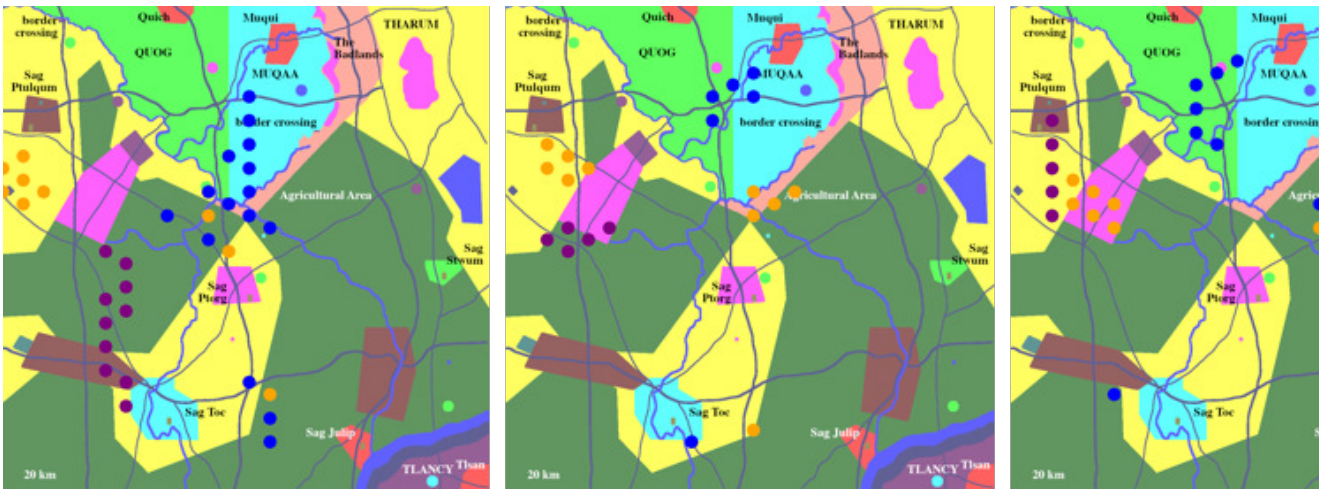
Fig. 9 Agent (Purple) evading Armed Opposition (Blue) and attracted to Government (Orange) while moving toward Sag Ptulqum or Sag Ptorg
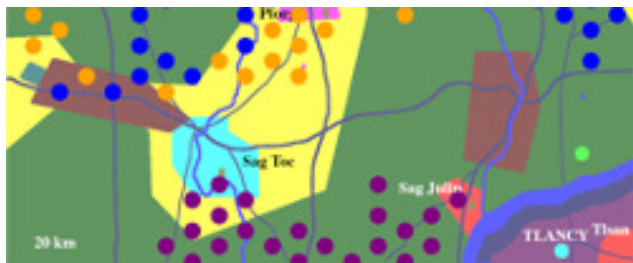
Fig. 10 Agent with no guidance from presence features

1. They discover a hex where the gradient for Sag Julip is stronger than the gradients for the other two large cities turned off, so they guide the agent to the east.
2. Because their guidance is more diffuse, the avatar's path is not direct, but more erratic.

8 Social Dynamics

SCAMP agents interact socially, and can change their basic group membership based on these interactions.

8.1 Social Interactions and Group Changes

Agents in SCAMP interact socially in three ways.

1. Though each agent has a home group, it can affiliate with other groups, based on similarity of preferences.
2. An agent can encounter other agents in event space (that is, the CEG) by participating in the same event type concurrently, or in geospace by being on the same tile concurrently. It builds relationships with other agents that it meets, relationships that increase in strength with the number and duration of encounters. These relationships modulate its own preferences based on the preferences of those with whom it interacts [4], implementing a form of group-think or conformity bias.
3. It constantly compares its actual probability of interacting with other agents with the probability of interaction it would expect based on group affiliations. If it finds itself behaving more like members of another group, it adjusts their urgency preferences to favor the goals of that group, and seeks to change its home group.

In keeping with the stigmergic model, agents do not exchange messages remotely with one another. But they do have access to the preference vectors of agents with whom they are colocated on the CEG or in geospace.

The group change mechanism can be attached either to event types or to geospatial locations. When an agent begins participating in such an event type or enters such a location, it triggers the change process, which may affect it or other agents, either collocated with it or elsewhere. For example, the participation of Government agents in Event 113 "government imposes its ideology or religion on its people" probabilistically triggers a group change that changes People who are participating in Event 492 "people radicalize" into Violent Extremists participating in Event 442 "violent extremists increase their numbers," and People who are participating in Event 18 "people arm themselves" into members of the Armed Opposition participating in Event 295 "armed opposition forces increase their numbers." Since Events 442 and 295 have high urgency for the Violent Extremists and Armed Opposition, respectively, People agents that develop a preference for one of those groups will

be attracted to these events, increasing the likelihood that they will change groups.

The group change mechanism also supports the virtual group Guf (in Jewish mysticism, the repository of souls). Executing a group change rule from Guf to a group adds a new agent to that group (for example, when a group recruits a foreigner, not previously represented in the model), while a change from a group to Guf removes an agent from the simulation (for example, a combat fatality). In a violation of the Jewish tradition, avatars do not exist in SCAMP's Guf before birth and after death, but are instantiated and deallocated as needed. Guf is simply the name used in in a group change rule in place of a group to indicate a birth or death event.

Group change rules are encoded in the model's Excel workbook.

8.2 Examples of Population Change

The result of the group change mechanism is visible in changing populations of the different groups over the course of a run. Fig. 11 shows the population levels in our Challenge 3 data. It shows several interesting and realistic features.

- The model defines death events (group change to Guf) for most groups, but birth events (changes from Guf) only for Violent Extremists and Armed Opposition. So the other groups lose population over time
- VE and AO do grow, but not uniformly. AO starts growing first, but then plateaus. Because VE is not present at the start of the war and originates outside the country, it does not begin to attract recruits at the start of the model.
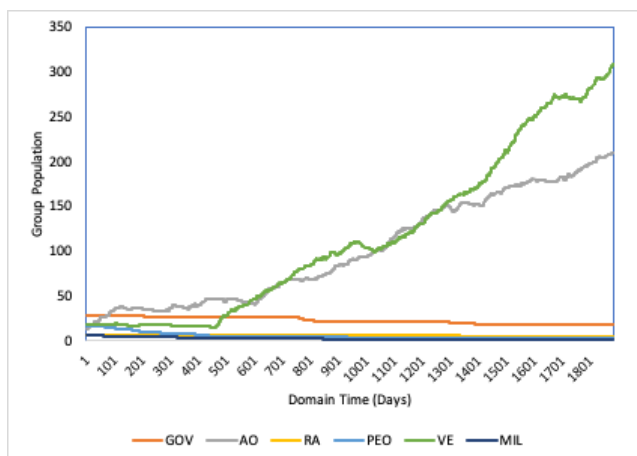

Fig. 11 Group populations over time

- Once VE begins to grow, AO also begins to grow. One can think of increasing competition for People interested in becoming more active in the conflict.
- For a while, VE and AO grow at about the same rate, but VE eventually outpaces AO and ends up with the stronger force.

9 Challenge Tests

Each challenge involved three tests: Explain, Predict, and Prescribe [20].

9.1 Explain

The Explain test was based on a causal graph, negotiated with the Test and Evaluation (T&E) team (Fig. 12). The nodes for Challenges 1 and 2 are white, while those for the geospace (yellow) and social (pink) perspectives were added for Challenge 3. These nodes are completely independent of the nodes in any of SCAMP's perspectives. Per T&E, we divided the nodes into three categories: the individual Actor, Groups of actors, and the System within which actors operated. Our ground truth does not include any details of the polyagent distinction between ghosts and avatars, since that is a computational mechanism that does not correspond to any natural social phenomenon or any data that we gave the research teams.

In Challenges 1 and 2 (the white nodes), the basic causal loop starts from node 8 (choose next event) to 7 (current event) (that is, choosing an event changes the current event). The loop continues to node 10 (event adjacency in the CEG), an important component of event eligibility (node 9a), which returns to 8. Other loops modulate both agent preferences (to the right) and features of events (to the left). The agency edges in the CEG implement the edge from 7 to 10, determining from the current event which events are accessible next.

In Challenge 3, agents participating in a geospatial event (node 7) drop into geospace (node 16) at their current location (node 18), making adjacent locations (node 22) accessible (node 9b) and leading to node 23 (choose next location), which in turn updates current location (18). When the agent reaches its destination (node 17a) or expends a limiting number of steps (node 24), it returns to event space to complete the current event (node 7). The pink nodes, representing the social perspective, modulate agent preferences based on other agents with whom they interact and enable agents to change their home group.
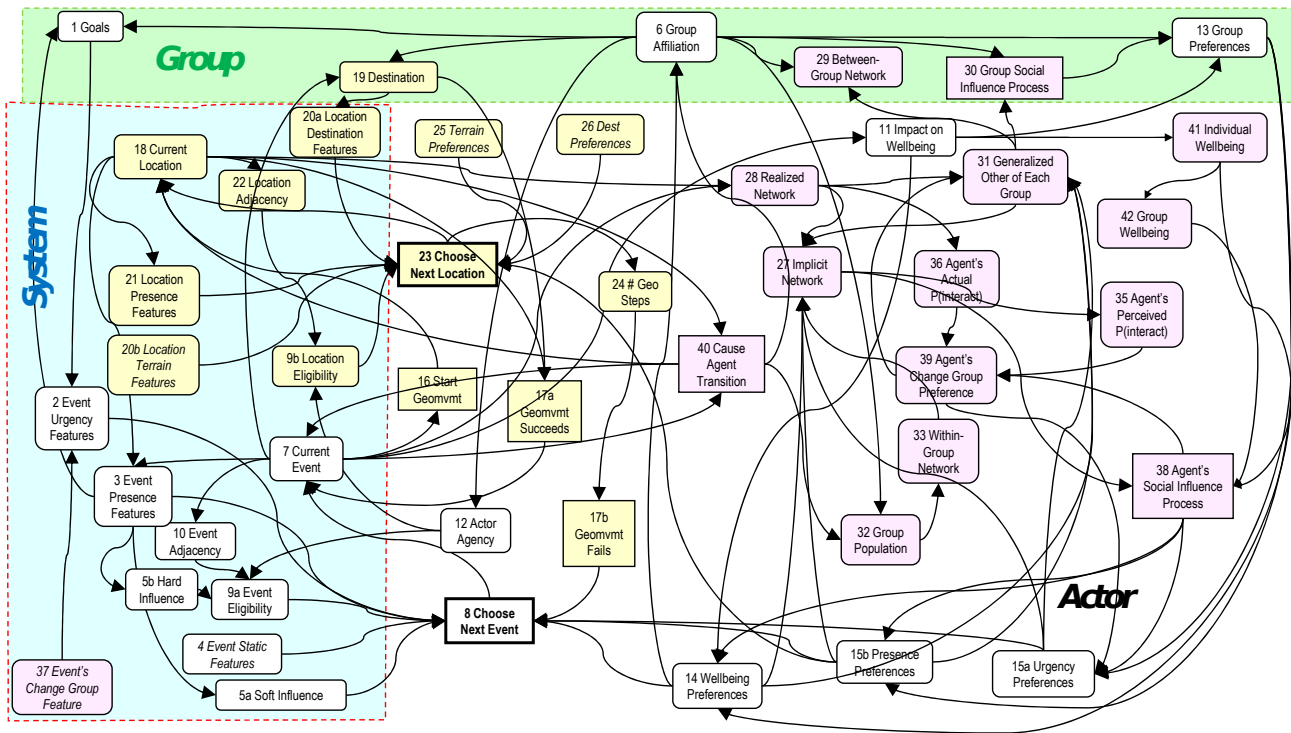
Fig. 12 SCAMP's Ground Truth

Research teams were scored based on the nodes and edges they identified from this diagram. Both teams identified numerous nodes that did not correspond to anything in our ground truth. Table 3 shows the number of nodes in each challenge, and the number of nodes that we could align, either exactly or (in most cases) approximately, with Fig. 12. Because most node identifications were only approximate, reporting the number of edges successfully recovered is very subjective.

9.2 Predict

As the Fourth MEB, we motivated Predict tests as information to guide upcoming decisions. In the Predict tests, we submitted the behavior of our system to T&E for comparison with the predictions from the research teams, but did not ourselves participate in the evaluation.

Table 3 Explain Results

Challenge	# Nodes	Chicago	JHU
1	17	1	7
2	17	1	8
3	46	9	19

9.2.1 Challenge 1

Here is how we introduced Challenge 1's Predict test.

Our CO has learned of the analysis you are performing with the data you have requested from the Donglap conflict, and has asked for your help in some upcoming decisions. The situation here is volatile and may change in several ways. In some cases, he is considering proactive operations. In others, we are aware of things that may change in theater, and we must respond appropriately. In each situation, our operational decisions will be affected by our estimate of the likely impact of each change.

The data that is most relevant to our decisions concerns levels of participation in a range of event types that we have identified, and we anticipate some of these will vary as a result of the changes in social, political, or military conditions that we may either initiate ourselves, or experience due to factors beyond our control. Our CO has asked us to report, for each possible change:

– For which event types will participation increase enough to merit our attention?
– For which event types will participation decrease enough to merit our attention?

– Which event types will have essentially the same level of participation that we are seeing now?

We defined a sufficient increase or decrease as the change in participation, divided by the square root of the original participation plus 1 (to avoid divide-by-zero problems). The denominator is roughly the standard deviation of a normal distribution around the original level, so the quotient approximates a standard score. We defined a sufficient increase or decrease as one for which this score was greater than 2 or less than -2, while we considered the level unchanged if it fell between -0.4 and 0.4. We also penalized wrong predictions, to discourage guessing. We defined the changes for which we sought predictions:

– Intel has identified event type E7930 as one that leads to a wide range of undesirable outcomes, and operations says that we could completely suppress it with appropriate deployment of our military and diplomatic resources. But if we do suppress it, it is critical to know how participation in other event types in Donglap might change if actors can no longer participate in E7930.

– We have recently observed that actors similar to A9979 are no longer participating in the society. We do not know why this is so, and do not expect you to be able to tell us. They have simply "disappeared" (unfortunately a not-uncommon event in this part of the world). We need to know how this is likely to change the participation levels of various event types.

– We have also observed a decrease (as much as 50%, though exact figures are uncertain) in the number of actors strongly similar to A4027. Again, we do not know why this is so, and do not expect you to be able to tell us, but we need to know the impact of this change on participation levels.

– As a result of rising tensions, actors avoid events likely to be visited by actors different than themselves. They apparently perceive that it's just not safe to be around people who are different. How will this increased cliquishness impact event participation?

9.2.2 Challenge 2

Though the underlying ground truth in Challenge 2 was the same as in Challenge 1, we altered the request for the Predict test:

Once again, we would be grateful for your help in predicting the likely future of events here in Donglap. We are interested in the levels of participation in certain types of events, and in the degree of satisfaction that different kinds of actors are likely to feel. In most of the queries we pose below, we are interested in the course of events from this point on in Donglap itself. That is, you have seen data from Donglap through day 730. We are interested in how that very same society will continue to unfold through day 930. Actor identifiers refer to the very same actors that they did for the first 730 days.

We are interested in levels of event participation for the following types of events: E8040, E6146, E1939. For these predictions, the level of participation in an event type at a given time is the number of actors who are participating in the event type at that time.

Satisfaction (generated by the HGNs, Section 6) was generated in Challenge 1, but not reported to the research teams, who were having difficulty recognizing the importance of groups goals in the ground truth. So in Challenge 2 we decided to report, unsolicited, a time series of this variable at the actor level (where it is the weighted average of the satisfaction levels for the groups with which the agent is affiliated).

We suspect that our informants are not homogeneous, and that the level of satisfaction they feel over time may not be the same across actors. We are interested in how their satisfaction changes over time. In particular, we would like predictions of actor satisfaction for the individual actor A3317, for actors like each of A3711, A7347, and A5939, and for the average satisfaction across the entire population.

All of the predictions requested so far vary over time. In addition, we would be interested in how the satisfaction level of actors similar to A7347 would change if there were twice as many actors of this sort present in the world, all other populations remaining the same. This prediction should be based on an alternate world that is like Donglap in all respect except for the population change. We seek a single number, the aggregate satisfaction of this class of actors for the first 500 days of such a world. By aggregate satisfaction, we mean that if you were to draw a curve of the satisfaction for a typical actor of this type over the 500 days, we are looking for the area under the curve. (A smart-aleck 2nd Lt in S6 mumbled something about "integral" and

"trapezoidal rule," but the rest of us boxed him pretty roundly around the ears for showing off.)

An important feature of prediction in complex systems is the existence of a prediction horizon beyond which any prediction is no better than random [28]. We suspected that the lack of accuracy reported by T&E for the predictions in Challenge 1 might be due to this effect, so we asked the research teams for time series rather than point predictions, allowing T&E to both accuracy and the window over which that accuracy persists.

> The last requested prediction (the aggregate satisfaction) does not vary through time, but please provide not only your best point prediction, but also pairs of predictions defining 50%, 90%, and 95% confidence intervals. For the other predictions, which vary through time, please provide predictions as a daily time series. Again, we would be grateful for your best point prediction and pairs of predictions defining 50%, 90%, and 95% confidence intervals.

9.2.3 Challenge 3

We posed three predictive questions in Challenge 3. Again, we asked for time series rather than point predictions, permitting T&E to measure prediction horizon as well as accuracy. We also included a prediction request involving manipulation of geospace.

> Your advice was so helpful in our last deployment to Dongglap that we would like to ask your help with some predictions about our current situation in Tharum.
>
> We would be grateful for three sets of predictions. All three concern the various groups of agents that you may have discerned are active in Tharum.
>
> First, you have seen data from the first 1886 days of our deployment. We would like to know what's coming in the next 100 days. Please give us time series for this extended period on the total population and satisfaction for each of the groups you have discovered. You may identify the groups to us by giving us the identifiers for one or more agents whom you believe exemplify the group, but we want the population and satisfaction for the overall group, not for any single agent you give us.
>
> Second, we expect that E297 is causing trouble. In a world that is like Tharum, what would a time series of the population and satisfaction for the various groups look like for the first 200

days if we were able to suppress this event, compared to what we saw in the first 200 days in Tharum?

> Third, four regions (shown in color in the attached map) also appear to be causing difficulty. Again for a world that is causally like Tharum, what would the population and satisfaction time series look like for the first 200 days in a world where we make these regions off-limits to all actors except those who happen to be there at the beginning?

9.3 Prescribe

In the Prescribe test, the research teams were given desired outcomes, and asked to prescribe actions that would achieve them. Available actions were those that we had hypothesized in the Predict tests, for example,

– Exclude an event type from the CEG;
– Make a particular spatial region inaccessible;
– Change the population of a particular group;
– Reduce interaction between specified groups.

The basic methodology in our Prescribe evaluation was to compare the distribution of five runs of the baseline configuration with five runs of each prescription, and assign two scores: an easy one (how far the median of the prescription runs dominates the median of the baseline runs) and a more stringent one (how far the worst result from the prescription dominates the best from the baseline runs). We also gave T&E the results of the best prescription we could devise for each problem, based on our full knowledge of the ground truth. We asked the research teams to achieve the best improvement they could with the fewest interventions and computed gain per intervention, but here we report only number of prescriptions that dominate the baseline, according to the two metrics. We consider a result to dominate the baseline only if the improvement divided by the baseline median is strictly positive.

Because SCAMP runs can be time-consuming and (at the time of these tests) we did not have the ability to cache a run and resume it later, Prescribe runs were all done on a world with the same ground truth as the initial data that the research teams received, but starting from the beginning rather than from the latest date that they had seen.

9.3.1 Challenge 1

We tasked the research teams with two problems. The first seeks to minimize or maximize participation on specified event types:

We would like to minimize the level of participation of the population in certain destabilizing types of events, while increasing their participation in others that we consider favorable. Below we list the types of events in each category. Within each list, the event types are listed from what we expect will be the easiest for you to impact, to the most challenging. The order below is our estimate of difficulty. However, we realize that your expertise may identify a different ordering.

We asked them to minimize ten events, and maximize 15.

The second problem concerns interactions (concurrent event participation) between actors from different groups:

> We observe that the interactions between different types of individuals are sometimes helpful to our mission, while others are detrimental, We would like to vary the interaction between actors of different types, based on their similarity to certain designated individuals whom we have been tracking. In some cases, we are interested in alternatives that could either increase or decrease interaction for the same two types of actors.
>
> Here are the types of actors of interest:
> - Type A shows behavior like that exhibited by A1686 around time 3100.
> - Type B shows behavior like that exhibited by A4927 around time 2200.
> - Type C shows behavior like that exhibited by A9773 around time 2000.
>
>   Please make separate recommendations to
> - increase interactions between actors of types A and B,
> - decrease interactions between actors of types A and C,
> - increase interactions between actors of types B and C, and also
> - decrease interactions between actors of types B and C.

Table 4 shows the results. A common intervention available to the research teams was to exclude certain events from the CEG. If too many events are excluded, the CEG may become disconnected and the model may fail to run, so in some cases both teams do not report on the same number of problems. We report results as fractions in which the numerator is the number of prescriptions that dominate the baseline while the denominator is the total number of runnable prescriptions provided.

Table 4 Challenge 1 Prescriptions that Dominate Baseline

Problem	Chicago Easy	Chicago Hard	JHU Easy	JHU Hard
Min/Max Participation	7/22	0/22	24/25	16/25
Inc/Dec Interaction	3/4	0/4	4/4	2/4

9.3.2 Challenge 2

The Prescribe test for Challenge 2 built on concepts from the prescriptions requested in Challenge 1 to minimize or maximize event participation, as well as on the notion of aggregate satisfaction introduced in the Challenge 2 Predict test. We posed eight problems:

- Maximize participation in E5451
- Minimize participation in E8368
- Maximize participation in E6739
- Maximize participation in E954
- Maximize the aggregate satisfaction of agents like A5205
- Maximize the aggregate satisfaction of agents like A3711
- Minimize the aggregate satisfaction of agents like A7347
- Maximize the aggregate satisfaction of agents like A2180

Table 5 shows the results. There are four problems of each type, and all prescriptions were runnable.

9.3.3 Challenge 3

Here is the tasking memo that the research teams received for the final Prescribe test:

> Our CO is considering several alternative tactical objectives in Tharum. We are hopeful that your knowledge of what is going on in the

Table 5 Challenge 2 Prescriptions that Dominate Baseline

Problem	Chicago Easy	Chicago Hard	JHU Easy	JHU Hard
Min/Max Participation	2	0	1	1
Min/Max Satisfaction	2	0	1	0

society here would enable you to recommend courses of action to achieve each of three outcomes.

These are independent recommendations: we might pursue 1, or 2, or 3, and want to know the best course of action in each case. We do not at this time seek a course of action to achieve all three concurrently.

By the time the Prescribe test for Challenge 3 took place, COVID-19 was in full swing, so we wove that into our request to the research teams, to motivate our request to achieve maximum impact with the fewest interventions:

As before, the value of a recommendation to us depends not only on the magnitude of the improvement over the status quo, but also the number of interventions required to achieve it. The larger the improvement and the fewer interventions required, the better. (You can think in terms of maximizing impact per intervention.) In addition, we just received a copy of the Stars and Stripes from early March (mail here arrives by camel train and is quite slow). It reports policies being implemented back stateside that have the effect of shutting down the society. Such recommendations are not practical in Tharum. (Recommendations that try to block too many events or locations can have this effect.)

In all cases, we are concerned with groups of agents of various types. As previously, we define each type by identifying a single agent representative of the type, but the questions we wish to address in all cases concern groups of agents, not just these specific individuals

The first of the three problems requires the research teams to maximize the population of the Armed Opposition over the Government. The second seeks to maximize the aggregate satisfaction of Armed Opposition over the Government, while the third seeks to minimize visits by military groups (Government, Armed Opposition, Violent Extremists, and Military) to refugee camps in other countries.

1. The US Government wishes to support agents like A6313, particularly in comparison with agents like A1854. What steps should we take so that by day 130, the difference in total population of the first type over that of the second type is maximal? Recommendations of the form "Reduce the number of agents similar to A1854" or "Increase the number of agents similar to A6313" are too vague to be useful to us in this case.

2. Recall the concept of aggregate satisfaction from our prescription request for Donglap. We wish to maximize the aggregate of the difference in satisfaction between agents like A6313 and those like A1854. How can we do this?

3. We have learned that the following locations in neighboring countries are crucial to humanitarian relief: Locations 10x3, 11x2, 14x4, 15x3, 20x20, 21x19, and 21x20. But the presence of military forces in these areas (agents like A6313, A1854, A3779, and A2763) is hindering this activity. What can we do to reduce the total number of visits of agents of these four types to these locations over the first 130 days? NB: One action available to us within Tharum is to blockade a region of interest to prevent agents from entering it, but since we do not have a Status of Forces agreement with the neighboring countries, this type of operation is not available to us in pursuing this objective.

The third problem illustrates how our military persona provides a rationale for asking for prescriptions that requires an understanding of the relation between events and geospatial locations, rather than allowing them simply to exclude the locations listed.

Table 6 shows the results. All prescriptions were runnable on all three problems.

10 Summary

Conflict World exhibits several important features of SCAMP.

– The entire model is constructed by domain experts, professional analysts who are not programmers, us-

Table 6 Challenge 3 Prescriptions that Dominate Baseline

Problem	Chicago Easy	Chicago Hard	JHU Easy	JHU Hard
Max Pop AO/Gov	0	0	0	0
Max Sat AO/Gov	0	0	1	0
Min Visits to Camps	1	0	0	0

ing three common desktop tools: the CMap concept mapper, GIMP, and Excel. Thus it can easily capture their insights without the knowledge acquisition barrier between domain experts and programmers common in many applications, and allows us to build models of realistic complexity.

- The resulting model artifacts give an unambiguous causal ground truth that underlies the data given to the research teams, and is easily modified to examine the causal impacts.
- In spite of stigmergy's simplicity, Simon's Law enables SCAMP to capture psychologically and socially realistic dynamics [25,32], including deliberate tactical choice guided by preferences over alternatives, non-deterministic decision-making [3], strategic (goal-driven) as well as tactical decisions [40], use of mental simulation to look ahead in time [16], interactions with other agents encountered on events or geospatial tiles as a mechanism for adjusting individual preferences [11], the centrality of narrative as a mental representation [10], and naturally bounded rationality [42].

Our experience in the program led to several lessons.

- Adopting a detailed backstory and persona was very helpful in managing our interactions with the research teams.
- The requirement repeatedly urged by T&E not to give away our ground truth often made it difficult to respond to questions that in the real world would have been legitimate. It would be interesting to consider alternative ways to validate causality that avoid this constraint. For example, we might give the research teams our ground truth together with a number of data sets, some generated by the ground truth and some by increasingly severe ablations of it, and see how well they could identify the real ones.
- Different disciplines in the social sciences talk about the world in different ways. The lack of a shared ontology for psychological and sociological phenomena was a major hindrance in our communication with the research teams.

Interested researchers can obtain the model under the Gnu Public License, and a detail user manual, describing the construction of the various configuration files, is available [33].

References

1. Argonne National Laboratory: Repast agent simulation toolkit (2007). URL http://repast.sourceforge.net/
2. Bonabeau, E., Dorigo, M., Theraulaz, G.: Swarm Intelligence: From Natural to Artificial Systems. SFI Studies in the Sciences of Complexity. Oxford University Press, New York (1999)
3. Busemeyer, J.R., Townsend, J.T.: Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. Psychological Review **100**(3), 432–459 (1993)
4. Camazine, S., Deneubourg, J.L., Franks, N.R., Sneyd, j., Theraulaz, G., Bonabeau, E.: Self-Organization in Biological Systems. Princeton Studies in Complexity. Princeton University Press, Princeton, NJ (2001)
5. Álvaro Carrera, Iglesias, C.A.: B2di - a bayesian bdi agent model with causal belief updating based on msbn. In: Proceedings of the 4th International Conference on Agents and Artificial Intelligence - Volume 2: ICAART,, pp. 343–346. INSTICC, SciTePress (2012). DOI 10.5220/0003746003430346
6. Cioffi-Revilla, C.: Introduction to Computational Social Science, 2nd edn. Texts in Computer Science. Springer, Cham, CH (2017)
7. Cox, T.F., Cox, M.A.A.: Multidimensional Scaling, second edn. Chapman & Hall (2000)
8. Dickau, R.: Counting lattice paths (2020). URL https://www.robertdickau.com/lattices.html
9. Epstein, J.M.: Generative Social Science. Princeton Studies in Complexity. Princeton University Press, Princeton, NJ (2006)
10. Fisher, W.R.: Human Communication as Narration: Toward a Philosophy of Reason, Value, and Action. University of South Carolina Press, Columbia, SC (1989)
11. Friedkin, N.E., Johnsen, E.C.: Social Influence Network Theory: A Sociological Examination of Small Group Dynamics. Structural Analysis in the Social Sciences. Cambridge University Press, Cambridge (2011)
12. Gilbert, N., Troitzsch, K.G.: Simulation for the Social Scientist, 2 edn. Open University Press, Buckingham, United Kingdom (2005)
13. Grassé, P.P.: La reconstruction du nid et les coordinations inter-individuelles chez bellicositermes natalensis et cubitermes sp. la théorie de la stigmergie: Essai d'interprétation du comportement des termites constructeurs. Insectes Sociaux **6**, 41–84 (1959)
14. Grimm, V., Berger, U., DeAngelis, D.L., Polhill, J.G., Giske, J., Railsback, S.F.: The odd protocol: A review and first update. Ecological Modelling **221**(23), 2760–2768 (2010). URL http://www.agsm.edu.au/bobm/teaching/SimSS/ODD_protocol.pdf
15. IHMC: Ihmc cmaptools - download (2013). URL https://cmap.ihmc.us/cmaptools/
16. Kahneman, D., Tversky, A.: The simulation heuristic. In: D. Kahneman, P. Slovic, A. Tversky (eds.) Judgment under Uncertainty: Heuristics and Biases, pp. 201–208. Cambridge University Press, Cambridge, UK (1982)
17. Levenshtein, V.I.: Binary codes capable of correcting deletions, insertions, and reversals. Soviet Physics Doklady **10**(8), 707–710 (1966)
18. Luke, S., Cioffi-Revilla, C., Panait, L., Sullivan, K., Balan, G.: Mason: A multiagent simulation environment. SIMULATION **81** (2005)
19. de Marchi, S.: Computational and Mathematical Modeling in the Social Sciences. Cambridge University Press, Cambridge, UK (2005)

20. Naugle, A., Russell, A., Lakkaraju, K., Swiler, L., Verzi, S., Keywords:, V.R.: The ground truth program: Simulations as test beds for social science research methods. Computational and Mathematical Organization Theory **(forthcoming)** (20210)
21. Parunak, H.V.D.: 'go to the ant': Engineering principles from natural agent systems. Annals of Operations Research **75**, 69–101 (1997)
22. Parunak, H.V.D.: A survey of environments and mechanisms for human-human stigmergy. In: D. Weyns, F. Michel, H.V.D. Parunak (eds.) Proceedings of E4MAS 2005, *Lecture Notes on AI*, vol. LNAI 3830, pp. 163–186. Springer (2006)
23. Parunak, H.V.D.: Between agents and mean fields. In: J. Sabater, D. Villatoro, J. Sichman (eds.) the Workshop on Multi-Agent-Based Simulation (MABS 11) at AAMAS 2011, LNAI, pp. 106–117. Springer, Taipei, Taiwan (2011)
24. Parunak, H.V.D.: Odd protocol for scamp. Report, Wright State Research Institute (2020). URL `https://abcresearch.org/abc/papers/ODD4SCAMP.pdf`
25. Parunak, H.V.D.: Psychology from stigmergy. In: Computational Social Science (CSS 2020), vol. (forthcoming). CSSSA (2020)
26. Parunak, H.V.D.: How to turn an mas into a graphical causal model. Autonomous Agents and Multi-Agent Systems p. (under revision) (2021)
27. Parunak, H.V.D.: Social simulation for non-hackers. In: K.H. Van Dam, N. Verstaevel (eds.) 22nd International Workshop on Multi-Agent-Based Simulation (MABS 2021). Springer (2021)
28. Parunak, H.V.D., Belding, T.C., Brueckner, S.A.: Prediction horizons in agent models. In: D. Weyns, S. Brueckner, Y. Demazeau (eds.) Engineering Environment-Mediated Multiagent Systems (Satellite Conference at ECCS 2007), vol. LNCS 5049, pp. 88–102. Springer (2008)
29. Parunak, H.V.D., Brueckner, S.: Concurrent modeling of alternative worlds with polyagents. In: the Seventh International Workshop on Multi-Agent-Based Simulation (MABS06, at AAMAS06), pp. 128–141. Springer (2006)
30. Parunak, H.V.D., Brueckner, S.A.: Engineering swarming systems. In: F. Bergenti, M.P. Gleizes, F. Zambonelli (eds.) Methodologies and Software Engineering for Agent Systems, pp. 341–376. Kluwer (2004)
31. Parunak, H.V.D., Brueckner, S.A., Matthews, R., Sauter, J.: Swarming methods for geospatial reasoning. International Journal of Geographical Information Science **20**(9 (Oct)), 945–964 (2006)
32. Parunak, H.V.D., Morell, J.A., Sappelsa, L.: The scamp architecture for social simulation (2020). In preparation for JASSS
33. Parunak, H.V.D., Morell, J.A., Sappelsa, L., Greanya, J.: Scamp user manual. Report, Parallax Advanced Research (2020). URL `https://www.abcresearch.org/abc/papers/SCAMPUserManual.zip`
34. Parunak, H.V.D., Purcell, M., O'Connell, R.: Digital pheromones for autonomous coordination of swarming uav's. In: First AIAA Unmanned Aerospace Vehicles, Systems,Technologies, and Operations Conference. AIAA (2002)
35. Parunak, H.V.D., Savit, R., Riolo, R.L.: Agent-based modeling vs. equation-based modeling: A case study and users' guide. In: N. Gilbert, R. Conte, J.S. Sichman (eds.) Multi-Agent Systems and Agent-Based Simulation, First International Workshop, LNCS, pp. 10–25. Springer, Berlin, Germany (1998)
36. Pynadath, D.V., Dilkina, B., Jeong, D.C., John, R.S., Marsella, S.C., Merchant, C., Miller, L.C., Read, S.J.: Disaster worlddecision-theoretic agents for simulating population responses to hurricanes. Computational & Mathematical Organization Theory **(forthcoming)** (2021)
37. Rager, S., Leung, A., Pinegar, S., Mangels, J., Poole, M.S., Contractor, N.: Groups, governance, and greed: The access world model. Computational & Mathematical Organization Theory **(forthcoming)** (2021)
38. Rao, A.S., Georgeff, M.P.: Bdi agents: From theory to practice. In: the First International Conference on Multi-Agent Systems (ICMAS-95), pp. 312–319. AAAI (1995)
39. Sappelsa, L., Parunak, H.V.D., Brueckner, S.: The generic narrative space model as an intelligence analysis tool. American Intelligence Journal **31**(2), 69–78 (2014)
40. Shivashankar, V.: Hierarchical goal networks: Formalisms and algorithms for planning and acting. Ph.d., Computer Science (2015)
41. Simon, H.: The Sciences of the Artificial. MIT Press, Cambridge, MA (1969)
42. Simon, H.A.: A behavioral model of rational choice. The Quarterly Journal of Economics **69**(1), 99–118 (1955)
43. Sterman, J.: Business Dynamics. McGraw-Hill, New York, NY (2000)
44. Sullivan, K., Coletti, M., Luke, S.: Geomason: Geospatial support for mason. Tech. rep., George Mason University (2010)
45. Sumpter, D.J.: Collective Animal Behavior. Princeton University Press, Princeton, NJ (2010)
46. Team, T.G.: Gimp: Gnu image manipulation program (2020). URL `www.gimp.org`
47. Vidal, J.M., Durfee, E.H.: Recursive agent modeling using limited rationality. In: First International Conference on Multi-Agent Systems (ICMAS 95), pp. 125–132 (1995)
48. Wagner, R.A., Fischer, M.J.: The string-to-string correction problem. Journal of the ACM **21**(1), 168–173 (1974)
49. Wilensky, U.: Netlogo (1999). URL `http://ccl.northwestern.edu/netlogo`
50. Züfle, A., Wenk, C., Pfoser, D., Crooks, A., Kavak, H., Kim, J.S., Jin, H.: Urban life: A model of people and places. Computational & Mathematical Organization Theory **(forthcoming)** (2021)