ODD Protocol for SCAMP

H. Van Dyke Parunak

This document describes the SCAMP simulator in the ODD (Overview, Design concepts, Details) protocol [2,3]

1 Purpose

SCAMP (Social Causality using Agents with Multiple Perspectives) was developed under the DARPA Ground Truth program to generate socially and psychologically realistic data with a known causal basis, for use in testing analysis methods in the social sciences.

2 Entities, state variables, and scales

2.1 Entities

SCAMP's entities are:

- Groups to which agents can belong
- Event types, organized into a directed graph (a Causal Event Graph, or CEG) that forms a narrative space [10]. That is, any trajectory through the graph is a coherent narrative from the perspective of an agent.
- Agents representing either individuals or groups. Following the polyagent model [8], they are of two forms: avatars, and ghosts.
- Goals and subgoals, organized into Hierarchical Goal Networks (HGNs)
- Hexagonal spatial tiles
- Spatial regions, associated with one or more tiles

2.2 State Variables

We introduce each entity and describe its variables. Many of these definitions involve feature space, so we begin by describing that.

2.2.1 Feature Space

The central action of a SCAMP agent is repeatedly choosing among available alternatives (which may be either events in which it may participate, or locations to which it may move). These choices are guided by its *preferences* over the *features* of the available alternatives. Both of these are vectors of scalars in $[-1,1]^n$.

Feature space is the vector space in which preferences and features are defined.

- *Exogenous features* are defined by the modeler or supplied in real time from an external data feed. In our model for Ground Truth, these features are three in number: physical, emotional, and psychological well-being. For instance, a preference of 0.1 for physical well-being means that agents affiliated with this group 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.
- Each group has one *urgency feature* in feature space, reflecting the state of that group's HGN. A group's preference for a given group's urgency feature reflects its desire to promote (preference > 0) or block (preference < 0) that group's goals.
- Each group has one *presence feature*. A group's preference for a given group's presence feature reflects its attraction to or repulsion from members of that group.

An agent chooses its next event by

- Computing the cosine distance (normalized dot product) between its preference vector and the feature vector of each alternative
- Exponentiating each of these values to get non-negative numbers
- Raising each value to a fixed determinism parameter (where 0 gives all alternatives equal weight and values greater than 1 favor the stronger elements)
- Doing roulette selection over the resulting set of values.

2.2.2 Groups

Every agent belongs to one group, and may affiliate with one or more others. In general, groups represent recognizable categories of agents in the problem domain (for example, Government, Citizens, Relief Agencies, Terrorists, …), but there are three special groups: Environment, Neutral, and Gufe.

- Environment is a group whose agents generate background events (e.g., droughts, financial collapse) that are causally relevant to the domain, but that we do not model in detail.
- Neutral agents are assigned random preferences, and then join whichever group is closest to their preference vector.
- Gufe (named for the repository of souls in Jewish mysticism) is a special group from which agents come when they are born, and to which they go when they die.

We also use Gufe as the name for a notional event to which dead agents can move, and from which new agents come.

Nominal preferences: Each group has a nominal preference vector, which is sampled by its agents when they are generated.

Preference variation: Each group has a variation v in [0, 1]. Each preference for a new agent is selected uniformly from the group's nominal preference $\pm v$.

Affiliation parameters: Groups can differ on

- Whether their members can affiliate with other groups or not (the Environment agent, for example, does not)
- Whether other groups can affiliate with them
- A threshold in [0, 1] specifying how close (1 - cosine distance) an agent's preference vector must be to the group's in order to affiliate (where 0 allows any positive proximity to affiliate, and 1 prevents all affiliation

Execution parameters: Three parameters control the polyagent model. In this model, each domain entity is represented by an *avatar*, which thinks ahead in time by sending out stigmergic *ghost* agents to explore possible futures. The avatar sends out a fixed number of *shifts* of ghosts on successive simulation ticks, each shift with a fixed number of ghosts. These ghosts run a fixed number of steps into the future, laying down pheromone corresponding to their group on the events or geospatial locations that they visit. Multiple shifts allow later ghosts to respond to the presence of previous ghosts. Each group specifies the number of ghosts per shift (more allows exploration of more possible futures), the number of shifts (more allows more ghost interaction and thus yields a more considered estimate of the future), and the number of steps to look ahead.

The ghosts follow the full preference-feature mechanism described above, while the avatars form their roulette based only on the presence pheromone of their group.

Home Regions: Each group can have up to three geospatial regions, each with an associated probability, to which its agents are initialized.

Movement Delay: Different groups may move at different speeds through geospace, and a group parameter indicates how much agents of that group are slowed down compared to the nominal transit speed of a spatial tile.

Relations: Each group maintains a list of the other groups and how close it is to each of them, based on the inter-agent relations reported by its agents.

Number of agents: Each group can specify the number of agents that should be initialized for it.

2.2.3 Event Types

Event types are nodes in the narrative space. In addition, for symmetry of coding, we allow Gufe as an event type. Strictly speaking, an event is a period of time during which agents are continuously participating on an event type node, and events can be repeated. But where there is no risk of confusion, we may call a node an event.

Agency: For each group, whether agents of that group can participate in the event.

Destination: For each group with agency, an optional region identifier. If present, any agent in that group participating in the event drops into geospace and moves to the destination before completing the event.

Features: A vector in feature space characterizing the event. In our present model, the exogenous features represent the impact of the event on an agent's physical, psychological, and economic well-being.

Edges: Two kinds of directed edges connect events to each other. *Agency edges* from one event to others indicate next events that an agent currently on one event can select next. They are of two types: *then* edges allow an agent to move from one event to another, while *thenGroup* edges allow an agent to move from one group to several concurrent events. *Influence edges* modulate the probability of the target event being chosen based on the total presence pheromones on the source event. Hard influence edges (Enable, Prevent) probabilistically make the target event accessible or inaccessible, while soft influence edges (Enhance, Inhibit) increase or decrease the probability of the event's being chosen.

Probability: Relative weight of an event, independent of agent preferences. In the absence of preferences, agents normalize these values across accessible events, and choose proportionately.

Zips: Two kinds of edges can connect events to the leaf subgoals of HGNs. A *support* zip means that total participation in the event (across all groups) increases the satisfaction level of the subgoal, while a *block* zip means that total participation in the event decreases the satisfaction.

Transit Time: The parameter of an exponential distribution used to compute the length of an agent's participation in an event (corresponding to the interarrival times of a Poisson distribution). If the event has a destination, the transit time is ignored, and the duration of the agent's journey through geospacer is used instead.

Effect Time: How long the agent's participation continues to influence the event after the agent begins participation. In general, Effect Time > Transit Time.

Exclusion flags: An event may be marked to exclude agents of a given group over some period of time.

Group Changes: Group changes are rules that are activated when an agent of the right group participates in an event containing the group change. The rule specifies:

- Trigger to which the rule is attached: An event ID in CEG, or a region in geospace (which expands to a set of tiles)
- What groups must be participating in the trigger to enable the transition
- The group(s) whose agents are vulnerable to the action of the rule; Gufe indicates the birth of an agent.
- The group to which vulnerable agents are transitioned; Gufe indicates the death of an agent.
- The location (event, region, or Gufe) from which the vulnerable agents are drawn
- The maximum number of agents to transition
- The location (event, region, or Gufe) where the agents reside after the transition
- What groups must be present on the from location to enable the transition
- The base transition probability for the event
- PromoterGroups: groups whose level of participation on either the trigger or the from location increases the base probability of the transition
- BlockerGroups: groups whose level of participation on either the trigger or the from location decreases the base probability of the transition

2.2.4 Agents

SCAMP's agents include one persistent avatar for each domain entity and the transient ghosts that it generates to reason about the future. Ghosts use the full preference-feature decision algorithm outlined above, while avatars follow the presence features laid down by ghosts in their group.

Group affiliations: Each agent has a home group, with which it is affiliated with strength 1. In addition, it can affiliate with other groups, with strength dependent on its preferences.

Wellbeing: This parameter is specific to our current model, which uses exogenous features to represent the impact of an event on the agent. Wellbeing has one component for each exogenous feature, and changes as the agent participates in events, based on the exogenous features of those events. Events with a negative wellbeing feature diminish the associated dimension of the agent's wellbeing, while those with a positive wellbeing feature increase that dimension of the agent's wellbeing.

Relations: Other agents whom this agent has encountered, either in event space or in geospace, and a strength of relation that increases with number and duration of encounters.

Preferences: A vector in feature space, sampled from the home group's baseline preference vector. Through time, the wellbeing dimensions of the preference vector change, based on the agent's wellbeing parameter. A positive wellbeing parameter reduces the agent's preference for that dimension of event features, while a negative parameter increases the preference for the associated event feature. In addition, the urgency preferences can change, based on the agent's relations, if its actual frequency of interaction with other groups differs from that which it would expect based on its group membership, and all of its preferences can be influenced by the other agents with which it associates.

2.2.5 Goals and subgoals

Each group can have a hierarchical goal network (HGN) [11] describing its overall objectives. The nodes in such a network are states of the world that it would like to accomplish. The HGN forms a tree, with a root goal successively analyzed into subgoals.

Satisfaction: A number in [0, 1] indicating the degree to which the goal is currently satisfied.

Urgency: A number in [0, 1] indicating how urgent the goal is to satisfaction of the root goal. At the root goal, urgency = 1 – satisfaction, and the urgency of subgoals depends on the urgency of their higher-level goals [7].

Edges: Each subgoal is connected to one or more higher-level goals by labeled edges. AND edges between a set of subgoals and a goal indicate that all of the subgoals must be satisfied to satisfy the higher-level goal, while OR edges indicate that only one of the subgoals is required to satisfy the higher-level goal.

Zips: Some subgoals are leaves in the HGN, without lower-level subgoals. Each of these subgoals is connected to events whose participation level can either support or block its satisfaction. These connections correspond to the zips defined above for events.

2.2.6 Spatial tiles

Geospace is divided into hexagonal tiles, and agents can move from one tile to any of its six adjacent neighbors. Movement off the edge of the grid is not permitted; edge tiles have fewer than six outgoing edges.

Each spatial tile has a feature vector to guide agents. The exogenous and urgency features have special meaning, but presence features continue to be defined by the participation of agents.

Terrain: The exogenous features all contain the physical gradient of the terrain (derived from NASA SRTM data, available through [14], for regions corresponding to the real world). This feature reflects the physical difficulty of movement, on the assumption that low gradients (level ground) are easier to traverse than steep ones.

Goal Gradients: SCAMP creates a gradient map leading to each region, which takes the place of the urgency features, guiding agents toward the appropriate goal for the event in which they are currently participating.

Exclusion flags: Like events, tiles can exclude agents of specific groups over specific periods of time.

2.2.7 Spatial regions

Spatial regions capture semantics such as countries, bodies of water, cities, topographical features other than physical gradient (e.g., agricultural land, badlands), and other information that might either be a destination for some event or impact the speed of movement of agents through geospace.

Falloff: SCAMP creates a gradient for each region, so that agents can orient themselves to it in making movement decisions. Falloff, in [0, 1], determines the distance from which the region is visible. A region with a falloff of 1 is visible anywhere in the map, while a falloff of 0 indicates that the gradient disappears as soon as one is outside of the regions.

DelayDiv: Each region can affect the speed of agent movement through it, and that speed may vary by group. This value divides the group's default geospatial delay, so that values < 1 increases the delay while values > 1 indicate that group members can move faster than their default.

2.3 Temporal and Spatial Resolution and Extent

The user can define the size of a spatial tile or the duration of a unit of processing. In our model for Ground Truth, spatial tiles were 20km across, and the entire terrain was 20 x 20 tiles, or 400 x 400 km. The duration of one unit of processing was a day, and we ran models for as long as 2000 days.

3 Process overview and scheduling

SCAMP is implemented in the Repast environment [1]. This environment advances time discretely, by ticks. Each tick, two methods are invoked: first runAvatars(), then (at lower priority) runGhosts().

Not all avatars run at each tick. Each avatar maintains its local time, the sum of the event delays it has experienced. Each time runAvatars() is called, the list of avatars is sorted, and only those with the lowest time are executed, so that avatars cannot run into the future with respect to one another.

The avatars' time variables reflect domain time, not ticks. It typically requires several ticks for an avatar to advance one increment in domain time. Domain time has the same units as event transit time; in our Ground Truth model, it is measured in days.

An avatar's basic execution cycle has two parts: send out ghosts to plan ahead, then move.

1. First it sends out the number of shifts of ghosts appropriate for its group, instantiating the ghost agents and adding them its context, where runGhosts() will find and run them. Each shift ends when all of its ghosts have run for the number of look-ahead steps called for by the group.
2. When the required shifts are complete, the avatar then builds a roulette based on the presence features laid down by the ghosts, and makes its choice.

Similarly, runGhosts() identifies the ghosts that are eligible to run, and lets them run for one step. If the group calls for more than one step look-ahead, the ghost will wait until the next tick to take the next step.

All other processes are driven by these two cycles. In particular:

- Each unit of domain time that an avatar is participating in an event or present on a tile, it triggers any group change rules associated with that event or tile, and updates its relationships with other agents that it meets there.
- A group change rule triggered by one avatar may impact other avatars, changing their location in event space or geospace, or even bringing them into being or killing them.
- Each time an avatar moves to a new event, it updates its wellbeing vector according to the exogenous features of the event.
- As its wellbeing preferences adapt, the avatar may change its affiliation with other groups.
- Each time a ghost queries an event feature vector to make a choice, the HGNs are invoked to update the urgency features, and the presence features are evaporated.

4 Design Concepts

4.1 Basic principles

SCAMP is a stigmergic model, which means that agents interact, not directly, but by making and sensing changes to a shared environment. Both the nature of interaction and the structure of the environment are central to SCAMP.

While agents can detect other agents with whom they are collocated on events or geospatial tiles, the do not exchange messages with them, and most of their interactions are by way of the presence features that agents deposit on locations, and the urgency features that the HGNs compute from them.

In spite of the very simple nature of stigmergic reasoning, SCAMP is able to produce realistic social data by taking advantage of Simon's observation that behavioral complexity may emerge from simple agents if they move in complex environments [12]. SCAMP's models of social behavior are all external to the agents, including the CEG, the HGN, and the geospatial model, and in Ground Truth were constructed by analysts with no knowledge of programming [6].

4.2 Emergence

Because of the underlying stigmergic mechanisms, SCAMP agents are embedded in multiple feedback loops with other agents, and virtually every observable feature of their behavior (Section 4.11) is emergent.

4.3 Adaptation

Individual agents adapt based on the exogenous features of the events they encounter (which modulate their wellbeing) and the other agents they encounter (through modification of preferences). In addition, by sensing the state of the external environment, they constantly adapt their behavior in light of the behavior of other agents who are making changes in that environment.

4.4 Objectives

Agents are not designed to optimize a single objective, but their behavior has the effect of pursuing several things concurrently:

- Changes in their wellbeing lead them to prefer events that are most helpful given their current state.
- In responding to urgency features on events, they seek to impact the satisfaction, not only of their own group, but of others, since the feature vector includes urgency for all groups.
- In geospace, they seek to reach a specified destination in a way that is consistent with their terrain and their attraction to or repulsion from other agents.
- If their encounters with other agents show that they are in fact behaving more like a group other than their own, they seek out events with the potential to change their group membership.

4.5 Learning

Learning is both internal and external to agents.

Internally, avatars update their preferences based on their wellbeing, which in turn depends on the events in which they participate. They also update their relations with other agents.

Externally, the environment (the CEG and the geospatial grid) learn the distribution of agents, and make that information available to agents who move over them.

4.6 Prediction

Avatars predict the near-term future by the ghosts that they send out, an emulation of the simulation heuristic documented by Kahneman and Tversky [5].

4.7 Sensing

Agents have three sources of information about their environment: the baseline preferences of the different groups in the simulation, the features of the events or locations that are accessible to them as they make decisions (Section 2.2.1), and the preference vectors of those agents whom they encounter by participating in the same event at the same time or by being on the same tile in geospace at the same time. We do not model any specific mechanism for the transfer of this information; agents simply have access to it.

4.8 Interaction

In keeping with SCAMP's stigmergic architecture (Section 4.1), virtually all agent interactions are indirect, through the environment, either by sensing the collocation of specific agents or by monitoring the urgency and presence features in the environment, which in turn are modulated by the movement of other agents. The only information that agents communicate directly with one another is their preference vectors, and those are shared only with agents that they meet.

4.9 Stochasticity

Stochasticity is pervasive in SCAMP. The preference vectors of individual agents are sampled stochastically from baseline preferences of their groups, and each agent decision is based on roulette selection among alternatives (though a system-wide determinism parameter can tune the degree of determinism in these choices). The agent choices drive the entire dynamics of the model.

4.10 Collectives

The identity and nature of agent groups, along with their initial members, are defined by the modeler. However, agents constantly adjust their affiliation with groups other than their own, and can change groups based on the other agents that they encounter.

4.11 Observation

SCAMP is designed to run in batch mode, logging a wide range of features of interest. These include (but are not limited to) the history of agent participation in different events and the tracks agents follow in geospace, their affiliation with different groups, their changing preference vectors and the changing feature vectors of events, the meetings among different agents, group changes that occur, the satisfaction of the root goal in each HGN over time, and the entropy of presence feature strength per group and overall through time.

Logs are in csv format, and we post-process them with a variety of Python scripts to generate yet further information about a run. Chapter 4 of the SCAMP User Manual [9] describes in detail the contents and analysis of the SCAMP logs.

In the Ground Truth program, we used these logs to respond to specific questions from the social scientists, and were able to respond to a large range of queries that they addressed to our agents, including, but not limited to,

- What were you doing on a given date?

- What was the last thing you were doing before your present activity?
- What other options did you consider at that time?
- What influenced your choice of this option?
- What options are you considering next, and how would you prioritize them?
- Whom have you met recently?
- How strong is your relation to them?
- How satisfied are you with your achievement of your objectives?
- How happy are you about your current condition (economic, physical, psychological)?
- How sympathetic are you to a specific group (e.g., the government)?

5 Initialization

SCAMP is not a model of a specific scenario, but a modeling framework, and the structure of the model and agent population is highly configurable. Analysts construct the models using CMapTools [4] for the CEG and HGNs and GIMP [13] for the geospatial model, supplemented by an Excel workbook to define the various parameters outlined above. Chapters 2 and 3 of the SCAMP User Manual [9] describe in detail the preparation of these files.

6 Input data

The present version of SCAMP directly uses NASA SRTM data to generate a realistic terrain, and the analysts who develop the models draw on published information relevant to the scenario we constructed. In the future, we expect to couple SCAMP's exogenous features to external data sources, but in the current version of the code, all dynamics are generated by the model itself.

7 Submodels

SCAMP has four perspectives, which may be considered submodels:

- The Causal Event Graph (CEG), which models possible narratives that an agent may experience;
- Geospace, which models the spatial environment of the agent;
- Hierarchical Goal Networks, which model the agent's strategic goals;
- Each agent's social network, maintained dynamically as the agent meets other agents.

But these do not run as separate processes. Except for the social network, which emerges dynamically as agents move, these models are static structures, whose parameters are described in detail in Chapter 2 of the SCAMP User Guide [9].

SCAMP's stigmergic structure allows these models to be constructed by domain experts rather than programmers. In the scenario we constructed for Ground Truth, we modeled a fictitious country, but were inspired by the ongoing conflict in Syria, and parameterized our models accordingly. For example, our geospatial map is roughly the same area as Syria, as were numerous geospatial features (such as the number of major roads, the amount of agricultural land, the areas of refugee and IDP camps, the military airbase, the de-escalation zone, and the oil fields. The movement delays for different groups were derived from timelines of the conflicts in Syria and Iraq.

Acknowledgements

The development of SCAMP was funded by the Defense Advanced Research Projects Agency (DARPA), under Cooperative Agreement HR00111820003. The content of this paper does not necessarily reflect the position or the policy of the US Government, and no official endorsement should be inferred.

References

References by Parunak are available at https://abcresearch.org/abc/papers, using a browser with Javascript enabled.

[1]Argonne National Laboratory. Repast Agent Simulation Toolkit. Argonne National Laboratory, 2007. http://repast.sourceforge.net/.

[2]V. Grimm, U. Berger, F. Bastiansen, et al. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, 198:115–126, 2006.

[3]V. Grimm, U. Berger, D.L. DeAngelis, et al. The ODD protocol: A review and first update. *Ecological Modelling*, 221(23):2760–2768, 2010.

[4]IHMC. IHMC CmapTools - Download. Pensacola, FL, 2013. http://cmap.ihmc.us/download/.

[5]D. Kahneman, A. Tversky. The Simulation Heuristic. In D. Kahneman, P. Slovic, and A. Tversky, Editors, *Judgment under Uncertainty: Heuristics and Biases*, 201-208. Cambridge University Press, Cambridge, UK, 1982.

[6]H.V.D. Parunak. Psychology from Stigmergy. In *Proc. Computational Social Science (CSS 2020)*, CSSSA, 2020.

[7]H.V.D. Parunak, T. Belding, R. Bisson, et al. Stigmergic Modeling of Hierarchical Task Networks. In *Proc. the Tenth International Workshop on Multi-Agent-Based Simulation (MABS 2009, at AAMAS 2009)*, 98-109, Springer, 2009.

[8]H.V.D. Parunak, S. Brueckner. Concurrent Modeling of Alternative Worlds with Polyagents. In, *Multi-Agent-Based Simulation VII*, *LNCS*, 128-141. Springer, Berlin, Germany, 2007.

[9]H.V.D. Parunak, L. Sappelsa, J.A. Morell, et al. SCAMP User Manual. Wright State Research Institute, Beavercreek, OH, 2020.

[10]L. Sappelsa, H.V.D. Parunak, S. Brueckner. The Generic Narrative Space Model as an Intelligence Analysis Tool. *American Intelligence Journal*, 31(2):69-78, 2014.

[11]V. Shivashankar. *Hierarchical Goal Networks: Formalisms and Algorithms for Planning and Acting*. Ph.D. Thesis at University of Maryland, Department of Computer Science, 2015.

[12]H.A. Simon. *The Sciences of the Artificial*. Cambridge, MA, MIT Press, 1969.

[13]The GIMP Team. GIMP: GNU Image Manipulation Program. 2020. www.gimp.org.

[14]USGS. Earth Explorer. 2020. https://earthexplorer.usgs.gov.