Modeling Multi-Agent Chaos: Killing Aliens and Managing Difficult People

W. Korey MacDougall (warrenmacdougall@carleton.ca)
Robert L. West (robert_west@carleton.ca)
Emmanuelle Hancock (emme.hancock@gmail.com)
Institute of Cognitive Science, Carleton University, Ottawa, Ontario, Canada

Abstract
Understanding macro cognition is important for understanding how experts and lay people function in the real world, and for building safer, more effective socio-technical systems. We present a framework and a methodology for creating and evaluating process models of highly dynamic expert tasks and illustrate it with two models.

Keywords: Macro Cognition; Cognitive Modeling; Cognitive Architectures; Expertise; Videogame playing; Mediation;

Introduction
Collectively we have developed innumerable forms of expertise, differing wildly in their appearances. These range from trapping animals to space flight, poetry to plumbing. Beneath the extreme diversity of forms, however, there may be consistency in the structure and acquisition of expertise. It may be that expertise (and our capacity to develop it) is rather simple and straightforward, but that its manifestations are so varied because it emerges in many vastly different domains. This is akin to the point made by Herbert Simon’s famous “ant on the beach” metaphor (Simon, 1969), in which he argues that the apparent complexity of an ant’s behavior as it moves in a convoluted path across a beach is largely attributable to the complexity of the environment, and not to any sophisticated scheming or strategizing by the ant. We think it is worth investigating whether the situation is similar regarding human expertise, but feel that the techniques and concepts which would permit such a study are in need of further refinement. We offer here a prospective research methodology that we are developing to address this using the SGOMS macro architecture (West & Pronovost, 2009; West et al., 2013), and describe two models which we have constructed to test the method. The first model is of people playing a video game, “Gears of War 3” (produced by Epic Games) and the second is of professional mediators.

Studying Experts
Expertise is a topic of theoretical and practical importance in a number of fields, including psychology, sociology, artificial intelligence, and education. Its relevance to multiple communities has led to the development of varied conceptual frameworks that distinguish experts from non-experts according to mental capacities, experience over time, representation and organization of knowledge, elite achievement, status within a community, or reliably superior performance (Ericsson, 2006). From these conceptual frameworks have developed a number of methods for studying experts, including, inter alia, laboratory based methods (Chi, 2006), naturalistic observation (Ball & Ormerod, 2000), the engineering of expert systems (Fink & Lush, 1987), and simulation of expert behavior and cognition (West & Somers, 2011).

Existing conceptions of expertise, and their associated research methodologies, have provided important insights into many aspects of expert performance, such as differences in knowledge representations between experts and novices (French et al., 1996), the psychological traits most frequently associated with the development of expertise (Shanteau, 1998), the importance of “deliberate practice” in acquiring mastery, (Ericsson, 2006), and the role of social factors and institutions in facilitating the development of expertise (Hunt, 2006). Additionally, AI-based research into expertise has led to the creation of a number of expert systems capable of supplementing or replacing the performance of human experts. Examples of these include systems for medical diagnosis (Saito & Nakano, 1988), hypothesis formation (Buchanan, Sutherland, & Feigenbaum, 1968), and chess analysis (Michie, 1972).

There are, however, significant aspects of expertise which are not as amenable to investigation using existing frameworks. Specifically, these methods do not lead to a process model. That is, a model that, given the current state of the agent, the task, and the environment, can predict what the agent will do next. We are most interested in applying this approach to: expert performance in complex, chaotic, and real-world environments; the cooperation and coordination of multiple experts (i.e., teamwork); and consistencies in the cognitive activities and aptitudes of experts in different domains.

Macro Architecture Hypothesis
Cognition can be divided into micro and macro cognition (Klein et al, 2003). Micro-cognition refers to the mental activities studied in traditional cognitive psychology experiments, whereas macro-cognition refers to the forms of cognition that underlie functioning in complex, real-world tasks (West et al., 2013; Klein 2003). Concerns have been raised over whether the theories and methods embodied in micro-cognitive experimentation and modeling can be reasonably scaled up out of the lab and applied to the study of real-world cognition (Klein, 2003). One avenue by which this has been addressed is the development of unified cognitive architectures, such as ACT-R (Anderson, 1996), EPIC (Kieras & Meyer, 1997), and SOAR (Laird, Newell, 2003)


& Rosenbloom, 1987), which constrain models according to micro-cognitive principles (e.g., speed of recall, memory structures, modularity, parallel processing, etc). These architectures have been used to model macro cognition, but the architectures “under constrain” the models. In other words, within the same architecture, multiple models can be built of a given task, and it can be difficult or impossible to determine which model is more psychologically plausible.

Our solution has been to develop the macro architecture hypothesis. We argue that there exists a macro level architecture built atop a micro level architecture, just as micro architectures are built atop a neural architecture (see Figure 1, below). This hypothesis then makes a claim about how micro cognitive models ought to be organized at a high level.

![Figure 1. Relation between micro cognitive and macro cognitive architectures](image)

**SGOMS**

We are using the SGOMS macro architecture (West & Pronovost, 2009: West et al, 2013), which is an extension of the GOMS (Goals, Operators, Methods, Selection rules) framework (Card, Moran, & Newell, 1983). SGOMS stands for socio-technical GOMS. GOMS models have been used successfully in a variety of applications, such as the development of telephone operator workstations (Gray, John, Atwood, 1993) and computer interfaces (John & Kieras, 1996). This class of models often fails, however, when applied to chaotic environments which involve multiple actors, imperfect information, and frequent interruptions (West & Nagy, 2007). SGOMS addresses this by extending the GOMS framework in two ways: to redefine unit tasks to include the condition that they are unlikely to be interrupted, and to introduce an additional cognitive control structure called a “planning unit”.

Unit tasks represent a way of partitioning the problem space into sub-goals in a way that will facilitate performance. In the GOMS framework, unit tasks are structured such that the agent can minimize downtime and avoid overload (Card, Moran, & Newell, 1980), both of which negatively impact performance. To this the SGOMS theory adds that unit tasks should be unlikely to be interrupted. Thus unit tasks become islands of predictability, with shorter unit tasks in more chaotic environments.

The second modification to GOMS is the introduction of planning units. A planning unit is a high level goal structure and set of associated contextual constraints that determine when the planning unit is appropriate. Planning units are typically structured as a list of unit tasks. As the name implies, planning units provide the units that are manipulated to create a plan, and they serve several important functions. First, by providing a common name for distinctly different modes of action they allow multiple agents to coordinate through the establishment of common ground (Klein et al, 2004; West et al., 2013). This reflects the fact that experts often have a specialized, shared vocabulary. However, because planning units are associated with constraints that specify when they can and cannot be used, they also allow individuals to make adjustments without consulting team members. In other words, planning units help individuals and teams balance the costs and benefits of reacting locally and individually versus globally and collectively. Second, planning units allow agents to deal with interruptions, by providing a structure that is resistant to failing when unexpected conditions arise. Unlike unit tasks, planning units are designed to be interrupted and restarted later. Third, planning units allow us to model complex tasks that do not have a single, optimal strategy or solution. In macro cognitive domains, there are often multiple, equally good strategies in a given situation. A model that predicts a single, specific response at each choice point will not work. In contrast, planning units can be rearranged to produce different strategies and/or react to unexpected events.

We are now working to further refine both the theory of SGOMS and the methods by which we develop SGOMS models. Below is an overview of the method we have developed to create these models, and two models that we have created using this approach. The first model is of playing the video game “Gears of War 3”, and the second is of professional mediators. The purpose of juxtaposing these two very different models is that we are investigating whether the SGOMS framework can be used to develop models across a broad range of tasks. If we can successfully model both tasks using the same architecture, we take this as evidence in favor of the macro architecture hypothesis, and of the principles of constraint built into SGOMS.

**Method**

We have two objectives in developing the methodology presented here. First, we want to be able to model experts in highly complex environments. Second, we wish to develop process models that can aid in the design of real-world, socio-technical systems. These two objectives are explored further below.
The first objective is related to our attempt to bridge micro and macro architectures. As mentioned above, cognitive psychology generally addresses micro cognitive functions (e.g., memory retrieval, task switching, attention). These are typically investigated by creating simplified tasks and environments, which are meant to isolate cognitive mechanisms or the variables that affect those mechanisms. Data is collected by running multiple participants who each perform many instances of the task(s), and statistical analyses (most commonly t-tests) are then applied to the collected data. This allows investigators to assess whether a particular condition or intervention has a reliable impact on some performance measure, such as reaction time or accuracy. All of this is well known; it is standard protocol in cognitive psychology experiments, and it is a robust, generative paradigm with countless applications. Of particular relevance to the present work, the collection of such data is useful in informing the creation and testing of micro cognitive models.

The above approach is of limited applicability, however, in studying macro cognition or experts in situ. The power of the method is precisely in its elimination of situational variability and complexity, through simplification, repetition, and averaging. But expert behavior is dynamic, adaptive, and often idiosyncratic, and these features are tightly linked to the complexities of the environments in which experts operate. Micro cognitive methods often obscure these important parts of the equation, particularly the integrated nature of expert cognition, the ways in which experts coordinate their efforts, and the importance of adaptation to the environment. We argue, therefore, that in the study of experts generally, and in the creation of macro cognitive models specifically, we ought to be capable of addressing expert behavior as it occurs in complex environments.

The second objective runs in the opposite direction. There has been excellent work studying experts in chaotic or multi-agent contexts. This branch of research employs a different set of methods, such as naturalistic observation (McAndrews, Banks, Gore, 2008), think aloud protocols (Ericsson & Simon, 1998), cognitive narratology (MacDougall & West, forthcoming), and iterative system design (more on this below). These techniques can provide useful information about how experts and teams function in the real world, but they generally lack the step-wise precision of techniques used in micro-cognitive research, and thus do not lend themselves easily to the creation of process models. The development of process models is a key component in using cognitive modeling theories and techniques to design safer, more efficient socio-technical systems, and thus we think that the information gleaned from this body of techniques ought to be integrated with the techniques of micro cognitive modeling.

Given these goals, we are working with a methodology that is more akin to systems engineering than to standard cognitive psychology experimentation. It is an iterative design approach, whereby we construct a model, test it against human performance using a model tracing technique (Anderson, Kushnerick, & Lebiere, 1993), and make changes to incorporate anything we may have initially missed. We repeat this process until the model is consistently capable of predicting what a human performer is likely to do next, given a particular situation. Because of the complexity of the contexts in which we are developing these models, we rarely see situations that will have one and only one possible next action. In both of the models presented here, the number of all possible actions in any given scenario is enormous, and it is often the case that there is not a single, optimal next action. This forces us to be flexible in our evaluation of whether the model is truly accurate in predicting behavior. We have adopted the stance that in order to be counted as a correct prediction, the action predicted by the model must be “reasonable”, given the context. By reasonable we mean that given what the agent (human or model) knows about its context and goals, its choice of action is coherent and is consistent with a specific planning unit. This reduces the solution space from infinitely many options to a handful of “reasonable” ones. This approach is similar to that often used in evaluations of the output of models in linguistics and in vision science (i.e., the sentence or the visual output is judged based on our own sense of what is right and what is reasonable).

We anticipate two primary objections to this method. The first is that our criterion of reasonableness is not rigid, and may lead to lax evaluation of a model’s accuracy. The second is that the iterative design process allows us to continue adding to the model until we are satisfied with its performance. These problems are those of over fitting and unconstrained expansion. We acknowledge that these issues must be addressed, but we feel that they arise from a particular theoretical and methodological orientation toward modeling, and are not weaknesses of method per se. More precisely, we are not interested so much in whether a model, at some particular stage of development, can predict an agent’s next move in some percentage of cases, in order for it to qualify as a valid model. We are more interested in exploring whether, through a process of iterative development, we can use the macro architecture to create models, across a range of domains, that exhibit high fidelity to human performance and can usefully inform system design.

This orientation is based on two ideas. The first is that evaluation of a theory or method requires identification of a specific “unit of analysis.” In Experimental Psychology the unit of analysis is the individual paper. A hypothesis must be supported within a paper in order to be accepted. In contrast, the unit of analysis for cognitive architectures is the results across all papers that pertain to the architecture. Thus, SGOMS is expected to apply across all forms of expertise practiced in real world settings, where interruptions and unexpected events are common. We must look at the framework as a whole, and the complete set of results (models) produced by its use, to determine its validity. The second idea is that if a theory or method
remains generative, or continues to inspire progress in some area, then it is valuable. This draws on the Lakatosian model of scientific evaluation (Lakatos, 1975; Cooper, 2007), and stands in contradistinction to the Popperian notion of falsification that underlies standard cognitive psychology research (Cooper, 2007).

Models
The first model presented here is of players attempting “Horde Mode” in Gears of War 3, and the second is of professional mediators in a staged conflict mediation scenario. The construction of these two models involved multiple iterations of development and testing. This consisted of a sequence consisting of preliminary archival research (manuals), development of pen-and-paper models, recording and analyzing video of experts performing, and subsequent modifications to the models. This cycle was repeated until the models were capable of consistently predicting what the experts’ might reasonably do next.

Video Game – Gears of War 3
Gears of War 3 is a 3rd-person shooter game, in which the player controls a soldier character and must defeat alien enemies in various 3d environments. The environments (or maps) are large and complex, and the activity is highly chaotic. The gameplay involves a variety of actions, such as running, sliding, taking cover behind objects, firing guns, reloading weapons, throwing grenades, finding ammunition, purchasing structural upgrades, controlling turrets, and reviving fallen teammates. Successful play requires excellent perceptual-motor coordination, effective strategic decisions, and (in team-play modes) regular communication.

The first step in creating this model was to determine the unit tasks, methods and operators used in the game. The operators comprise the full set of discrete, low-level actions that can be performed by a player, such as moving, looking around, firing, reloading, etc. These map closely on to the control scheme of the game and the functions of the controller buttons (we used the version of the game for the Microsoft XBOX 360). The methods are combinations and sequences of operators that are executed routinely (e.g., scan heads up display, pick enemy from group, take cover, retreat, aim and fire, quick reload). Unit tasks are the next level of structure, and are constructed atop methods. These involve a well defined sub-goal and a set of actions (methods) that can be taken to accomplish the goal. Example unit tasks are: find safety (methods: retreat, fire, take cover); set up crossfire (methods: find position, take cover, aim, engage); collect ammunition (methods: move to supply, cover fire, retrieve). As mentioned above, unit tasks are meant to be islands of work that avoid downtime, prevent overload, and avoid interruption. Given the chaotic nature of the game, the unit tasks in this model are completed very quickly (within the space of a few seconds), unless it is between stages (when no enemies are present).

The next step was to determine the planning units that players were using. These are high level strategic plans, and include such things as find defensible position, hold ground, replenish supplies, repair stronghold, or revive teammate. Planning units include a goal and contextual factors (or constraints) that dictate when the planning unit is appropriate.

We chose to refine and clarify the methodology using “Gears of War 3” for two principal reasons. First, it provides a strong test of the validity and usefulness of the planning unit construct. The chaotic gameplay entails frequent interruptions which force the player to modify their strategy, and we treat such strategic changes as shifts in the current planning unit. Additionally, although the main objective is quite straightforward (eliminate all enemies), the ways in which this can be accomplished differ enormously. Players can choose to set up a stronghold and defend it, use stealthy maneuvering to flank enemies, move in circuits to guide the enemies into the open, lead the enemies into chokepoints, etc. By developing the model with six different human players, we found significant variety in play styles and strategies. After developing the model through several iterations, we observed two principal strategies (hunter/aggressor and defender) which differed according to which planning units were favored. Because of these two factors, the frequent interruptions and the potential variability in strategy, rigid models of this task will consistently fail, or will exhibit artificially uniform behavior. For example, a standard GOMS model can describe any of the individual units, but cannot predict or replicate the richness and fluidity of the behaviors that the human players exhibited across an entire game. The introduction of planning units to the model was necessary to capture this richness.

Second, the co-operative two player mode also allowed us to examine the usefulness of the macro architecture in modeling team play and communication. These are important macro level activities that have received considerable attention (Klein et al, 2003), and we have examined them in (West et al., 2013). When playing cooperatively, players established common ground by verbally agreeing on which planning units to use. By agreeing on a common planning unit, players were able to coordinate and communicate more quickly and effectively. Interestingly, information that was relevant to planning unit choice (e.g., hold ground) was conveyed differently than information that was relevant to unit task choice (e.g., retreat). When switching planning units, players would discuss and negotiate, sometimes at length, but when both players were within the same planning unit the conversation consisted mostly of short bursts of speech, often a single word. We believe that this highly efficient form of communication was possible because of the common ground established by being in the same planning unit. One of the predictions of the SGOMS framework is that choosing a shared planning unit saves time and effort, and
we are working to further investigate the degree to which human players are aware of this cost-benefit trade-off.

**Mediation**

Mediation is a complex macro cognitive activity. It requires one to be sensitive to other parties’ motivations and cognitive styles, to understand how these might differ and clash, to construct a narrative that runs from past to present (on various time scales), and to have a strategy for reaching a resolution that satisfies all involved parties. We chose to model professional insight mediation, which is a form of mediation that attempts to help the involved parties to gain a deeper insight of the problem or situation, and thereby to resolve the problem (Melchin & Picard, 2008). The construction of the model involved the same iterative process as that presented above for the video game. We began constructing the model with textbooks detailing insight mediation techniques, and further refined the model using video recordings of professional mediators acting out a mock conflict resolution scenario. Limitations of space prevent us from describing the model in full (see West el. al., 2013, for more detail), but we were again able to iterate to a point where we could reasonably predict all actions.

**Juxtaposing different models**

There are two reasons we have presented these models together. The first is that, while the two tasks appear very different, we were able to use the same framework (the SGOMS macro architecture) to successfully model both. This is important, as it is central to the macro architecture hypothesis that the same system can be used to inform model creation in different domains. The second reason we are presenting these two models together is as evidence that the methodology we are developing is not useful only in a narrow set of tasks. The process of iterative model design, constrained by the criterion of reasonableness, has allowed us to develop models of two very different, complex tasks. The models are extremely robust in the face of interruptions and general chaos, and have allowed us to address important questions concerning how to model multi-agent activity. Concerning the latter point, we hope to further develop our approach to test ideas about distributed cognition as it occurs in socio-technical systems.

**Future Directions**

There are a number of remaining avenues that we wish to pursue using the SGOMS framework and the methodology presented here. One goal is to replace our paper and pencil representations of SGOMS models with a high level computer language. Currently, we have SGOMS implemented in Python ACT-R (Stewart & West, 2005). This is to demonstrate that SGOMS is cognitively plausible and also to provide a template for building SGOMS models in ACT-R. Furthermore, we are currently developing a framework for creating standardized, high-level representations of SGOMS models, which can then be directly compiled into Python ACT-R code (i.e., into computational process models). This step is necessary to communicate more clearly the structure of our models, and to allow other investigators to reproduce and run the models if desired.

Another project that has developed out of this work is the investigation of the importance of narrative in the development, exercise, and transmission of expertise. This concerns the ability of actors to understand and anticipate sequences of events, to make sense of complex contexts, and to understand other actors according to their role in these contexts. Work in cognitive narratology (Herman, 2007) has indicated the importance of constructing narratives for making sense of the world and functioning in it, and we believe this line of thought can be profitably applied to the study of experts.

**Conclusion**

We began by drawing a parallel between Simon’s (1969) fable about an ant on the beach and human expertise across different domains. An ant’s convoluted movements across a beach can be interpreted as chaotic and random, or else highly sophisticated and subject to some arcane set of heuristics. But these are both illusions. The complex path arises simply because the ant avoids climbing hills. In the mediation example and the video game example, our models indicate that dissimilar trajectories through the task can be accounted for by differences encountered during the task (i.e., from other agents) and from different high level strategies (e.g., aggressive versus defensive). However, the experts’ knowledge and the architecture (SGOMS) for systematically applying that knowledge, given the contextual constraints, were identical across domains and experts. Moreover, when dealing with different types of expertise, the only thing that needed to be changed was the knowledge content and the high level strategies. Overall, this suggests there is some level of simplicity or consistency underneath the diversity of forms.

A survey of the literature concerning human expertise reveals an enormous diversity of theories and methods, with particular combinations of these often being tightly tied to the domain which is being studied. While this has been a fruitful approach in many ways, our results suggest there may be aspects of expertise that are common across all forms of real world expertise. SGOMS provides a single unified way to potentially understand factors such as high level strategizing, co-operation, stylistic differences between experts, similarities between experts in different fields, potential simplicity underneath apparent complexity, and the processes involved in adapting to complex environments.

**References**


