# Realistic Agent Populations for Large-Scale Virtual Training Environments


**Art Pope, Ph.D.**
**SET Corporation**
**1005 North Glebe Road, 4th Floor**
**Arlington VA 22201**
**apope@setcorp.com**

**Peter Selfridge, Ph.D.**
**SET Corporation**
**1005 North Glebe Road, 4th Floor**
**Arlington VA 22201**
**pselfridge@setcorp.com**

**Col. John R. Surdu, Ph.D.**
**DARPA IPTO**
**3701 N. Fairfax Dr.**
**Arlington VA 22203**
**John.Surdu@darpa.mil**



**Keywords:** Agent-based simulation, background population, crowd simulation, human behavior representation, human terrain



**Abstract**

This paper explores the problem of effectively simulating realistic populations of background agents in large-scale virtual environments for training and mission rehearsal. It explains why such populations are needed, and surveys the behaviors one would want to see them exhibit. It argues that a successful simulation of a background population ought to be credible, scalable and maintainable, while defining what is meant by those terms. It identifies the key technical challenges as being the efficiency of behavior authoring, and the efficiency of simulation. Finally, it describes some ways such a background population simulation can be verified and validated.


## 1. INTRODUCTION

Virtual environments generated by computer graphics technology have been used for many years to train military personnel, catalyzed in part by the SIMNET program sponsored by DARPA [i]. More recently, inspired partly by the rise of computer gaming, new applications have been developed that include training for squad tactics [ii], cultural dialog with non-combatants [iii], and multiple uses [iv]. The next generation of these applications will provide more complex virtual environments in which to train warfighters for urban operations, low-intensity conflict and asymmetric warfare.

In these environments, *realism* will depend not just on higher polygon counts, faster networks, and better physics simulations, but also on the incorporation of realistic *human terrain*: the background population of citizens who go to work, take their kids to day care, shop at markets, do chores, meet friends for lunch, walk their dogs, etc. Members of this population exhibit complex, time-dependent behavior based on their individual beliefs and intentions, their relationships and interactions with other individuals, and their perceptions of the dynamic, physical environment. Importantly, their behavior is also influenced by customs and norms that depend on both region and culture.

In this paper, we explore the problem of effectively simulating realistic background agent populations in large-scale virtual training and mission rehearsal environments. We explain why such populations are needed, and briefly catalog the behaviors one would want to see them exhibit. We then identify key qualities that simulations of background populations ought to have, and the key technical challenges that must be met in order to achieve those qualities.

## 2. BACKGROUND POPULATIONS

To achieve their potential as flexible and powerful military training experiences, virtual environments must portray the physical terrain of interest in a realistic fashion, be it so called “geo-typical” (e.g. a generic Middle Eastern city) or “geo-specific” (e.g. a specific block of Sadr City in Baghdad). Technologies to create typical urban areas or specific physical terrain exist today (see, for example, any number of military-style first-person-shooter games for generic terrain, and “Kuma\War” [v] for specific terrain created by hand). These games also illustrate adequate physics, and effects like smoke, explosions, and weather.

However, what is missing is the simulation of realistically large background populations. Instead, small numbers of non-player characters are simulated using simple scripts. But realistic background populations are *essential* to training realism. The background population provides the key to learning the sense of “normality” that allows one to pick up subtle social indicators crucial to tactical decisions and long-term intelligence. For example, population animosity may be gleaned from glances, facial expressions, and body language; lack of children on a normally busy street; and unusual vehicle behavior. All of these are important indicators for a patrolling warfighter or his or her commander needing a sense of the area of operation.

Any discussion of simulating a background population quickly comes to the question of what behaviors that population ought to exhibit. There is no universal answer, of course, for the simulation’s purpose must determine what behaviors (and other phenomena) are relevant. An environment for training helicopter pilots might avoid any

**Figure 1.** A virtual environment's background population should exhibit a range of behaviors that is normal and appropriate for the modeled region and culture.

need for the simulated population to engage in verbal communication, for example, whereas one used to train civil affairs officers will find it essential. Nevertheless, even in a general discussion such as this, it is helpful to have in mind some specific examples of the types of behaviors that need to be simulated. What follows is a quick survey.

One broad class of potentially relevant behaviors includes those involving various forms of travel. Individuals travel to get from one point to another (e.g., from home to work), to move toward some attraction (e.g., a street performance), or move away from something distasteful or dangerous (e.g., combat or an explosion). They may do so at various rates (e.g., ambling or sprinting), by various conveyances (foot, bicycle, taxi, bus), and while either following established routes (sidewalks, roads) or not. En-route, they may deviate to avoid obstacles, queue behind others, respond to traffic signals, pass through checkpoints, and embark or disembark a vehicle or animal; as pedestrians they may be herding or leading domestic animals.

Communication, another class of behaviors, includes not only reciprocal conversation, or dialog, but also such acts as demonstrating, threatening, applauding, jeering, yelling, chanting, singing and praying. Individuals communicate through a coordinated combination of verbal utterances, facial expressions, gestures, and other body language. A virtual environment may be required to represent not just the dialog among simulated agents, but also that among agents and human players.

Behaviors are often collective, and almost always at least influenced by constellations of other individuals, both proximal and distant. Many behaviors, including the forms of travel already mentioned, can be performed either individually or in small groups of family members, companions or colleagues. Group members inherit aspects of the group's common purpose and affect, yet exhibit varying behaviors depending on their roles within the group (such as one finds on any family stroll). Even among strangers in a crowd, emotional state propagates among individuals, affecting the behavior of others.

Recreation and commerce are the theme of a diverse class of behaviors. These include loitering, observing a performance, browsing, shopping, looting, passing an object from one individual to another, begging, eating a picnic or restaurant meal, playing a game, and gambling. A final class of behaviors, those involved in combat, includes threatening and attacking with firearms and projectiles, using positions of concealment and cover (including bystanders), and reconnoitering and retreating.

As this survey illustrates, the variety of behavior one might expect to observe in any populous area is rather large. Yet this variety is made even larger by additional dimensions of regional and cultural specificity. For example, different cultures inculcate different preferences for which side to step to when passing someone, whether to queue up when awaiting a bus, and how close one should stand to a stranger. Culture determines daily schedules of work, rest, and other activities, how long you exchange pleasantries before getting down to business, the meaning of a gesture like "thumbs up", and many other specific aspects of behavior.

## 3. KEY QUALITIES

There are certain performance qualities that technology for simulating background populations must achieve in order to be both useful and economical for military training and mission rehearsal applications. We shall now describe the three we consider most important: *credibility*, *scalability*, and *maintainability*.

### 3.1. Credibility

A virtual environment's simulated background population is *credible* if, to knowledgeable human observers, it behaves correctly in those ways that matter. This is a somewhat lesser demand than that it be realistic, for whereas a realistic simulation must be one that closely approximates objective reality, a credible one need only appear to do so, subjectively, to its human observers.

To be credible, the background population must exhibit a realistic variety of behavior that is consistent with the simulated environment, particularly its specific geography, structures, and culture. Individuals must exhibit evidence of mental state and processes, such as beliefs, desires, intentions and values, so that behaviors are appropriately purposeful (or not). Actions must be suitably paced, with those that ought to be instinctual or automatic proceeding more quickly than those that require deliberation.

Credibility is often particularly evident when simulated individuals are made to respond to novel situations. In new or unexpected situations, humans respond in diverse and complex ways, each according to his or her particular knowledge, goals and perceptions. To be credible, simulated agents must also exhibit an appropriately rich variety of responses.

Because credibility is subjective, what is required to achieve it depends on what is perceptible to observers. For human players who can only view the background population from a considerable distance, it may be sufficient that the population congregates and moves in a credible way. However, when closer observation occurs (or, in some cases, is even possible), additional behaviors must be simulated.

Due to its subjective nature, credibility itself is difficult to evaluate. One might judge the credibility of a simulation with a form of Turing test in which human observers are asked to classify various agents as either human players or simulated members of the background population. As credibility increases, human and simulated agents will become less distinguishable, and the accuracy of such classification will approach the level of chance. In any such test, interactions between the observers and the agents they are to classify would have to be carefully moderated so that they remained within the scope of modeled behaviors (e.g., not allowing observers to question agents on just any topic whatsoever). Also, the human players would have to be given some incentive to maintain their interest in participation, yet kept unaware of the true purpose of that participation so as not to inadvertently skew results. Thus, while perhaps possible, a Turing test of credibility would have to be very carefully designed and administered.

### 3.2. Scalability

A technology for simulating background populations is *scalable* to the extent that it accommodates growth in the size, density and complexity of the simulated population. One would like the technology to be able to provide a small-scale environment on some common computing platform, such as a PC, so that such environments can be made widely available. However, one would also like the technology to be able to sustain larger and more complex environments with a commensurate increase in computing resources. Those additional resources should be ones that are commonly available and can be added incrementally, as nodes are added to a computing cluster or servers, to a network.

An example will serve to illustrate what a large-scale simulation might entail. Suppose one wanted to simulate an urban region of 5 $km^2$ in which 20 human participants could travel the streets over a period of a few hours, not only observing the background population but also able to interact with any simulated individual they encountered. At typical urban densities,[1] a region of that size might contain 50,000 individuals. Fortunately, it may not be necessary to model all individuals, each with a specific state and dynamics. If we suppose that only one in five might actually travel outside during the few hours of the exercise, thus becoming potentially observable by human participants, then only $10^5$ need actually be modeled. However, of these, we might reasonably expect that perhaps as many as 1000 could be observed at any one time by all human participants combined. The simulator would therefore have to be capable of sustaining somewhat detailed modeling (e.g., path-planning, collision detection, and pose generation) for as many as 1000 individuals concurrently. Similarly, we might suppose that as many as 50 simulated individuals could be involved in direct interactions at any one time with the 20 humans; thus the simulation would have to be capable of generating appropriate interactive responses on behalf of 50 individuals concurrently. In summary, a simulator for this virtual environment would simultaneously need to model $10^5$ individuals, generate appropriately detailed paths and poses for 1000, and engage in interactions with human players on behalf of 50.

The scalability of a simulation technology may be characterized by a function (i.e., curve) that relates an amount of processing power to the size of background population that can be simulated with that power, given various fixed conditions. Those fixed conditions should include the complexity of the simulated environment, the density and complexity of the population, the nature of the simulated behaviors, and the manner in which the population is observed. For a highly scalable technology, the function must span a broad range of population sizes with moderate (and perhaps even linear) growth in processing power.

### 3.3. Maintainability

If a simulation of background populations is successful, there will inevitably be considerable interest in adapting, extending, and refining it. There will arise needs to model

[1] Urban population densities of 5,000 to 10,000 individuals / $km^2$ are common. As of 2003, the population density of Baghdad was estimated to be 9,250/$km^2$; that of Mogadishu, to be 12,500/$km^2$ [vi].

additional locales with their associated cultures, correct the way certain behaviors are performed, and add new behaviors and other phenomena. A technology for simulating background populations is *maintainable* to the extent that it facilitates such changes.

Our concern about maintainability regards, in particular, the manner in which behaviors are represented, specified, and performed, for we see these as being among the most difficult aspects of the maintenance problem. The ISO 9126 standard for software quality evaluation [vii] provides a definition of maintainability that applies as well to behavior specifications as it does other forms of software. It defines maintainability as consisting of *stability*, *analyzability*, *changeability* and *testability*. These are precisely the qualities needed to ensure maintainability in behavior specifications.

It must be possible to specify new behaviors and modify existing ones quickly and efficiently (changeability) while avoiding unintended consequences, such as disrupting seemingly unrelated behaviors or degrading simulation performance (stability). Behavior specification must be something that can be done not just by software engineers familiar with the inner workings of the simulation, but by those who have most knowledge of the behaviors to be modeled, such as cultural experts, social and behavioral psychologists, and military personnel with relevant first-hand experience.

To promote maintainability, behavior specifications ought to form a distinct and well-defined part of the simulation system—like the content on a web server or the courseware in a tutoring system. Behaviors should be specified in a way that supports analysis and reuse (analyzability), perhaps by employing certain practices developed to facilitate other forms of software analysis and reuse. For example, formal specifications of a behavior's pre- and post-conditions may permit analyses that identify unused or conflicting behaviors. Factoring out aspects of a behavior that vary across cultures, making them parameters or delegates, may permit the behavior to be reused in various cultural settings. Organizing behaviors into hierarchical libraries of primitive and composite elements may permit reuse opportunities to be identified more easily, and simplify the task of maintaining common behaviors. And specifying behaviors in a language that is independent of any particular simulator or implementation may permit greater returns from the behavior specification effort.

Finally, maintainability calls for testability. For example, while writing behavior specifications, one should be able to easily preview the consequences of any choices under consideration. In the course of a simulation, one should be able to obtain an explanation for any action (or non-action) performed by a simulated member of the background population, including what behaviors the individual is performing (or not), what goals they are attempting to achieve, and what stimuli they are perceiving. The simulation must be transparent to inspection, and simulated actions must be attributable to their underlying behavior specifications and run-time conditions.

Maintainability can be characterized by the amount of effort required to perform certain incremental changes. Here one might measure the amount and type of effort required to implement some new behavior or phenomenon under certain, well-defined conditions, including testing to ensure that no regression has occurred.

## 4. PREVIOUS WORK AND REMAINING TECHNICAL CHALLENGES

Technology for simulating agents for virtual environments has been undertaken in the areas of computer graphics, urban planning, and cognitive architectures, among others; related scientific work has modeled actual or theoretical pedestrian traffic.

Realistic behavior is a central requirement for agent populations, but defining this requirement is dependent on the level of detail involved. At the level of crowds, realism depends on appropriate mass movement and (if visibly rendered) non-uniformity of individual appearance and gestures. Work has been done at the level of crowds, including academic work [viii–xi] and commercial products [xii, xiii]. At a somewhat lower level, agent populations need to include small groups of individuals like family groups and friends sharing a meal [ix, x]. Finally, realism at the individual level requires appropriate appearance, gestures, emotion, and reactive behavior [ii]

Some work has been done in capturing and modeling cultural aspects of behavior like interpersonal distance and conversational dynamics [xiv], and other work has demonstrated high-level behaviors based on intrinsic cultural values [xv, xvi].

Cognition is important both for dynamic problem-solving and realistic interactions with humans. Cognitive architectures and models have been used in simulated agents to create realistic behaviors [xvi]; however, the amount of engineering effort required can be substantial.

The bottom line is that today we can simulate a background population consisting of a very small numbers of highly credible agents (e.g., [xvi]), or we can simulate a large population of relatively simple ones (e.g., [xvii]). But we cannot yet simulate a population that is both large and credible—a simulation that achieves all three key qualities of credibility, scalability and maintainability. The technologies yet needed are of two general kinds: those that support efficient behavior creation and maintenance, and those that support efficient simulation. We shall characterize each in the following paragraphs.

### 4.1. Efficient behavior creation and maintenance

With today's technologies, specifying a moderately complex set of behaviors for a few simulated agents typically requires months of effort by engineers and scientists (often graduate students) who have detailed knowledge of simulation software. While such efforts have been valuable demonstrations and tests of principles and approaches, they have each stood alone, not significantly contributing to a shared corpus of reusable behavior specifications. The substantial effort required to author behaviors presents an obstacle to the development of new simulations, and the extension and diversification of existing ones.

We need technologies that enable those who are knowledgeable about cultures and behaviors, while perhaps knowing little about simulation engineering and science, to efficiently author behavior specifications for large numbers and varieties of background agents. Several technologies that may help in this regard have already been mentioned in section 3.3, Maintainability. Also helpful would be technologies for inducing behavior specifications automatically from some examples illustrating those behaviors, such as anecdotes, newswire reports, or direct observations of the population to be modeled, either in its natural environment (the region to be modeled) or a virtual one (as human participants in a simulation). This form of learning from observation is one that humans often rely upon, and simulations ought to employ as well.

### 4.2. Efficient simulation

Economically simulating a large, sophisticated background population will also require more efficient simulation methods than those in widespread use today. Recent research has demonstrated several ways in which simulation efficiency may be improved. One idea, already alluded to in section 3.2, Scalability, is to expend simulation effort on an agent only to the extent that it can actually be observed. An agent that is traveling unseen, for example, need only be modeled coarsely, foregoing such processes as collision detection and inverse kinematics except when the agent is being observed closely. Similarly, little cognitive modeling need be performed on behalf of most agents except when they are confronted by novel stimuli. Successfully exploiting this idea requires some means of smoothly transitioning among levels of modeling detail, and, in particular, doing so at unpredictable times, as it cannot always be predicted when a human player will approach and interact with a simulated agent.

A complementary approach to achieving greater efficiency is to off-load certain calculations to stream processors, such as graphics and physics processing units (GPUs and PPUs), that can perform these calculations more economically. Data-intensive operations that consume a large proportion of simulation time for simple agents, such as line-of-sight determination and route planning, have already been shown to be amenable to GPU implementation [xviii]. It remains to be seen whether operations involved in maintaining cognitive models for more complex agents can be similarly accelerated.

## 5. VERIFICATION AND VALIDATION

A simulation will not be accepted and employed unless users believe it is authentic. It is therefore essential to have verification and validation approaches that yield high confidence in simulation correctness. This is especially so in the simulation of background populations, where the correctness of simulated actions may not always be readily apparent.

Verification, the process of determining whether a simulation meets its specifications, can be helped by some of the ideas mentioned in section 3.3, Maintainability. A formal specification of behaviors, including their pre- and post-conditions, will permit reasoning about consistency during authoring, and verification of post-conditions during simulation. Aggregate statistics, such as the rate of arrivals at a checkpoint or the density of pedestrians on a street, can also be used for cross-checking during simulation. Moreover, the simulation itself should be able to explain the basis for any action by a simulated agent, to allow verification of any that might seem incorrect.

Validation, the process of determining whether a simulation meets users's needs, can be based partly on the three key qualities mentioned in section 3—credibility, scalability, and maintainability. Validation should also consider more specific user requirements, although the manner in which it might do so is beyond the scope of this paper.

## 6. CONCLUSION

Training for urban operations, low-intensity conflict and asymmetric warfare can be performed in virtual environments, provided those environments include realistic background populations of civilians who exhibit diverse and complex behaviors that are culturally and regionally appropriate. We have argued that, to be successful, a simulation of background populations must be credible, scalable and maintainable. Achieving such simulations will require technological improvements that support more efficient authoring of agent behaviors and more efficient simulation of them. To ensure their acceptance, simulations must undergo comprehensive validation and verification, especially since the correctness of simulated actions may not always be readily apparent.

**References**


[i] S. Hoxie et al. 1998. "Developments in Standards for Networked Virtual Reality." IEEE Computer Graphics and Applications, 18, no. 2: 6–9.

[ii] S. Marsella; J. Gratch. 2001. "Tears and Fears: Modeling emotions and emotional behaviors in synthetic agents." In Proc. of the 5th International Conference on Autonomous Agents. ACM Press, New York, NY, 278–285.

[iii] W.L. Johnson; H. Vilhjalmsson; S. Marsella. 2005. "Serious Games for Language Learning: How Much Game, How Much AI?" In Proc. International Conference on Artificial Intelligence in Education (Amsterdam, The Netherlands, July 18-22). IOS Press.

[iv] America's Army. http://www.americasarmy.com.

[v] Kuma LLC. "Kuma\War". http://www.kumawar.com.

[vi] Demographia. "Demographia World Urban Areas (World Agglomerations)." March 2007.

[vii] International Organization for Standardization. 2001. "ISO/IEC 9126: Software engineering—Product quality".

[viii] G.A. Kaminka; N. Fridman. 2006. "A Cognitive Model of Crowd Behavior Based on Social Comparison Theory." In. Proc. of the 2006 AAAI Workshop on Cognitive Modeling. AAAI Press, Menlo Park, CA, 25–35.

[ix] S.R. Musse; D. Thalmann. 2001. "Hierarchical model for real time simulation of virtual human crowds." IEEE Transactions on Visualization and Computer Graphics, 7, no. 2: 152–164.

[x] M.D. Petty et al. 2004. "Developing a Crowd Federate for Military Simulation." In Proc. Spring 2004 Simulation Interoperability Workshop. Simulation Interoperability Standards Organization, Orlando, FL, 483-493.

[xi] M. Sung; L. Kovar; M. Gleicher. 2005. "Fast and accurate goal-directed motion synthesis for crowds." In Proc. of the 2005 ACM SIGGRAPH/Eurographics Symp. on Computer Animation. ACM Press. 291–300.

[xii] MASA-SCI. "SWORD Packaged Simulation Software". http://www.masa-sci.com/sword.htm.

[xiii] Massive Software. "Massive Prime". http://www.massivesoftware.com/products/prime/.

[xiv] J. Dusan et al. 2007. "A Computational Model of Culture-Specific Conversational Behavior." In Proc. Intelligent Virtual Agents (Paris, France, Sept. 17–19). Springer, 45–56.

[xv] B.G. Silverman et al. 2006. "Human Behavior Models for Agents in Simulators and Games: Part I: Enabling Science with PMFserv." Presence, 15, no. 2: 139–162.

[xvi] W. Swartout et al. 2006. "Toward Virtual Humans." AI Magazine, 27, no. 2.

[xvii] W. Shao; D. Terzopoulos. 2005. "Autonomous pedestrians." In Proc. of the 2005 ACM SIGGRAPH/Eurographics Symp. on Computer Animation. ACM Press. 19–28.

[xviii] M. Verdesca et al. 2006. "Using Graphics Processor Units to Accelerate OneSAF: A Case Study in Technology Transition." JDMS: The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology, 3, no. 3: 177-187.

**Biographies**

Dr. Art Pope has more than 25 years of experience in computer vision, video exploitation, machine learning, distributed processing, and interactive simulation systems. While at Bolt Beranek and Newman in the 1980's, he was the principal author of the SIMNET protocols that later became the basis for IEEE standard 1278, protocols for distributed interactive simulation (DIS). More recently, he has been a senior member of the technical staff at Sarnoff Corporation, and a senior scientist at SET Corporation. He has S.M and Ph.D. degrees in computer science from Harvard Univ. and the Univ. of British Columbia.

Dr. Peter Selfridge received his Ph.D. in Computer Science from the University of Rochester. He has worked as a researcher in many areas, including computer vision, robotics, artificial intelligence, knowledge representation, knowledge discovery, and 3D web-enabled technologies. He has published numerous technical articles over the years and has been awarded 9 joint patents. After 20 years at Bell Laboratories and then AT&T Laboratories, he worked for a small computer graphics company and then joined the Science, Engineering, and Technology (SET) Corporation in 2003, where he primarily provides advanced technical support to DARPA/IPTO.

Dr. John Surdu's research interests include simulation, battle command, semi-automated agents, and novel human interface technologies. His current focus is on advancing the science of military simulation and on applications of simulation capabilities to battle command. This includes new paradigms as well as algorithms for military modeling; technologies to interleave anticipatory planning and adaptive execution of military operations; and more intelligent, light-weight, multi-resolution SAF entities. Prior to his assignment at DARPA, COL Surdu was the Army product manager for the One Semi-Automated Forces simulation effort. He also worked in the Information Technology and Operations Center (ITOC) at the United States Military Academy as an assistant professor and senior researcher. He also led the virtual sand table project at the Army Research Laboratory. In addition to military education, including Command and General Staff College, Dr. Surdu holds a B.S. in computer science from the United States Military Academy, West Point, an MBA from Columbus State University, a M.S. in computer science from Florida State University, and a Ph.D. in computer science from Texas A&M University.