Cosmopolis: A Massively Multiplayer Online Game for Social and Behavioral Research

Marc Spraragen¹, Peter Landwehr², Balakrishnan Ranganathan¹, Michael Zyda¹, Kathleen Carley², Yu-Han Chang¹, and Rajiv Maheswaran¹
¹University of Southern California, GamePipe Laboratory
²Carnegie Mellon University, CASOS Center
³USC Information Sciences Institute
{sprarage,brangana,zyda}@usc.edu,{plandweh,Kathleen.carley}@cs.cmu.edu

ABSTRACT

The community aspects of Massively Multiplayer Online Games (MMOGs) are an exciting opportunity for studying social and behavioral models. For that purpose we have developed Cosmopolis, an MMOG designed to appeal to a wide variety of player types, and that contains several key research-oriented features. The course of development has revealed several challenges in integrating behavioral models with an MMOG test bed. However, the Human Social, Cultural, and Behavioral (HSCB) research value of Cosmopolis has been demonstrated with a number of prototype studies, and based on these studies and challenges we propose an ongoing experimental plan largely driven by collaboration with HSCB researchers.

INTRODUCTION AND MOTIVATION

A 2008 study by the National Research Council entitled “Behavioral Modeling and Simulation – from Individuals to Societies” [NRC 2008] discusses how we need to expand research in modeling and simulation to include models of individual and societal behaviors. In the study, it is pointed out that a technological infrastructure needs to be developed for behavioral modeling such that we can properly develop, test and then deploy such models. The study, in fact, suggests the development of a massively multiplayer online game (MMOG) for that infrastructure. Such an MMOG can be utilized as a test bed for models of individual and group phenomena.

Cosmopolis is an MMOG we have developed for this purpose. In designing the game, we have been motivated by the need to balance the diverse interests of players and researchers: players want an engaging game experience, while Human Social, Cultural, and Behavioral (HSCB) researchers need the flexibility to perform various experiments of their own design. To accomplish these goals, we’ve designed the game to incorporate specific features to work towards these ends.
For players, *Cosmopolis* is an MMOG built around an outer game world and a collection of subgames; these subgames may be of any genre (action, puzzles, sports, etc.). For researchers, *Cosmopolis* is a unique test bed and data source for studying social and behavioral models, particularly via custom-designed experimental subgames. The models can be of individual players or multiple players over time, as well as of non-player (AI/Artificial Intelligence) characters (NPCs), or combinations of the above. *Cosmopolis* provides various and flexible methods to facilitate these needs. *Cosmopolis* also has a novel approach to information channels, with multiple real-world and game world sources being combined to create effects on the game AI, and customized output for players.

Section 2 of this paper will discuss the theoretical framework of Cosmopolis from game design and social/behavioral modeling perspectives. Section 3 will cover the approach to development of the game engine, in terms of both design and engineering. In Sections 4 and 5, the first set of experiments in Cosmopolis will be described and their results analyzed. Section 6 will largely be devoted to future work on Cosmopolis, particularly relating to challenges discovered during development, and a formal plan for upcoming social/behavioral modeling experiments. Section 7 will summarize and draw conclusions from our findings during the Cosmopolis project.

**THEORETICAL FRAMEWORK**

Research into video games and their scientific uses has currently taken several paths. Work that has taken a Human-Computer Interaction (HCI) perspective has looked at issues of real world reaction to virtual appearance, [Yee2007]. Other researchers have begun investigating the broader issue of how real world social phenomena translate into virtual spaces. Castronova specifically proposed that virtual worlds might serve as ideal platforms for experimenting with a wide variety of individual, organizational, and societal (IOS) models [Castronova2006]. To demonstrate this point, he carried out a small scale experiment demonstrating that the real world concepts of supply and demand mapped reasonably to a virtual space [Castronova 2008].

A variety of researchers have also looked at the social structures that form in games and their relative strengths and weaknesses. Such analyses have been derived from qualitative and ethnographic observation of player interactions, from surveys of player opinions, and from social network analyses of the strength of ties between different players. One notable ethnography-based analysis is Pearce’s long term study of “Uruvian expatriates” (players of Myst: Uru Online who migrated to There.com after the first game’s abrupt closure), and the roles that emerged among them. [Pearce2009] Social network analyses have been specifically conducted using both Everquest and World of Warcraft, [Ducheneaut2006,Huang2009,Huffaker2009,Williams et al. 2006] demonstrating the relatively small levels of interaction among players within the same guild structures, while Johnson et al. have developed a model of guild formation patterns that also helps to explain the formation patterns of offline gangs. [Johnson2009]

All of these efforts fall into the general category of mapping, as described by Williams: researchers want to know how virtual actions and representations serve as analogs of real world actions and representations. [Williams2009] By implementing a new MMOG, as opposed to relying on working in the diversity of extant MMOGs and virtual spaces, we can establish a unified mapping
environment in which a variety of different phenomena can be explored, linked across a single space. Then, as our understanding of mapping principles develops, we will be able to implement and test different mappings with autonomy unavailable to a corporation beholden to a much more fixed game structure.

Mapping is a difficult phenomenon to deal with in MMOG development because it is difficult to predict exactly how different experiences will map for different individuals. That said, game designers have done considerable work to try and understand the differing natures of play styles practiced by different individuals in online environments. Designers have long been aware of the emergent values and behaviors of different MMOG communities and attempted to foster a broader awareness of this fact in the community at large. Raph Koster, one of the designers of Ultima Online and lead designer of Star Wars Galaxies, famously formulated that “[An MMOG is] a community. Not a game. Anyone who says, ‘it's just a game’ is missing the point.” [Koster2009] Morningstar and Farmer, developers of LucasArts’s social game Habitat, encountered the same phenomenon and noted that “a cyberspace is defined more by the interactions among the actors within it than by the technology with which it is implemented” and that from a design standpoint “detailed central planning is impossible.” [Farmer1990] While the rules and incentive structures for certain behaviors can be incorporated into MMOGs, players will be driven by their own motivations as well. Instead of looking at player growth as a process opposed to these rule structures, however, community development should be considered in tandem to them. Players’ reactions to the IOS models as implemented in the game environments will help to evolve our understanding of these models themselves.

That said, to apply these ideas to our specific development of an MMOG for the study of individual, organizational, and societal (IOS) models, it is still necessary to develop both an understanding of the community that will play the game and a method for allowing investigators to translate the salient features of IOS models into game dynamics. While measuring players’ reactions to model implementations is essential, it is impossible to engage in accurate study without any theory of the base population. Bartle notably broke down players into four types based on discussion within a game’s forum about what people want out of a multi-user dungeon or MUD (a precursor of modern MMOGs): Achievers, Explorers, Socializers, and Killers. Bartle posited that these player groups can exist in various stable states of flux, determined by the type of MUD that had been created [Bartle1996]. Yee later followed up this work with an attempt at a multi-factor analysis of the survey results from players of different MMOGs, identifying three salient factors in players’ motivations for play: Achievement, Socialization, and Immersion (into a virtual environment). Yee also noted that these factors did not suppress each other, but might actually coexist within an individual at equal intensities, bolstering each other. [Yee2006] As noted earlier, significant research has already been done to determine the demographics of several conventional MMOGs, though Aschbacher’s report on Whyville demonstrates the possibility for considerable demographic variability based on design, a phenomenon also at the heart of Pearce’s study. [Aschbacher2004, Pearce2009] Game designers can appeal to all or some subset of these perspectives via design choices, and in creating Cosmopolis we have sought to provide a framework that would support multiple combinations of desires. Additionally, given our expected ability to segment players based on their play habits and associations, we can hope to provide a more refined breakdown of play habits than has been previously found.
Another important aspect of creating and instrumenting a game, particularly an MMOG, as a tool for social research is that of using competition among players as a means to reduce search complexity over a large-scale problem space. Our “human heuristic hypothesis” is the assumption that competing human players in increasing numbers and at growing levels of expertise will be able to find better solutions in large-scale social scenarios than would brute-force AI methods assigned the same problem. This is particularly true if the conditions discussed above hold in order to keep the players engaged and on task.

**APPROACH**

As a research test bed, Cosmopolis offers a critical degree of experimental flexibility beyond the data-logging capability of the standard MMOG. Our overall design comprises a federated model architecture: each subgame is a potential lab for a different social and behavioral model, maintaining interoperability with the outer game world model. Subgames may be added, and gameplay of the outer world can be tweaked, all to meet the needs of researchers who use our game to validate or collect data for their models. While all in-game events will be logged, we will be specifically providing appropriate data export, reporting, and visualization capabilities so that researchers can easily analyze the experiments that they design and conduct in the game environment. Exported data would include player characteristics and activities, relational information such as who played with whom, performance outcomes, geo-temporal activity sequences and so on.

From an engineering standpoint, the subgames’ content and logic is completely isolated from the outer game except for a controlled data access pipeline. This enables administrators and researchers to determine exactly what parts of the outer world (or other subgames) that a subgame can modify, to prevent any unexpected behavior. This also means that in the event of bugs or design inconsistencies in a subgame, it can safely be taken down without affecting the rest of the game.

The event-based networking model enables efficient logging management, which is vital for researchers using Cosmopolis as a data source. The logging parameters can vary from player to player, subgame to subgame, based on the needs of various researchers. The current networking model has a separate gameplay and analysis server. The analysis server can be tasked to perform near real-time processing in addition to logging the data.

To separate the account management from gameplay logic, Cosmopolis uses web services to perform authentication and initiate game connection. This also paves the way for enabling researchers to control different parameters of the game from the browser—anything from adding characters to changing the weather.

**TRIBES EXPERIMENT SYNOPSIS**

One of the key ideas in the development of Cosmopolis was to have an AI system underlying the world and providing feedback for player actions. Such an AI would open the world for possible large-scale experiments involving the game’s entire population. As a first step towards this, we began work on a basic
experiment using player interaction with a computational model to investigate the possibility of crowdsourcing policy development.

The research question underlying Tribes was whether individual players, when presented with a computational model with too many configurations to be completely solved by a computer, could find optimal configurations through reasoning. In the same way that researchers have built custom scenarios in SimCity, accepted the model’s assumptions, and analyzed the results [Pescohon1996] we sought to build a model, have individuals play it, and would then analyze their results.

In the case of this game, we developed a model of inter-tribal relationships in Sudan. Our model was coded using Construct [Schreiber2004], a simulation engine that has been used for a variety of different social simulations. Construct posits that agents interact and exchange facts; these facts have positive or negative associations with particular beliefs. Correspondingly, agents choose to interact with each other based in part on defined affinities and in part based on overall similarity of belief.

![Figure 5: Player comparing tribal beliefs in the Sudan Tribes subgame](image)

We modeled Sudan as a set of 14 prominent tribes, each with a different degree of affinity to the others, and each tribe with a leader. The tribal prominence and affinity data is based on a semi-automated analysis of the corpus of the Sudan Tribune newspaper over the period of 2003-2010. The newspaper transcripts are reduced to a pre-defined list of relevant entities that are then linked together based on their relative proximity. Our data thus assumes closer relationships between tribes discussed in the same context and more distant relationships between those that are not mentioned in the same context. We correspondingly assume that the most important Sudanese tribes are those mentioned most often in the paper. The process of vetting and analyzing this data was not part of our own research and is more thoroughly covered in [VanHolt2011]. Tables describing the key components of the model can all be found in Appendix E at the end of the paper.
The Sudanese are not only defined in terms of tribe membership but also as possessing a set of eight beliefs, each consisting of ten facts – five positive and five negative. The set of beliefs and their distribution across the different tribes were derived from selected readings on Sudan and consultation with a Subject Matter Expert (SME) on Sudan [Carley2010,ICS2007,Johnson2006]. We posit that a central cause of tension in Sudan to be that different tribes possess markedly different beliefs, and that to rectify this it is important to bring the tribes’ beliefs into alignment. We measure hostility by considering the net difference in average beliefs between each tribe pair. If at least four of the beliefs of the tribe in the pair differ more than a particular threshold, we consider the tribes to be hostile. All of these hostilities are then normalized across the entire set of possible pairs, providing an overall hostility indicator.

To support the propagation of beliefs and the corresponding reduction of hostility, our model contains two different interventions: one where tribal or national leaders give speeches to their particular regions and one where leaders meet in conference to increase their own knowledge (approximating the Tamazuj forums which occurred in Sudan before the referendum [ICS2010]). To “win” the model, a sequence of interventions needs to be chosen that will successfully reduce hostility below a particular threshold. (Our theoretical minimum is 0.2, but this is not based on a fully play-tested implementation of the game.) While it may seem that by incorporating only two forms of intervention our model keeps the complexity limited, this is not the case. Given that our model incorporates 14 tribal leaders and six national leaders, eight beliefs that can be construed either positively or negatively, and the ability of leaders to meet in groups of two, three, or four, there are:

possible interventions at any time period. This number grows exponentially over time, making complete computation impossible.

Successful solutions to the Sudan game can be considered high-level policy solutions for Sudan - an ordered mixture of conferences and interventions with specific tribes. While one player’s successful game should not be a sole determiner of policy, a winning combination of choices that is validated with multiple simulation runs outside of the game could be used to make a recommendation about an appropriate course of action. Players thus serve as individual policy analysts, their opinions and choices informed by the information that is included within the scope of the simulation and the game that is wrapped around it.

The Tribes game is a graphic interface to this model. A player takes on a role as a member of the UN, talking to the different tribal leaders mentioned above, and choosing interventions based on their understanding of the current situation in Sudan. This understanding is cultivated in several ways. First, all leaders provide a modicum of contextual text explaining their relationship to their tribe or the nation as a whole. The player can also ask the leaders about the precise
effects of their interventions on Sudan. Secondly, and more importantly, the player can access visual information about the beliefs of the different Sudanese tribes, the beliefs about which a leader is informed (that is, about which they could give an effective speech), and the current hostility in the country.

Beyond tracking players’ solutions and their relative effectiveness, we plan to look at players’ dialogue in Cosmopolis relating to Tribes, as well as their interactions in the section of Cosmopolis’s online forum dedicated to Tribes. By doing so, we hope to better understand players’ motivations for why they make the choices that they do when playing. Players will be developing their solutions based on their understanding of Sudan as presented in the game. It is not essential that they develop a deeper understanding of the policies of the country -the game is based on the assumption that such an understanding should not be required of the players- but to not have any insight into their process is anathema to common sense. As such, we will log and analyze this conversation in order to better understand not simply players’ decisions but also why they make them.

Because Tribes has not been launched, no formal results currently exist for it. However, we have tested out the model in simulation and as such have some understanding of the game’s difficulty. To prepare, we have carried out two alternate “greedy” test simulations of the game. By “greedy”, we mean that in analyzing the simulation we commit to the best choice for a particular time period and then determine the best choice for the next time period, thus building a complete set of best choices. We do not, however, limit our determination of the best possible option by only looking at the current time period. Rather, we look at the long-term consequences of an intervention (as if no future interventions occurred) and thus choose interventions based on net performance over time. If no intervention can be considered the best, we choose randomly between them.

A downside of the greedy approach is that it makes it impossible to study the impact of conferences; a conference between leaders will have no immediate positive impact on the beliefs of the Sudanese. To compensate for this, we have run two alternate sets of simulations, one in which all of the leaders have their natural knowledge, and one in which all leaders have perfect knowledge – the maximum possible outcome from a set of conferences. These two set of simulations thus provide bounds for the problem space – the maximum and minimum impact that can be had by the best intervention at a particular time period.

Our results from these tests are shown below. They suggest several key elements that will need to be taken into account when actually releasing the game. The first of these is that the difficulty level of the game needs to be significantly refined. Our initial plan was to ask the player to reduce hostility to a level of 0.2; this was not possible in even our optimal case, where the lowest hostility value seen was 0.46153885 and occurred because of the first intervention. That is, we never see a hostility value lower than this initial case, suggesting that it may be impossible to dramatically alter the hostility levels in the country. Any model that is this difficult to manipulate needs to be reworked to be made
amenable to gameplay. The player needs to think that they are making headway on a problem, not running into a brick wall.

Secondly, we need to develop other methods of probing the model. This initial map provides us with some boundaries for the space, but a better map could be made by using AI agents to make greedy decisions on the fly as opposed to by locking in a particular solution at each time period. Our current method works best to find a single, optimal solution for each time period. Given that in both the imperfect and perfect knowledge cases we ran into a situation where more than 100 possible choices performed equally well at one time period, randomly choosing a particular path will not necessarily guarantee any long term measure of success. Looking at the output of a host of AI agents will certainly not yield a single best path, but will definitely provide better bounds on the space.

| Intervention | Imperfect Knowledge | | Perfect Knowledge | | |
|--------------|---------------------|------------------------|-------------------|------------------------|
|              | Min. Hostility Seen | # of Minima | Max. Hostility Seen | # of Minima | Max. Hostility Seen |
| 1            | 0.4725275           | 1           | 0.5274725          | 1           | 0.5274725          |
| 2            | 0.4615385           | 1           | 0.5274725          | 1           | 0.5274725          |
| 3            | 0.4725275           | 2           | 0.5274725          | 1           | 0.5274725          |
| 4            | 0.4725275           | 1           | 0.5384615          | 2           | 0.5164835          |
| 5            | 0.4615385           | 1           | 0.5274725          | 1           | 0.5274725          |
| 6            | 0.4725275           | 4           | 0.5384615          | 2           | 0.5164835          |
| 7            | 0.4725275           | 1           | 0.5274725          | 1           | 0.5274725          |
| 8            | 0.4725275           | 3           | 0.5274725          | 1           | 0.5274725          |
| 9            | 0.4835165           | 283         | 0.5274725          | 1           | 0.5274725          |
| 10           | 0.4725275           | 1           | 0.5274725          | 1           | 0.5274725          |

Table 1: Impact of interventions in Tribes game

CONCLUSION

MMOGs are widespread and popular online communities; *World of Warcraft* alone boasts millions of subscribed players. The significance of *Cosmopolis* is its uniqueness as an MMOG designed specifically as a research testbed for social and behavioral models, with a correspondingly high degree of researcher control over experiments performed in and data gleaned from the game world. A few of the key features that *Cosmopolis* incorporates are a multi-genre system of subgames, a dynamically modifiable outer world, and a channel-based information system featuring real-world feeds and game-world effects. While these features help to make the game novel and engaging, they also have specific applications for scientists opting to use *Cosmopolis* as a research platform: subgames are a way for researchers to conduct isolated experiments; the modifiable nature of the game world allows for events to occur that may dramatically alter the main game environment, providing fodder for scientists interested in the evolution of online communities; and the information broadcasting systems will allow different messages to be broadcast to different portions of the community to help manage experiments conducted on the entire player community. Also, any and all *Cosmopolis* actions and interactions (including the internal processes of AI-based non-player characters) may be logged into our databases, and may be used to explore the mappings between game world and real world societies. Ready access to a high-fidelity data set
means that researchers will have an easier time determining the impacts of different experiments on the community in Cosmopolis than do researchers of more closed gaming environments. It is impossible for one MMOG to be considered the definitive online game, and Cosmopolis is not intended to be that. But it is an important step in opening up game environments for use by researchers, and one that can help support the work of scientists interested in studying game environments and how different social and behavioral phenomena manifest within them. We hope that the public presence of Cosmopolis will encourage other researchers to look to our game environment as an avenue for research into human behavior.

REFERENCES


Carley, K.C. Answers on Sudan, 26-Jul-2010.


Landwehr, P., Karley, C., Spraragen, M., Ranganathan, B., and Zyda, M. Planning a Cosmopolis: Key Features of an MMOG for Social Science Research. In Proceedings of CHI-10, Atlanta, GA.


Schneider, M., Carley, K., and Moon, I. Detailed Comparison of America's Army Game and Unit of Action Experiments, Pittsburgh, PA: Carnegie Mellon University, 2005.


Williams, D. The mapping principle, and a research framework for virtual worlds. University of Southern California, Los Angeles, California, 2009.


Yee, N. Motivations for Play in Online Games. CyberPsychology & Behavior 9, 6 (2006), 772-775.

