GIS, MODELING AND HUMAN CIVILIZATION: THE BIRTH OF GEO-SOCIAL ENGINEERING

E. Scott Morris

Thesis Prepared for the Degree of

MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

December 2011

APPROVED:

Pinliang Dong, Major Professor
John Neuberger, Minor Professor
Robert Pavur, Committee Member
Paul Hudak, Committee Member and Chair of the Department of Geography
Victor Prybutok, Associate Dean of Academic Affairs in the Toulouse Graduate School
James D. Meernik, Acting Dean of the Toulouse Graduate School

Morris, E. Scott. *GIS, Modeling and Human Civilization: The Birth of Geo-Social Engineering*. Master of Science (Interdisciplinary Studies), December 2011, 98 pp., 7 figures, references, 17 titles.

Computer-based, mathematical models have significant value in describing the processes behind urban development and its inhabitants. The following research describes the theories and concepts behind modeling and offers insight into the potential future of the field. First, the research covers a brief history of applicable modeling strategies. This is followed by a summary of current popular approaches. The numerical background of geo-social engineering is developed through mathematical techniques. Geo-social engineering is the integration of modeling into the basic design human civilization. The mathematical models will be incorporated into a design of a computer program. From this, a possible geo-social model structure is presented and its architecture is described.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

BACKGROUND

Introduction to GIS

The global positioning system (GPS) was developed by the United States in the early 1970s. By the mid-90s, it was fully operational with a constellation of 24 satellites. Each satellite broadcasts its location and the time at that location. Ground based receivers triangulate their location based on the signal from four or more satellites. The Cold War spurred the investment into the technology and by 2000 non-degraded signals were made available to the global public. By the middle of the same decade, GPS technology was being integrated into cell phones. GPS accounts for the technology providing specific (x, y) coordinates, but it is a geographic information system (GIS), which contains the information associated with each location. Layers of data can address population, development, vegetation, topography, meteorology or any other set of information the user wishes to include. GIS is a burgeoning industry capable of applications within nearly every industrial sector. The data structure of GIS is well suited for integrating modeling into daily life. GIS facilitates daily tasks like weather forecasting and reporting traffic congestion or complicated problem solving and data analysis. GIS can be perceived as a spatially linked database in which elements (entries, rows, features etc.) contain attribute data combined with a location in space. Therefore, users can project data onto maps and incorporate spatial relationships into their analysis.

GIS offers every benefit of database management with the added facet of describing, discovering and displaying spatial relationships. This fact makes GIS an ideal tool to store individual behavior, log events and track movements through an environment. Individuals can be clustered or combined into groups and communities such as neighborhoods, census tracts, cities, counties etc. Economic entities can be considered as functions of employment, members of a larger industry sector or an individual entity with its own behaviors and attributes. In this fashion GIS applications can be tailored to specific intentions of the user. Municipalities, businesses and academic institutes are discovering that GIS is becoming necessary to provide or purchase.

The geographic foundation of GIS and urban development makes GIS the best option for modeling humans and the environment. Models of civilization incorporate geographic elements pertaining to specific location, activities occurring at that location and the connection between locations. Geo-social boundaries can be explored as the relational parameters give rise to the geographic classifications, which include physical, personal and statistical associations. GIS can assign a topology to networks and can thus describe the connections of objects, areas and the behavior. GIS data structure also increases its suitability to not only describe human behaviors, but also monitor and improve the physical accessibility of the areas in which those behaviors occur. In GIS, data can be presented as features (points, lines and polygon). These could be roads, rivers, buildings or any other physical object the user incorporates. In addition, GIS can store in present data in a raster format, in which areas are divided into cells. Each cell can assume a range of values and display information based upon the assigned coding scale. This specific aspect is naturally conducive to some of the modeling techniques described by this research.

Models are built to describe reality, although each model provides insight into a specific sliver of the entirety of human existence. Models focus upon a specific arena, which is abstracted and idealized. Nonetheless, models enhance our understanding of reality and provide a framework for the creation of more sophisticated modeling techniques and testing a variety of parameters and scenarios. As human technological ability increases, models become better able to describe reality. Feedback from these results can then update a database, improve algorithms and motivate more accurate decisions.

Urban Model History

Models are classified as abstractions or simplifications of reality. They can take on a physical or digital form. They can reflect a current state or forecast projections of future scenarios. The historical meaning of the term model was that of an ideal situation, but acceptance into the scientific realm has extended the definition. The purpose that models serve increases as the capacity and sensitivity with which data is revealed evolving to a near real-time status. Most importantly, models are the bridge between theory and reality and can be used to test or implement theories. Potential implications of such research can help authorities and policy makers render decisions and implement monitoring systems. They allow the exploration of scenarios and alteration of variables without the expense and implausibility of manipulating the real world. Decisions then can be made with this added insight that efficient models provide. This is of particular value when analyzing complex urban situations in which there are numerous factors and large-scale environments that are not easily altered.

In 1826, Johann Von Thunen introduced an early version of urban model. It focused on the cost of production, rent and transport relative to the distance from the central business district (CBD). These concepts were based predominantly upon agricultural production, spoilage and refrigeration costs. The result is concentric circles around the CBD. These respective zones were indicative of areas financially viable for production of certain commodities based upon market value (Batty, Longley, 1994, 125). Cities and the market place have changed drastically since von Thunen. Likewise, the dynamics of urban development have a much different form and complicated behavior. However, the notion of describing cities in relation to location and economic concerns is a foundation for developing sophisticated computer models. The implementation of such models offers insight to the nature in which humans, geography and public policy manifest the physical form of cities and quality of life of their inhabitants.

Considering multiple economic centers enhanced the CBD concept. Based upon his study of cities in southern Germany during the early 1900s, Walter Christaller developed the basis of central place theory. This concept works on some of the similar assumptions as von Thunen, such as assuming flat land without geographical obstacles (Batty, Longley, 1994, 55). In addition, central place theory introduces the notion of hierarchy in that the larger cities serve larger areas. The resulting geometry is a lattice of hexagons describing the areas surrounding cities.

Systems can be optimized according to market principles, transport principles or administrative principles.

Thomas Schelling introduced a segregation model in 1971. He described an area as a grid of cells or nodes with each assuming one or two values (colors). Each unit was coded with a simple set of rules or preferences regarding the adjacency of units with similar values. As the model iterates through time, the neighborhood of each cell is evaluated. If a certain threshold of satisfaction is maintained, there is no resulting action. However, if the number of neighbors, which are of different orientation, exceeds the threshold, then the cell is reassigned either in value or location depending on the behaviors of the system. Schelling demonstrated that even slight tendencies or preferences can dramatically impact the entire system (Batty, 2007, 57). Schelling's approach is a significant contributor to the development of agent-based modeling (ABM) and cellular automata (CA). Incorporating these concepts renders more detailed and powerful interpretations of the reciprocal relationship between geography and development.

Early models of urban development merged ideas from social physics and location theory. Simulations represented the central business district, agricultural resources and transportation costs between the two. These patterns are then combined with microeconomics to view locations as points connected via a network of transportation and finance. Current research is still addressing these same issues while addressing environmental dynamics. Fortunately, the technologies to process and theories to interpret the complex socio-economic environment have also evolved.

The manner in which land is developed and occupied greatly impacts the overall quality, health and livelihood of an urban environment. Meanwhile, the inhabitants make choices regarding time and space, which establish the viability of a city. While this process occurs, evidence of the health of the occupying civilization is visible in the structures built and relative quality of inhabitant activity. Current techniques focus less on the notion of equilibrium models, which don't appropriately incorporate the significance of time and interconnectivity upon social development. The equilibrium view is associated with the top down approach, which involves a dominant hierarchical perspective. In contrast, dynamical systems develop global patterns via the interaction of rather simple agents. These agents can be perceived as individuals making choices while navigating their socio-economic environment. The "bottom up" approach develops models and systems based on this amalgamation of activities and their locations over time.

GIS is the revolutionary step to follow in this process. The ability to spatially link a database is essential to connecting the reality collected through sensors and observation with the processing background code used to monitor and modify. Movement towards a grid environment (rasters) renders enhanced processing and graphic capacity. Classification of physical location as cells creates the ideal platform for cellular automata. Associated with this view is the ability to consider the surrounding neighborhood and the interactions between adjacent cells. Regardless of the system of study, this deeper degree of leverage lends itself to the bottom up method of description and development, which are vital to the health and longevity of human civilization. CA and object-oriented programming are the best tools for modeling change in dynamic areas and behaviors of agents. All of this builds upon previous modeling techniques and creates a less abstract representation of the systems, inhabitants and improvement is quality and satisfaction, respectively (Batty, 2007).

Geo-Social Engineering

Geo-social engineering is an inter-disciplinary undertaking. It incorporates aspects of mathematics, computer programming, sociology, politics and geography. These disciplines incorporate both the problem solving techniques and implementation of decision based applications. Geo-social engineering is defined as the integration process of human existence, technology and policy while minimizing time or maximizing efficiency between situation and response.

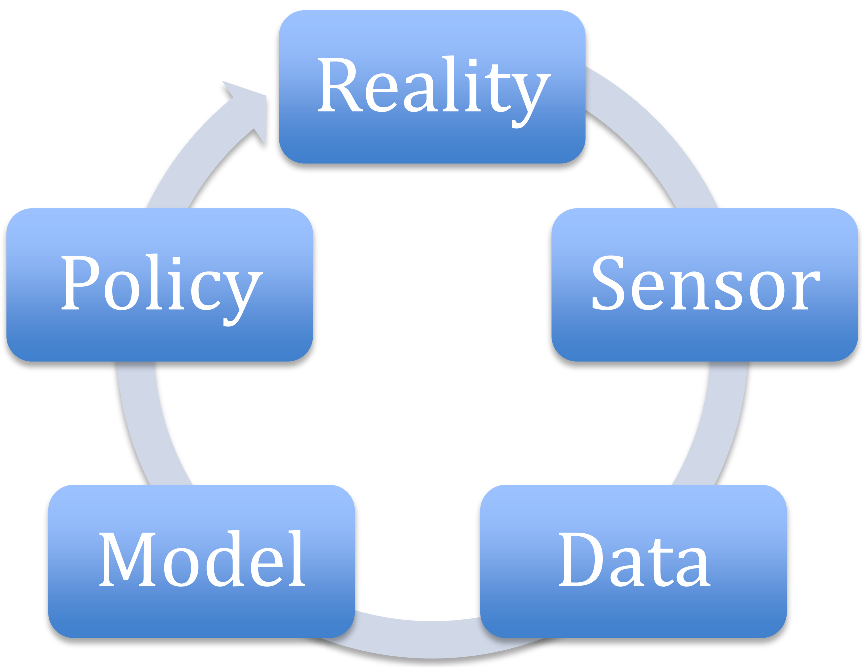

Figure 1. Elements ofGeo-Social Engineering.

As the human population grows, new problems and talents emerge. Technology evolves in response to the manifested problems. This typically occurs only when the financial constraints of scarcity are mitigated by the decrease of costs due to technological innovation. For example, as fossil fuels are consumed, new energy sources are developed. This concept is evident throughout history and has defined warfare. Power resides among those with the most effective combination of strategy and weaponry. Therefore, dominance is maintained until opponents develop more effective tactics or the dominant approach runs its course failing to adapt to inherent limitations or shifts in paradigm (see Netflix v. Blockbuster).

This combative approach is exemplified in competition for resources, economic marketplaces, technology production, development of intellectual concepts and a variety of other aspects of humanity. Essentially, competition embodies the current state of human development. This fact is exemplified by the popularity of professional sports, countless awards classifications and the constant pursuit of financial success. The intention here is not to criticize human heritage. The competitive worldview has been the basis of species survival and lead to the vast intellectual and technological resources available. However, the inherent limitations of a competitive worldview may hinder future success. These limitations offer insight to the development of a new paradigm, which includes the benefits of competition and expands to cooperative measure addressing the limits of competition. The primary limit of competition is disparity. This disparity is most evident in economic and intellectual realms. By definition of competition, there must be a loser. However, if the losers starve to death or the winner burns the library at Alexandria or thousands of documents from Mesoamerica, the loss to humanity is incalculable. Disparity reaches a critical level when the separation between winners and losers is so vast that inertia creates an unbridgeable social chasm.

The objective of geo-social engineering is to develop the theory of a unified socio-economic system incorporated into a global database. The focus of this research is to present a generalized theory and architecture. It suffices to say the primary concern is to address disparity, specifically by quantifying a humanitarian minimum threshold below which human existence is inconceivable. Thus, the basic geo-social variables are location and economy, which serves as a standardized unit of wellbeing.

CHAPTER II

PROCESS

Human Geography

Historically, traditional map making was a proprietary service restricted to those controlling the information. As the world was first documented in terms of geographical catalogs and standardization, map making marked the onset of the information as economic or military power beyond the immediate sphere of influence. Whereas, humans could only control that what was seen, by mapping and recording what existed in newly charted regions power began to consolidate among those regulating these early databases.

In England a study entitled "Children's Activities, Perceptions And Behaviour in the Local Environment (CAPABLE)" was conducted. The study investigates children's physical exertion in structured and non-structured activities. Research includes fitting children with a global positioning system (GPS) tracking devise and activity monitors. This is predicated by children's abilities and parents' confidence for kids to navigate the environment and manage situations. The children also kept diaries and engaged in mapping exercises. In addition, parents and children completed questionnaires. The researchers analyzed data describing age, gender, activity, surfaces type, levels of exertion and mapping. The unit of intensity of energy consumption is 10^{-2} calories per kilo per minute (Mackett et al., 2007).

Researchers made field visits to survey environments and geo-location data relative to land use. Classifications include natural environment, general surface, road or track, roadside, structure etc. This information allows researchers to coordinate types of activities with GPS coordinates. Map drawing assignments are conducted to reflect children's imagery of the route from home to school and the area around school. Detail and accuracy are evaluated and compared to questionnaires. The tracks formed are put into a geographic information system (GIS) and color-coded based on activity and energy exertion (Mackett et al., 2007). The CAPABLE study includes general elements similar to the basis of this research and pending survey. The fundamental link is using GPS to monitor location and behavior. Combined with individual specific data, the GIS information reveals a manner in which geography mitigates action.

As made evident by the CAPABLE study, technological advancement and social trends have revolutionized the means by which data can be collected, shared and processed. Previously, the required infrastructure prevented non-professionals from accessing the capacity of GIS, but this in no longer the case. The future of the industry relies on the ability of individuals to collect data from their experiences, upload or share the information, which, with processing and output will reflect their current state and prospective decisions. As humans increasingly integrate with the virtual world, what was once the object of entertainment becomes one of everyday use and eventually that which improves the quality of life. In a sense, the efforts of the video game industry to create a sense of realism within the fictitious gaming environment are melding with the increasing public access of geo-demographic data.

Neo-geography, as put forth by Eisnor, encompasses the practices lying beyond the edges of the realm of professional geographers. This field intends to make relevant the use of geographic models for personal expression more than merely scientific methodology. First-person gaming advances the technology and concepts by simulating an avatar moving through a digital environment. The arena of neo-geography does have the potential to contribute to the geographic community at large. Albeit fictitious at times, it embodies the "bottom up" methodology of users mapping their environment from an individual perspective. Websites such as www.secondlife.com allow users to control their avatar, manipulate their environment and interact with others via the virtual world. Comparatively, www.openstreetmap.com allows a similar reality based process. Through this website users are able to load paths and photos of their environment with the link to geographic coordinates. Data can then be aggregated. As with www.wikipedia.com, accuracy is increased by multiple revisions as more and more users compile data (Hudson-Smith, Crooks, 2008). Contributors continually refine the representation of their surroundings, which leads to both accuracy and relevance. If a multitude of users are allowed to aggregate data, the average of these approximations tends towards a more accurate solution, although the data may be composed of several inaccuracies.

This two-way online interaction (Web 2.0) allows users to map their surroundings both as they move through the environment and as external conditions are altered, such as a natural disaster. Data such as this is known as volunteered geographic information (VGI). Quickly, VGI is becoming a resource with which businesses can improve marketing strategies and policy makers can make more informed decisions. Web 2.0 is the concept of using the Internet for interaction between users rather than disseminating information in a single direction. Social networking has given rise to a digital community less dependent upon geographical proximity. By use of tagging metadata, individual users can aggregate information and people, whereas historically this was available only through third-party research. The development of databases by users themselves as opposed to a centralized management agency is one step beyond the two-way interaction. In this manner, the grid concept develops into an amorphous cloud. The concept of cloud computation includes data, resources, networks and users linked remotely without individuals consciously housing or maintaining data and software in specific locations.

Crowdsourcing is the process of database creation that is available via web-based applications. Crowdcasting is a similar mechanism, which results from a specific push by owners or managers to draw in users or participants. In crowdcasting, incentives or prizes reward contributors (Hudson-Smith et al., 2008). While the "reward" participants gain from crowdsourced data is the value of participating and being considered as contributor to the greater good. Most often, contributors to crowdsourcing are also users of the service to which they offer VGI. Crowdsourcing affords the ability to construct more elaborate data structures while also rendering them more dynamic in regards to their accuracy, flexibility and time sensitivity. Coupled with VGI, crowdsourcing delivers the capability to create maps within real-time by those exploring the environment.

The rapid evolution of technology is evident in nearly every walk of life. There are three main concepts, which encapsulate this trend. According to Moore

(1965), the speed and memory of computer chip technology doubles every 18 to 24 month. Also, the informational value of a telecommunications network is proportional to the square of the number of users of the system, where each user gains access from a single computer. First presented by Gilder (1993) Metcalfe's law estimates this relationship to be quadratic. Connected with the network performance is the trend presented by Gilder that the total bandwidth of a communication system triples every 12 months (Hudson-Smith et al., 2008). These trends make the informational age an exciting and critical phase of human development. Information is a resource that will pervade regardless of energy concerns, environmental policies, global warfare or any of the other first-world problems humans construct for themselves. Fifty years ago physical and material energy flows dominated the global economy. Today, the flow of information is the most critical economic component rendering computers and communication technologies essential to survival. According to Kurzweil, his estimate for achieving singularity is prior to 2050. The actual form this will take is uncertain, but the ability to record, quantify and map nearly every conceivable data point is inevitable.

Geo-social engineering expands beyond the concept of map-making, models can be used to investigate the types of services and structures that may be associated with specific locations. The random spin-offs of individual experiences, preferences and tastes coalesce in a complex and nondeterministic fashion indicative of the evolution of species and systems. Therefore, the technologies and techniques used to mediate and mimic this process must also invoke similar dynamics. Many of the archaic conventions and impending social limitations or hardships that ensue are a result of techno-physical reality. As humanity reaches an age at which there are few imaginable limits to technology, systems and policies, can be designed to achieve a level of humanity proportional to technological advances.

Cell Phones

Cell phones are a powerful tool for the collection of data related to populations, economies and patterns of development. Sophisticated interfaces can be customized to particular inputs depending on the area of study or pertinent variables. Cell phones also function as mobile probes or sensors. Concentrations of cell signals demonstrate population densities and mobility patterns. The pending survey design will be best implemented on a mobile platform. This allows those surveyed a convenient interface and capitalizes on the imbedded location technology. Mobile phone usage serves as proxy to the standard human interaction. Call genres are indicative of activities and include work, family, leisure and service. The connections between cell phones are the ultimate data because it contains user, time, location, types of activity and network connection within a single instance.

Telematics refers to the integration of telecommunications and information interface. An example of telematics involves the integration of vehicle control and monitoring with tracking and communication. New policies have emerged in response to this technology. E-911 is an FCC regulation requiring cellular providers to report time and location of emergency calls as well as a phone number associated with the call. The requirements of this regulation originally addressed network-based technologies, but transferred over to handsets as improvements were made. Locations are determined either by self-positioning (handheld) or remote positioning (network). Location can occur by triangulation between multiple towers (Ygnace et al, 2000). The area of closest vicinity to a tower is its cell. Cellular phones broadcast to the closest tower to access the network. This requires less power and allows for frequencies to be used by multiple users in different cells. Voronoi diagrams can represent the location of towers as declaration zones.

Cellular activity is already being applied representations of human mobility patterns. One example employs anonymous call records in which volume per cell are displayed and compared to peak hour patterns. Peaks occur in about 12-hour cycles or 10-hour cycles with 14 hours between cycles. This pattern reflects the business workday and also coincides with the probability of return (Candia et al., 2008). Connections can then be made between activity, location and time. By establishing baseline behaviors, anomalous events can be identified via deviation from the average (standard deviation). A tolerance threshold is established as sensitivity measure. Cellular plots are then made consisting of the three categories "no activity," "normal activity" or "anomalous activity." Anomalous events are clustered according to proximity (neighborhoods) and relativity to threshold. There is a greater tendency towards clusters during anomalous events. In this fashion, call volume can be used as an indication of an extreme situation or emergency which requiring authority intervention. This study also demonstrates a connection between the radius of gyration and time between calls (Candia et al., 2008).

Generally, the three methods of gathering traffic info are sensors, probe vehicles or cell probes. The basic model presented in this study is an exponential relationship between coverage with the sample size, density and link length.

Greater densities reveal higher coverage. This is a preferential caveat as these are also the areas of higher congestion. The negative aspect associated with areas of congestion is that there is an increase in directional errors since velocities tend to be lower (Ygnace et al., 2000).

Time and location of connecting cell tower were recorded, sampled and used to establish baseline behaviors. These were compared to samples and parameters calculated. Similar equations are used to model the radius of gyration occurring between consecutive calls. These paths are described as random walks (RW), Levy flight (LF), and truncated Levy flight (TLF). Levy flight follows an inverse power law and can be applied to humans as a diffusors of bank notes. Human trajectories can be modeled as continuous time random walks with wide displacement and waiting time. But cells phones are better since typically carried by single person. Radial growth over time is elliptical or circular. Time dependence of radius is logarithmic. Small radius of gyration tends towards users travelling small distance. High radius of gyration reveals users with multiple short trips coupled with occasional long trip.

Return probabilities are consistent with the cyclical patterns of location. By ranking the most visited locations, an inverse relationship to the probability of being at a particular location can be formed by exercising concepts of tension likelihoods of specific coordinates. When scaled by the variance of location, similar patterns emerge. Users can be assigned to a radius of gyration dependent upon the population distribution of the particular region (Gonzalez, Hidalgo, Barabasi, 2008).

Cell phones are saturating both the marketplace and daily human life. They have sparked the telecommunications boom and facilitate business in all sectors.

Cell phones allow users to maintain connections with personal and professional relationships. Via Internet access and powerful graphic functions, cell phones offer a portal to global information. As the capacity and convenience of these mobile devices increase, so shall the integration into all facets of human civilization. Merged with databases and modeling techniques, cell phones have the potential to serve as the primary human interface. Extending this trend, cell phones can evolve into a commonality of economic, political and behavioral aspects of human civilization.

Code

Code can be generalized to include the objects, infrastructure and processes composing the human/non-human interactions of life. Currently, focus is largely placed on the specific technologies and tangible hardware. However, an underlying code has the potential to evaluate, correct and manifest the intentions of the users. Hence, it is the fundamental logic deserving of detailed investigation in each of the following three areas: "code/space (a problem cannot be solved without code), coded space (code mediates the solution, but is not the only solution) and background coded space (potential code until activated)." These facets of code dictate the unfolding development, manipulation of space and human interaction with the environment. Each step solving an existing situation and revealing new problems of "navigation, negotiation and management" (Dodge, Kitchin, 2004).

Coded objects are non-networked and range from clocks to computers. Their hardware and software are each critical to serving the intended function. Coded infrastructure includes the networks connecting coded objects. These can be as localized as an individual automobile, as expansive as the highway system on which it drives or as inclusive as the GPS technologies that locate and coordinate the entire global traffic system. Coded processes regulate coded objects within the coded infrastructure. The processes occur via interaction between linked databases. They provide access to and feedback from objects across the infrastructure. Withdrawing money from an ATM is a coded process. Coded assemblages are the convergence and coordination of multiple processes. Individual households are created by the assemblage of the process of construction, utility services and the financial institutions both providing for the acquisition and maintained payment of a specific home.

Code manifests in human social life through five major forms. These forms are "domestic living, travelling, working, communicating and consuming" (Dodge, Kitchin, 2004). Transduction is the process of iterative solutions of urban interaction. The transduction of the human experience is mediated, supplemented and augmented by technologies. This technicity is written in the code made possible by human knowledge and skill. Thus, the technological items are integrated into their purpose to shape the human space according to the natural environment. Code is manifested via urban technicity in four primary means: objects, infrastructures, processes and assemblages (Dodge, Kitchin, 2004). It follows that solutions to long range, complex problems associated with urban development are best addressed and implemented by the technologies and codes, which also serve to facilitate the development.

Data Structure

The data model is a tiered structure. Fundamental data include background layers such as counties, cities and other political boundaries as well as roads, rivers, lakes and additional elements of infrastructure. For reference purpose, this environment is known as MAP(0). The layers included are pre-loaded or selected based upon relevance. From these sources, geographic factors are obtained which describe either suitability or limitation of the area pertaining to development opportunity.

First level data incorporates user entry and variables associated with caloric intake, economic transactions and commuting. Associated information, such as the coordinates of a location, are included in the first level. MAP(1) houses these inputs. The related interface serves as the basic survey tool. Users log their activities (nutrition, economic or transit). An individual instance is known as an event and has an associated location, time, group and value. Based upon these parameters, theoretical “containers” are constructed with a volume parameter reflecting the impact of the respective event location.

MAP(2) provides the environment for second level data that is generated via modeling and mathematical operations or essentially derived from first level data. This level of processing connects events or locations forming the network foundations. Theoretical “pipes” are constructed which emulate the flow of data, populations or common economic factors. From a physical perspective, this network is indicative of the regional transit system. However, the network can be extended to simulate flows of commodities, services, people or other facets of the system. Comparison with the roadways and other transit options can reveal tendencies or inefficiencies.

The transition of data into the MAP(3) environment is facilitated by the mathematic model and processing of data input or derived from the preceding maps. The features created in MAP(1) and linked in MAP(2) are converted into raster values. From this sources, sinks, boundaries and other modeling elements are coded into the database. Then, projections are made across areas and/or time periods through simulations.

Research at the Centre for Advanced Spatial Analysis resulted in development of a data structure applicable to this survey. Objects relate to a coded entity, which responds to agents and the environment according to properties and behaviors. People, companies, buildings and structures are examples of objects. Fields describe the relevant facets of objects such as number, value, size or location. Object-field models intend to combine the data models of objects and fields. Thus, each point in space can be mapped to a discreet object. Ideally a fusion of agent-based and object-field models can be incorporated into the data structure.

An object-based model is an empty space occupied by phenomena, entities or things. The phenomena are composed of behaviors, attributes and geometry. Complex views of phenomena require that models maintain both a functional (interactions) and structural (forms) view. Object based models have specific geometric boundaries; whereas field based views consider a range of possible values of a specific measurement.

Field-based models are characterized by the domain and range of the defining function. Fields are singular values associated with a given location. Clusters can then be formed from components, which share like values. In this sense, phenomena are formed by the compellation of cells containing like values of a specific field, which itself is the collection of measurements over a study area. Although included in the model diagram, the geometry of a field is an abstraction as field-based models operate without explicit connection to the objects they may in fact describe.

Object-field models incorporate properties of elementary geo-particle. They are flexible models capable of emulating both objects and fields. Each point within an object-field model is mapped both to a value and an object. However, the geometry is not the primary basis of the object-field model. This enables the description of complex phenomenon without new identifiers associated with dynamic changes. Elementary geo-particles can be aggregated to form objects by simple rules while relationships are formed at the phenomena level.

Cellular automata (CA) models can generate an environment on which agent-based models are conducted. The agent interaction occurs as decisions towards accomplishing goals or satisfaction by association with other individuals, organizations or institutions. The limits of these options define the bounded rationality. The physical representation of this reality is structured by the key principles of CA; cells, states, neighborhoods and transition rules. Space (although typically referring to physical location) can be represented in multiple spheres, often simultaneously. In addition to the mathematical and physical space (distance between agents and cost-based distance) socioeconomic space (location of market elements), behavioral space (fear) and experiential space (senses) impact dynamics (Voudouris 2011).

Agents exhibiting goal-oriented behaviors, primarily to reach a higher level of satisfaction contribute to the complex adaptive systems. Their performances dictate an emergence over relative scales. This emergence is not necessarily predictable as changes in seed properties reach divergent results. Perhaps, it is also likely that dominant outcomes have a gravitational property that is likely to be sought, regardless of origination. These large numbers are brought about by the random performance of human parameters regardless of ethnic, social or economic configuration. Attributes and behaviors are coded into the agent psyche as well as the organizational structure and environment. These rules ultimately manifest the behaviors of agents and structure of cells with an implied degree of uncertainty. The key distinction between agents and objects is the self-recognition that an agent is aware of its ability to control itself. Otherwise, both objects and agents reflect at least a degree of reactivity (response), social ability (language) and proactivity (satisfaction).

Holarchy

Hierarchy is a concept intrinsic to civil structures and systems. Therefore, it must be incorporated into the modeling concepts intended to describe urban development. Before discussing these in more detail, the notion of hierarchies needs clarification. First is the notion of a dominant hierarchy, which places itself somehow superior to other hierarchies. Examples include political philosophies, economic systems and religious doctrines. Often these incorporate an “us and them” perspective aligning members in opposition with those not adhering to the same paradigm. However, if our cluster includes inhabitants of planet Earth, there is only us.

Ken Wilber issues he distinction between dominant hierarchies and the natural hierarchical organization, which pervades the natural order of the universe. Borrowing the term from Arthur Koestler (1967), Wilber contends that the universe is comprised of “holons.” A holon is both an individual whole composed of preceding holons as well as a part of a subsequent holon. A holon is a whole/part and it’s holons all the way up and all the way down. A holarchy is the systematic relationship of holons and is synonymous with the natural hierarchies, which overlay existence. Each holon emerges as the result of a system of lower holons and exists as a member of a system providing for the emergence of higher holons. There is no qualitative measure associated with higher and lower orders. They merely distinguish between constituents (Wilber, 1996, 29-44). In reference to urban models, particular ranks are not associated with intrinsic value. Rather, the notion of hierarchy refers to a particular position in a sequence or rank order.

As cities develop over time and space, population distribution assumes a natural hierarchy. Ordering populations by rank of population size reveals there exist fewer cities with large populations. The simplest example utilizes Euler’s method in which a growth rate is randomly assigned. The subsequent iteration results in a population size equal to the sum of the previous size with the product of that size and its associated rate. This approach tends towards a singularity given sufficient iterations unless a limit is placed at which, when reached, the population is reset to the limit. The resulting unbounded hierarchy has a lognormal configuration, while the bounded form is near linear. The rank of populations over a variety of areas conforms to an inverse power law (Zipf's Law). This empirical pattern reveals the self-similarity of cities in that even when rescaled over a variety of populations and sizes their growth conforms to Zipf's Law. This renders a relatively simple model since it does not account for interaction or competition between cells.

A step towards improving the hierarchy approach is to introduce a diffusion component. Given a rate of diffusion, each population of a given cell migrates to one of the adjacent cells in it von Neumann neighborhood. This is a fundamental step, and even without accounting for migration via networks, it does result in density distributions centered about the cells with the highest populations. Although an improvement, there are still limits associated with this method of diffusion. One notable symptom is that each cell has the ability to rank at the top of hierarchy, whereas, in actual cities this is not the case. Individual cells or agents can be connected via networks, which influence the performance dynamics. Networks associated with nodes and cells are built by links added (or with the probability of being added) once per time period. Initially, cells are selected as candidates to form nodes. There is an indicator function, which determines whether a node is created at this cell (Batty, 2004). Indicator functions determine if and which cells are linked to the node. This is done based upon an inverse distance relationship and considers frictional effects of this distance.

As the model is developed, the hierarchy concept will be incorporated. The simplest example is the notion of rank order. Considering the same variable (area, cost, distance etc.), values are listed in descending order. The largest is then assigned the rank of 1, the second 2 and so one. Hierarchies appear in nature and are seen in geological features, tree branches, planetary bodies, rivers and countless other examples. Not surprisingly, hierarchies also emerge in economic worth, building sizes, roadways and many other social constructs. It follows that models geared to unite geography and society would also incorporate the hierarchy concept.

Complexity and Development

Complexity is the interaction of variables resulting in non-linear implications. Complex systems are comprised of these magnifying relationships receiving influence from seemingly peripheral sources. Durlauf (1997) lists the criterion for complex systems as “non-erodicity, phase transition, emergence and universality.” The criterions are respectively defined by shocks, tipping points, self-organization and self-similarity. Complex systems cannot be precisely predicted due to the vast number of combinations of a nearly infinite number of states or conditions. Solutions are therefore sought by sampling random number distributions indicative of particular situations and projecting the results via Monte Carlo simulation.

In discussing complexity, perhaps the clearest metaphor is the distinction between a machine (simple system) and an organism (complex system). The former may have a level of detail and sophistication, which render it difficult to understand, yet remain repeatable and consistent over a given range of parameters. Whereas, organisms, which are subject to certain biophysical operations and follow systematic processes, evolve with random likelihood and unpredictability. The creation, development and decay of living bodies are a confluence of events dependent upon unique interfaces with subtle, often immeasurable, circumstances of their environment. Mechanisms are built or developed, and organisms are created and evolve. Urban configuration has previously been interpreted as a deterministic configuration of infrastructure and utility. It is increasingly evident, however, that the randomness and volume of human need and satisfaction have manifested a fluctuating fabric intimately woven into the geographic environment.

Human structures and space are evolving organically along with the culture and species it fosters. These trends beckon reasoning to reexamine the modes by which urban development is designed and occurs. Specifically, the impetus is moving away from the "top down" approach of fashioning society via directive and expectation that human development conform to a predetermined design. The anticipation is more towards a "bottom up" approach in which the physical and systematic civilization is an increasingly more accurate reflection of its dwellers or participants. Technologies and theories are also evolving and available to facilitate this process. Human tendency is to merge in cooperative and competitive fashion, thus forming groups and organizations. Civil structures provide the framework upon which the fabric of human existence is woven. The choices and decisions individuals make in reaction to their situation provides motivation and solutions unpredictable through theory and experimentation. This trait highlights the complex nature of civilization.

There are three primary complexity patterns that are displayed in the evolution of cities. Each of these tendencies has an associated element of the geographic environment and code. The first exemplar is that generative systems build order in space and time from the bottom up, in respect to physical location. Harken back to the Schelling checkerboard of segregation. Given a random distribution of states, agents are provided with a mechanism by which to move towards a higher state of satisfaction. As iterative steps of decision and action progress, order is established respective to the time and space of the progression.

The second trend addresses network configuration. This can also be seen as the links between the locations described in exemplar 1. There are advantages of high local cluster density, as well as a necessity to maintain connectivity with the global environment. There exists a threshold where the benefits of the local organizations can be retained while maintaining efficiency through the entire network.

The collection of locations and the network linking them essentially comprise a system. The final paradigm reveals the tendency of these systems. Fundamentally, there is the notion of exponential growth of organisms. Built from the exponential growth model is the concept of an ecological logistic which incorporates the notion of a carrying capacity. This capacity serves as a limit by which the growth is limited. Idealized urban development demonstrates the tendency for this logistic to be "reset" at the point which convexity changes from positive to negative. Technological advancement or shifts in perception or preference bring about shifts in the prior carrying capacity. Mathematically, there is a discontinuity in the derivative of the logistic growth. Essentially, the timetable of the proceeding logistic is reset to zero and a new one emerges with updated growth rates and limits.

These tendencies help describe the theoretical underpinnings but since the ultimate goal is facilitation of civil evolution, controls need to be developed for applying leverage to critical aspects. As mentioned, the process is moving away from a strict, mandated top-down approach and merging towards a bottom-up strategy. However, there exists a hierarchy of controls. One function of such controls is to identify critical points within the network and influence them within predetermined parameters to move the process towards established goals. The departure from the enforcement of control from a top down philosophy towards the perspective that urban structure embodies an evolution driven from the agents in a bottom up fashion.

One objective of controls is to alleviate what Riddell and Webber deemed "wicked problems." These are situations in which the problem and solution are often indistinguishable. The problem-solving environment contains no definitive test or stopping rule to determine if an optimal solution has been reached. Wicked problems lead to more wicked problems, infinitely. Drastic intervention is not always as straight forward as the removal of an identified undesirable result. Perhaps the removal an identified issue prevents the creation of an alternative that would pervade a future long ranging and more severe dilemma. If problems are not clearly understood, planners must realize the nature of complex evolution involves trial, error and random emergence (Batty, 2007). Traffic congestion is an example of wicked problems manifested by urban development. Roadways are “improved” to accommodate more traffic. During construction, congestion is increased or reassigned, contributing to congestion elsewhere in the system. Once completed, expanded roadways allow more traffic, inevitably leading to more congestion. These sorts of “first-world” problems are indicative of the current human condition. Although they result directly from social conventions, cause stress to millions and hinder the collective quality of life, such issues are allowed to fester since they are indicative consequences of the system at large. The impudence of this research is to suggest an approach by which social efforts capitalize upon the convenience our culture affords.

Concepts from biology, physics and cybernetics have been implemented to develop theories and models of urban development and social physics. The elemental view is that cities function as a system of components tied together through interactions. These interactions are based primarily upon the land use of a particular location and the economic and transport influences of the surrounding area. Historically, urban models have conceived patterns as existing fixed at a given point in time. Social systems were seen as existing in a state of equilibrium and progressing in a predictable manner rather than evolving over time.

In contrast, dynamic systems are out of equilibrium. In dynamic systems, change, improvement or accidents are often unanticipated and chaotic. Feedback based upon the interaction between agents and the environment impact the entire system. This mechanism describes the manner in which local actions contribute to global tendencies. Networks are based upon these interactions between agents, between agents and their environment and between locations. The characteristics of an interaction or network connection are indicative of the size and proximity of the agents and locations. The hierarchy of these objects is established by a rank order classification relative to the variable of interest.

Cities in equilibrium have activities arranged based upon the number of locations and the amount of activity at each location. The total number of arrangements is proportional to the entropy (information) of the system subject to relevant constraints. Social analogies led to formal models describing interaction between locations and attraction to locations. Many probability distributions are derived from these concepts and when coupled with the population densities associated with the specific locations facilitates the differential equations governing the system.

Distributions and densities are based upon the assumption that the area of each respective zone is equivalent. Therefore, by ranking the locations based upon populations yields a hierarchy creating the expected structure that there exist fewer zones of high density and more at lower densities due to competition. It can also be demonstrated that there is a negative exponential distribution describing the relationship of population, land value, trip densities relative to the distance from the center of the city.

Two-dimensional models can account for the relationship between multiple locations. The sum of the originating interactions is proportional to the probability to originating trips at that location likewise to the sum of departing interactions. Specific goals entail minimizing cost/distance from the central business district

(CBD). The two dimensional framework can be generalized to address connected pairs (origination and location) of locations. The model is scalable. Constraints are applied to consider trip costs and models can be scaled to real distances or populations.

Traditional urban modeling viewed situations as static since the physical structures were well formed and seemingly stable in the near future. Also, land use zones were well defined and concentric about the CBD. However, as our understanding of urban development and city systems are being enhanced at increasingly rapid rates, it is evident that development and thus, the associated models, are dynamic functions of time. The concepts of the equilibrium models are expanded to an exponential growth model in which growth is proportional to the population and dependent upon the rate of change. The growth rates themselves of cities are likely also functions of time.

These growth patterns are restricted due to carrying capacity and minimums of survival. Introducing such parameters curbs the exponential behavior and results in a logistic configuration. There also exist functional limits upon the other parameters. For example values in growth rates between 2 and 3 yield increasing chaotic behavior. Cost and frictional parameters have upper limits which, when reached, yield fluctuations in the economic network severe enough to significantly alter behavioral patterns. Additional lagged variables may also have ramifications brought about through previous time periods. Although this approach is more dynamic and considers space and time, it remains postulated from a top down perspective based upon aggregation into zones, regions etc. (Batty, 2008).

The next step from equilibrium to dynamic models is to deal with objects and individuals on smaller scales and simulate the process of decisions made based upon physical parameters and satisfaction principles. In this sense, the patterns of urban development are emergent from the participants or agent within the system. The background established in the approach to static models can be expanded with the change in population being expanded into two components. The first (reaction) is built on the concept that growth is dependent upon the population present at a given location. The second (diffusion) incorporates a rate of growth from neighbor influence. The combination of these along with a random component denoting the anomalies at given time periods form the basis of the reaction-diffusion equations.

Cellular Automata and Agent-Based Modeling

Segregation is the separation of agents or areas along respective lines of division. Individuals can be segregated age gender, age, income, geography or any other coded value. Area segregation, typically zoning, include divisions such as, residential, commercial, agricultural, industrial, open space and transportation. Within these divisions, clusters are formed of entities serving similar functions. In this context, segregation has no negative connotation but refers to the degree of physical grouping of locations or agents. Locations are segregated by the CA rules governing their state. Agents operate according to their mechanisms of choice or preference seeking a higher degree of satisfaction. Michael Batty and CASA offer an excellent model describing the dynamics of segregation at *casa.ucl.ac.uk*. The resulting actions tend to move agents to select locations and behaviors in accordance with the social segment best supporting their definition of satisfaction.

Modern models can be seen as an extension of the Schelling's "checker-board" model. Agents are programed with rules and behaviors, such as parameters governing "residential tipping". Residential tipping occurs when an agent's action results in a particular threshold being met. For example, in Schelling's model, there was a preference to be next to neighbors of similar preference. Once the ratio of similar/different neighbors falls below the desired level, the agent will change location in search of a higher level of satisfaction. Simulations are projected over time to demonstrate outcomes of aggregates. Representation is varied by the size and shape of neighborhoods and grid structure (von Neumann, Moore etc.). Agents tend to move based upon satisfaction and mobility opportunities. Schelling's contributions facilitated the pioneering of the field of Agent based modeling (ABM). Essentially, cells behave as agents when granted mobility and the "choice" to respond environmental circumstances.

ABM simulates individual agent activities and projects the system resulting from their actions. Linking ABM with GIS enables modelers to associate specific geographical information with respective or theoretical agents. The utilization of these two technologies allows for geometric detail to be assigned and thus allows for more accurate scalability and realistic models. ABM embodies the bottom up approach in similar fashion as CA. Whereas, CA build upon local space (areas), ABM build upon local activity. In a GIS format, real world infrastructures and inhabitants can be represented as features (points, lines, polygons) or as areas (rasters, Voronoi polygons, triangulated irregular networks (TIN)) or both. (Crooks, 2008)

The evolution of the model process includes the representation of space. Rather than that of zones, locations represent a specific cell consisting of a single household, individual, entity or cell state. Since these will be standardized relative to the field of interest, populations and or structures exist as true or false (1 or 0). Likewise, rates of change assume the values of -1, 0 or 1 depending upon if the population has migrated, remained unchanged or occupied. These switches are based on rules, constraints and thresholds comprising the structure of the cellular automata (CA). Given a cell (i), the state (presence or absence) will be influenced by the surrounding neighborhood (surrounding 4, 8 etc. cells). Setting a max threshold defines the rule that if contributing factors exceed the given value, then the value of the location (i) is 1 otherwise it is 0. In the strictest CA diffusion only occurs by influence of immediately adjacent cells. However, urban development of areas is often influenced by actions at a distance. The access to information, technology and mass transit provides resources beyond the immediate vicinity. Cell-Space or CS models consider this tendency. The modeling procedure remains similar to that of CA with the added caveat that migration occurs beyond the adjacent neighborhood.

An example of the concepts discussed above is the dynamic urban evolutionary model (DUEM) developed by Xie and Batty. The five states of land use are residential, industrial, commercial, transportation and vacant. Rules are implemented to govern which states can influences the onset of other states, the life cycle of particular states (ie. initiating, mature or declining). Thresholds are utilized to mark the onset of these state changes over spatial scales (neighborhoods, fields or regions). Neighborhood triggers are exercised dependent upon the street network. Densities and distances from relative origins exhibit influence on the field level. Constraints on the field level determine which cells are available for development. In this particular model, seeds are specifically or randomly assigned by declaring land use values of selected cells. Development then occurs based upon the rules and established guidelines. Due to the life cycles of the programmed land values, the space fills and then space is vacated as the carrying capacity is reached and then filled again and so on. Well-suited models emerge through the land use perspective (residential, commercial or industrial). Ultimate goals incorporate land use and transportation by building frameworks on the economic relationships between activities. Employment essentially is the connection between locations and agents. The population commutes to areas of intended land use as a function of the regional economy.

Employment can be divided into two categories. This first is "basic" in that these operations are conducted for the direct creation of goods or services to enter into the economy. "Non-basic" employment refers to sectors that provide for the population. Frictional and attraction parameters are incorporated to estimate values. This relationship leads to an iterative structure in which predicts populations from the basic economy the non-basic employment and so on. This approach can be tailored to different populations, employment types and sector interaction such as education, leisure and so on. Transportation is the real world manifestation of the employment network. Roadways present a physical imprint and commuting patterns reflect the human behavior. Four-stage transportation modeling (trip generation, distribution, modal split and assignment) can be added to enhance sophistication (Batty, 2008). This class of model is effective represented in www.urbansim.com.

The majority of CA models are geared to mimic urban sprawl. In addition efforts have been made to also address aspects of intra-urban change and segregation. In other attempts the concepts of classical mechanics, scale, hierarchy, 3-dimensional, complexity and self-organization have also been incorporated. Agent-based models typically focus upon small-scale transportation behaviors are beginning to link individual behaviors with spatial location. These sorts of approaches are burgeoning and lack standardization, consistent terminology or consensus regarding their best application. Although these models perhaps lack detailed urban morphology, they do replicate the bottom up process by which development evolves. Adding strict detail would hinder the speed by which these simulations are generated.

CA models of urban development are also able to capture the fractal nature of growth. Since they contain imbedded rules of behavior, the results are self-similar across variant scales. However, growth progresses with a degree of random uncertainty as the effect of “noise” and technological leaps alter the step-by-step progression, allocation of resources and capacity. These issues are the same impacting daily human interaction and quality of life. Hence, a more intimate connection between modeling processes and the systems they replicate can ultimately contribute to the establishment of more sustainable cities.

Conclusion

The bulk of this effort has addressed the history and theory behind models of urban development. The mathematical underpinnings are addressed in the next section. Subsequent steps in the process involve creating and implementing the survey. This data and historical records of development will be incorporated to estimate parameters and enhance the models. The final step is to merge the mathematic representation into an approach of policy development. Computer code is at the heart of policy reform and problem solving, since the math models are executed within the computer code. There are three major areas of human civilization, which can be unified through a centralized code structure. These areas are the human element, the technological element and the policy element.

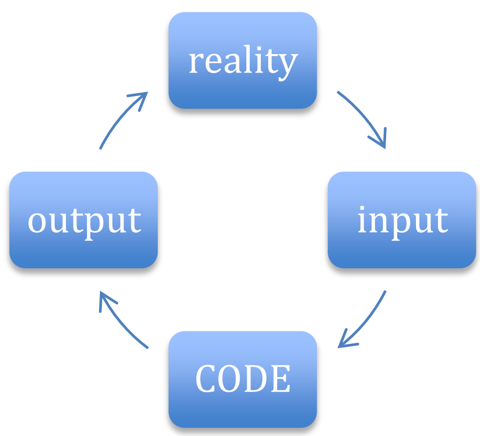

Figure 2. Sphere of Human Function

The human sphere considers the needs and functions of agents (people). Humans offer value to the modeling process as sensors, researchers and actors capable of implementing policy. Ultimately, it is the longevity and livelihood of these agents this effort intends to maximize. Agents behave with a quest of satisfaction. This can be simple sustenance or security for family or more sophisticated measures such as quality of life.

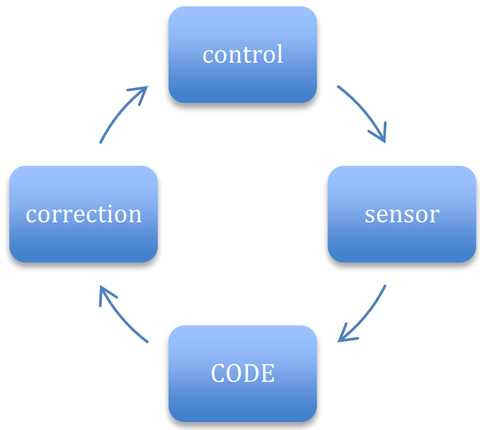

Figure 3. Sphere of Technology

The technology sphere refers to the mechanical and digital components designed to implement policies into civilization. These are the components that render data for processing. Through this, adjustments are made to sensor location or sensitivity. Corrections are adjustments made to the physical location or mechanical functioning of technology. Controls are the mechanical manipulation or software based changes occurring to launch, recall, adjust or otherwise alter the position or operation of technology. The technological aspects can be expanded to enforce guidelines of agents, objects or structure.

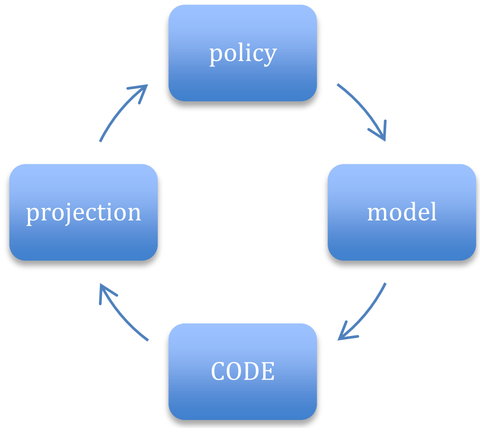

Figure 4. Sphere of Infrastructure

The infrastructure reflects the structures and policies, which will mark the legacy of human civilization. From the code, projections are made and policies can be developed. In the digital environment, regulations and standards can be embedded into operations of social systems. Projections generated from code operations represent the compilation of current data and processing. They signify more than simple output by incorporating time factors and multi-variant interaction to issue statistical predictions.

Sensor types have three classifications human interface, stationary and mobile. Hand-held GPS units and smart phones are the most powerful advancement in the human interface. Examples of stationary sensors are fixed or pivoting cameras, motion sensors etc. Mobile sensors are burgeoning because they enhance perceptive abilities and facilitate the abundance of implementation issues. Inherent with Human culture are the rules and policies, which serve to moderate behavior. They fall in ranges from specific laws and recourse to accepted norms, ingrained behaviors to fashionable trends. Government structure, safety parameters and cultural processes are examples. In general, policies include procedures and processes, which either through cultural evolution or physical limitations has been implemented on a civic basis.

Geo-social engineering assumes a scientifically objective approach. Although current situations, dilemmas and advantages are often indicative of the socio-economic and political factors, Geo-social engineering seeks solutions without regard to these mechanisms. This distinction is the difference between a top-down approach imparted by dominant hierarchies and a holistic bottom-up approach. As computer code is at the heart of this strategy, Geo-social engineering seeks to merge the technical code with an ethical code based upon empirical parameters rather than idealistic philosophy.

CHAPTER III

TECHNIQUE

Data Sources

Initial investigations are focused on the region spanned by the north central Texas council of governments (NCTCOG). The Dallas/Fort Worth area is well suited geographically. There are few, if any, mountains, natural water sources or other obstacles, which hinder development. Keeping scalability in mind, other metropolis areas such as Houston, Austin, San Antonio and Oklahoma City share economic proximity with few geographic impediments between. Distance is perhaps the most significant environmental factor. This fact is conducive to the modeling approach, which bases projections upon the radius of influence of economic entities. Given the geographical parameters of the area, the model looks at behaviors pertaining to significant questions of human existence.

Ultimate objectives include expanding the model to include multiple regions and countries. Historical data can be used for calibration, behavior description and identification of influence and commonalities. User input fuels the modeling mechanism. In this sense, the user describes a perspective of reality with analysis tools and modeling techniques. The three units selected generalize the cultural mechanisms present within the social context of human civilization. These are nutrition, economics and transportation. Essentially, the user perspective will be described in elemental terms of calories, dollars and miles - the three units of interest. Their significance should be noted now for the variables' relevance across every current, previous and future human civilization. The three American units selected generalize the cultural mechanisms present within the social context of human civilization. User inputs will be collected and compiled via digital surveys. Data points include events, which are associated with a location and time.

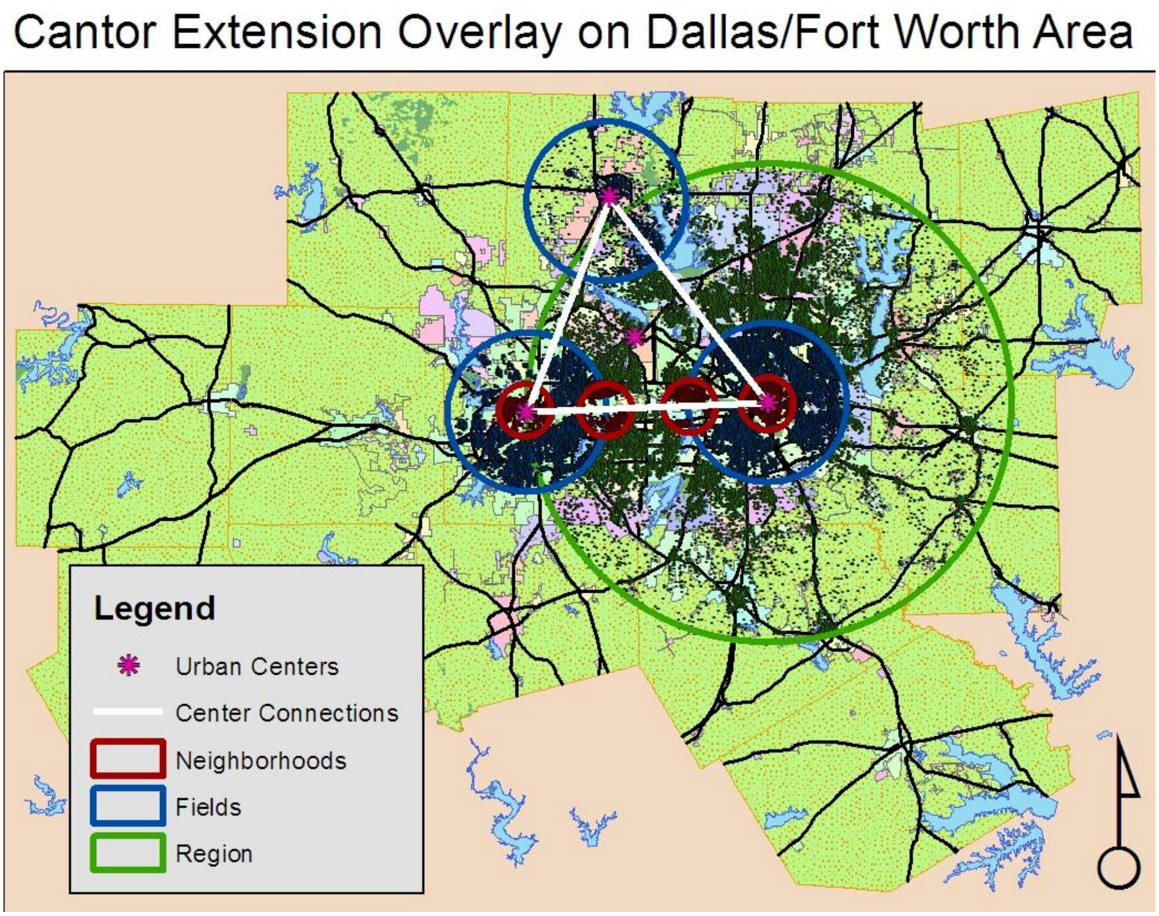

Figure 5. Model Geometry

The model is housed within a geographic information system (GIS) framework. The interface assimilates user inputs and makes projections based upon parameters input via survey or predetermined values. Applications will be created through the use of a GIS environment and imbedded coding language. Interface design allows user input through text, drop-down menus and input files. The three primary environments are databases, features and rasters. These titles refer to the sorts of files created via interface. The objectives of the individual interfaces are to be used in junction as a tool to create and project models over specified times and regions. The mathematical model is the principle component of the entire computer code. Whereas, the majority of functionality is dedicated to user interface and output display, the mathematical model is the system of equations whose operations ultimately dictates the ensuing experience and its impact on the environment and agents. There is a reciprocal evolution between the code and model in which the code design reflects the model and enhances the understanding, accuracy and effectiveness of the model.

Fractals

Fractal concepts serve two primary purposes concerning this research the first of these is as an iterative process by which cell values are estimated or simulated. The second purpose is related to the notion of fractal dimension. The dimension of a fractal is the exponent value describing the growth dynamics. For example the dimension of a linear fractal is $0 \leq D \leq 1$, an area filling fractal, $1 \leq D \leq 2$ and a volume filling fractal $2 \leq D \leq 3$. Essentially, the dimension indicates the extent the extent to which a fractal "fills" an available range. If a particular fractal describes an area and $D = 2$, then the entire area is filled by the fractal. Likewise, $D = 1$ implies a complete line and $D = 3$, a full volume.

Perhaps the best-known example of a fractal is the Mandelbrot Set (M), named for the mathematician Benoit Mandelbrot who is credited with popularizing the concepts fractals and their applications in mimicking processes in nature. The initiator of the set is defined as

$$Z_0 = 0$$

$$Z_1 = c$$

Where c = a + bi, such that (a) is the real part and (b) is the imaginary part of the complex number (c).

Hence, the generator of the set is defined as

$$Z_n = (z_{n-1})^2 + c$$

This recursion is applied until the iterative limit n = N is reached. The value of (z) is evaluated and if $z < 2$ for all $n \leq N$, then (c) remains in the set, M. (a) and (b) are incremented until all (a, b) coordinates are evaluated such that,

$$\text{Real min} \leq a \leq \text{Real max}$$

$$\text{Imaginary min} \leq b \leq \text{Imaginary max}$$

If c is in M after the iterative limit is reached, the cell graphically representing the position (a, b) is assigned a color (typically black). If $z_n > 2$ for some n, then $z_n > 2$ for all $m > n$. In this case, c is not in M. The graphic representation assigns a color-coded value associated with n, the iteration at which z_n is said to have escaped. All coordinates c which escape at the same iteration n are assigned the same color for graphic display.

The iteration behind the Mandelbrot set is similar to the model design presented in this paper. If the function grows relative to a radius of influence, area or density, it will be an iteration of a previous value in a space-time sequence. The resulting values will be checked in accordance to coded rules such as the life cycle of urban development. These factors will influence the probability of remaining in a current state or transitioning into the next phase in the development cycle.

Fractal Cities, by Michael Batty and Paul Longley at the Centre for Advanced Spatial Analysis (CASA), provides a research basis for fractal simulations. Batty describes the foundation of fractals and applies these concepts to models of urban regions and processes. Fractals are a geometrical repeating motif they are self-similar on any scale with components described as a hierarchy. Fractals are irregular, continuous but non-differentiable. There is inherent an initiator or geometric object used to start a fractal. While the generator is the motif repeated at every scale. Thus, fractals are constructed applying the generator to the initiator (Batty and Longley, 1994, 60). As a curve becomes more irregular and fills more of the two – dimensional space, the fractal dimension increases. Early fractal studies focused upon irregularities of coastlines. If its length is measured by a number (N) of unit lengths (r) Rugged coastlines have higher dimension than smooth coasts. The value of the fractal dimension has strong implications for the underlying processes of weathering and erosion, which dictate form.

Dimensions of coastlines:

Britian (Koch)	1.262
Australia	1.13
South Africa	1.02

A dimension above 1.5 tends toward self-intersection (Fractals, 102).

To calculate the dimension of an associated length

$$N_k = (r_k^{-1})^{D} = r_k^{-D}$$

$$D_k = \log N_k / \log (1/r_k)$$

D – dimension

N - # of parts

L – length of parts

r – scale, proportion along line where cut is made, i.e. r = 2: ½ L segments

Samples of N and L are taken to establish a regression for different scales r:

$$L_k = N_k r_k = r_k^{(1-D)}$$

If the fractal generator is self-similar then the dimension will be same for all (k).

Dropping the index and incorporating a proportional multiplier (G) yields

$$N(r) = Kr^{-D}$$

$$L(r) = N(r)r = Kr^{(1-D)}$$

$$\text{Log } L(r) = \log K + (1\text{-}D) \log r$$

(Batty and Longley, 1994, 66-69)

Let (R) be an object of relative size. In this case an area. Then the impact of (r) will be proportional to the size of (R).

$$L(R/r) = K(R/r)^D r = Kr^{(1-D)}R^D$$

$$L_i = GR_i^D$$

$$G = Kr^{(1-D)} \text{ and } R_i = A_i^{(1/2)}$$

Then the relationship between perimeter and area becomes

$$L_i = GA_i^{(1/2)}$$

$$\text{Log } L_i = \log G + (D/2) \log A_i$$

This estimate of dimension is often used for physical objects such as clouds, craters or islands. With a fixed (r) and z=(k/r) the dimension is dependent upon the (R)

$$N(R) = K(R/r)^{D}$$

$$N(R) = z\,R^{D}$$

Produces the density relation

$$p(R) = (N(R)/A(R)) = z\,(R^{D}/piR^{2}) \sim R^{D-2}$$

(Batty and Longley, 1994, 70)

As more sophisticated models are built, they will consider both radius and physical or economic size of objects. The above model can be applied directly towards the edges of development where transitions and new development are most likely. In this sense, the relationships of perimeter, area and density can describe the zonal barriers between land-use, industry sectors, demographic values etc.

The fractal utilized in this model is an extension of the Cantor Set. The formal definition of the Cantor Set is available in the appendix. It is constructed on the domain [0, 1] (initiator) and proceeds by removing the middle open interval one-third the initial length (generator). The cantor set consists of all values, which are never removed. The dimension of this set is

$$D = \log(N)/\log(r^{-1})$$

$$D = \log(2)/\log(3)$$

$$D = .6309$$

Where (N) is the number of parts generated at a subsequent level, in this case 2 and (r) is the relative length of each segment, in this case (1/3).

The extension of the Cantor set incorporates properties of urban development in an effort to create a generative process indicative of the way cities grow. Spatial considerations must be made to make the step between simple geometric fractals to a model with more realistic behavior. Likewise, land use types are classified and display different tendencies, thus also impacting the shape of cities.

The von Thunen bid-rent principles state that there exists a relationship between real estate value and the distance from the relative economic center or Central Business District (CBD). The theory was developed in relation to crops and the rate of spoilage versus transportation and refrigeration costs (Batty and Longley, 1994, 123).

Consider three key urban activities, which are 1:1 with land use

Commercial – industrial (u=1)

Residential – housing (u=2)

Open space – recreational (u=3)

P(u,r) is the probability of activity (u) occurring at distance (r), the radius from the CBD.

$$p(u,r) = a(u) + b(u)\,(r - R(u)),\ u = 1,2,3$$

Where a(u), b(u) and R(u) are parameters controlling the distribution of land use. Changing these parameter values can reform space of all cities – the shape of a city (Batty and Longley, 1994, 131). If these are constants, then there is a deterministic piece-wise function of linear inequalities describing the likelihood of activity. Additional economic functions and housing standards can be implemented. For example, u = 1,2,3,4,5,6, considers basic and non-basic economic sources and three level of housing density. More realistic behavior is obtained by development occurring with a given probability rather the deterministic approach.

Let P(u,r,s) be the probability of activity type (u) occurring at radius (r) on hierarchy level (s), p(u,r,s) is the predicated probability prior to iteration.

$$P(u,r,s) = [1 + p(u,r,s-1)]^{v(u)}p(u,r,s)$$

The most realistic simulation achieved by Batty with

$$[v\ (u) = 1 \text{ for all } u],\ s = 4$$

Based on fractal perturbation, randomness in land use allocation and very slight hierarchical conditioning where s = level of hierarchical simulation. These versions lack any real calibration or behaviors indicative of real parameters. However, these foundations provide leverage with which other more detailed elements may be incorporated.

Thus the influence of hierarchy is employed, where v(u) = 0 implies no influence from the prior level and yields the classic Von Thunen rings (Batty and Longley, 1994, 135-137). This begins the cycle between model development and understanding the reality the model describes. This is a cyclical process of induction, building general theories from the bottom up, and deduction, where these universal are used to predict specifics.

Von Thunen's concentric bid-rent model is a foundation of describing cities as economic sources. Bid-rent theory postulates tradeoff between housing space and type versus proximity to central urban functions, but only considers a single, mono-centric metropolis. Dwellings and neighborhood types are clearly related to distance from the CBD and time. As these models evolve to describe more dynamic systems, time will be incorporated as a critical variable for development, land-use and agent behavior. A cities growth, success and survival depend upon the logistic viability of the flow of goods, services, capital, population and information in and around the city. More realistically, the entire market influences this logistic efficiency with those economic sources closest in proximity being most closely related. Walter Christaller's Central place theory considers multiple economic centers and the opportunities they provide the population. Depending upon the element to be optimized, rules involving market area, range of goods, population served, variety of goods and services are considered. A tessellation of regular hexagons is formed around the city centers. The hexagonal network describes the areas served by particular services, with larger cities having a proportional larger area of influence (Batty and Longley, 1994, 55).

The central place theory implies a hierarchical structure inherent in its design. The hierarchy of land-use has also been introduced. It will also be essential to identify the hierarchy of market area, transportation routes and population centers. These structures could include traffic routes, public and private services, business organization from regional to local classification. The model explains spatial structure inductively using clustering methods as simulation begins with top level by subdivision and fractal rendering around centers and activities; lower levels depend on upper levels (Batty and Longley, 1994, 141). Parameters are updated and the process repeated.

As time, land use and location are implemented into the description of the process of urban development, a process of aging and renewal according to simple rules controls life cycles (Batty and Longley, page 161). Agents are given choices and behaviors and have impact on the landscape as they move through. The hierarchical structures lend themselves to sequential and nested approaches involving entropy models. Regular, non-random fractal patterns from cell space models can be used as initial approximations from which cellular automata (CA) and artificial life models are built (Batty and Longley, 1994, 162). Likewise, random distributions, survey data or projections can be used as seed values from which iterations are run.

Model Elements

At this point, the generalized concepts are refined and built upon. The following model is taken primarily from *Cities and Complexity* (Batty, 2005). In contrast to the text, the model is consolidated. Notation is updated to reduce conflicts in variables and functions. Basic fractal relationships involve the size of an object (x) and the frequency (y) with which a particular size is observed. If (α) is a scaling parameter then,

$$y = x^{\alpha}$$

And

$$dy/dx = \alpha\, x^{(\alpha - 1)} = \alpha\, (y/x)$$

Or

$$f(x) = x^{\alpha}$$

$$f_1(x) = \alpha\, x^{(\alpha-1)} = \alpha\, f(x)/x$$

So, (α) can be interpreted as the proportion of percent change between (x) and f(x). If (x) is scaled by constant (c), where x' = cx, then

$$f'(x) = (x')^{\alpha} = (c\, x)^{\alpha} = c^{\alpha}\, f(x)$$

Consider the observations to be a population (p) of a city, or other geographical classification, and f(p) to be the frequency of the observation. Let (β) be the scaling exponent such that

$$F(p) = p$$

(β) can be estimated by fitting the above equation to data. Hence, f(p) is a probability density function based on a power law. The cumulative frequency (F(p)) for which F(p) is the proportion of values less than a population (p) and

$$F(P>p) = 1 – F(p) = P ^(1-β)$$

This counter cumulative function (Pareto's law) describes the tendency for the number of observations to be inversely proportional to the size of the value. If events are ordered via a descending hierarchy, the number of events greater than a value (p) is the rank order (r) where the largest value of (P) is associated with r = 1. If,

$$r = P ^(1-β)$$

Then

$$P = r ^[1/(1-β)]$$

Zipf's law applies this principle for city sizes and claims that the exponent

$$1/(1-β) = -1 => β = 2$$

The population as a function of its rank is therefore

$$P(r) = P(1)/r$$

(Batty, 2005, 462)

Zipf's Law, modified by Mandelbrot, yields the q-exponential

$$P(r) = P(1)/[(c+r)^1/(1-β)]$$

Gibrat's law, which is often used to describe the growth of business firms, assumes that growth is proportionate to the size. However, the rate of growth is an average rate (λ) applicable to all members of the classification. Each member has an associated deviation ε(t,i) which is normally distributed. So,

$$P(t, i) = \lambda\ P(t-1,i) + \varepsilon(t,i)\ P(t-1,i)$$

$$P(t,i) = [\lambda + \varepsilon(t,i)]\ P(t-1,i)$$

With initial conditions, t = 0 and P(0,i),

$$P(t,i) = P(0,i)[\ \lambda + \varepsilon(t,i)]\ [\lambda + \varepsilon(t-1,i)]\ \ldots\ [\lambda + \varepsilon(1,i)]$$

(Batty, 2005, 466)

Starting point: P(t,i) at time (t) in cell (i) for a uniform distribution at t = 0, P(0) = 1 for all (i).

$$P(t+1,i) = \lambda(t,i)\ P(t,i) = [1+ \varepsilon(t,i)]\ P(t,i)$$

The deviation from λ(t,i) is

$$\varepsilon(t,i) = z(t,i)\ \text{random}(\delta)$$

(δ) is the max bound on growth and z(t,i) is a switch

$$z(t,i) = \text{random}\{1,-1\}$$

To add local diffusion within the 8 cell Moore neighborhood Ω (i), a proportion (YY) of diffused population then,

$$P(t+1,i) = \lambda\ (t,i)\ (1-YY)\ P(t,i) + YY/8 \sum_{m\epsilon\Omega i} P(t, m)$$

With truncation, such that (w) is a minimum threshold below which an area is no longer defined as a city.

If,

$$P(t,i) < w \quad \text{(cuts off tail)}$$

Then

$$P(t+1,i) = w = P(0,i) = 1$$

(Batty, 2005, 470)

Introduction of Epidemic

The next step is to create a systematic model capable of handling development at varying stages. The classic differential model, which simulates the movement of disease through a population, is modified for this purpose. In the differential model, the population is considered to exist in one of three stages; susceptible, infected or recovered. In this initial description, once a member is recovered, it is no longer susceptible. The translation considers the areas "susceptible" to development as undeveloped or available. The "infected" are areas of new development, which have recently transitioned from undeveloped. Typically, this is indicative of cells that border areas, which have not yet been developed. The "recovered" are those which are areas of mature development. These are cells surrounded on all sides by development.

U(t) – undeveloped, available land

K(t) – new development (recent transition)

P(t) – established (development surrounded)

Development is divided into 2 constituent parts, new and established (mature). Assuming space/resources are capacitated, the is limit (C) is defined as

$$P(t) + K(t) + U(t) = C$$

for all t, t= 0, 1, 2, …T

And of course

$$dP(t)/dt + dK(t)/dt + dU(t)/dt = 0$$

The one directional process implies

$$U(t) \rightarrow K(t) \rightarrow P(t)$$

So,

$$A(0) = C - E,\ N(0) = \varepsilon(0,i)$$

If, dk(t)/dt represents the addition of new development, then

$$dK(t)/dt = dk(t)/dt - dP(t)/dt$$

And, if (α) and (γ) are proportional coefficients of new and established development respectively, then

$$dK(t)/dt = \alpha\ K(t)\ U(t) - \gamma\ K(t)$$

$$dP(t)/dt = \gamma\ K(t)$$

And

$$dU(t)/dt = \alpha\ K(t)\ [U(t) - p]$$

Where

$$p = \gamma/\alpha$$

A(0) > p, serves as a minimum threshold for initiating civilization.

(Batty, 2005, 390-2)

(p) is a relative removal rate

Threshold epidemic

$$dU(t)/dP(t) = -(\alpha/\gamma)\ U(t) = -U(t)/p$$

Steady state

$$A(inf) = A(0)\ \exp(-p(inf)/p) = A(0)\ \exp\ [(A(inf) - C)/p]$$

$$\text{Int}\,[A(\text{inf})] = N(t) = C / \{1 + [c/N(0) - 1) \exp(-\alpha\, C\, t)]\}$$

If capacity is a function of developed land

$$dC(t)/dt = \beta\, P(t)$$

where β = rate of growth in capacity and

$$dU(t)/dt = -dn(t)/dt + d\,C(t)/dt = -\alpha\, K(t)\, U(t) + \beta\, P(t)$$

Introducing a function which distinguishes available land as suburban fringe, F(t) and periphery land relative to the core, A(t), so,

$$P(t) + K(t) + F(t) + U(t) = C$$

$$dF(t)/dt = \mu\, dK(t)/dt$$

where (μ) is the transition parameter so,

$$dU(t)/dt = dk(t)/dt - dF(t)dt$$

(Batty, 2005, 395, 7)

If developed land is allowed to cycle back into available land and (λ) is the rate at which established development converts to available land,

$$P_1(t) = \gamma\, K(t) - \lambda\, P(t)$$

and

$$A_1(t) = \lambda\, P(t) - \alpha\, K(t)\, U(t)$$

then

$$P_1(t) = \gamma\, K(t)$$

$$n_1(t) = \alpha\, K(t)\, U(t)$$

$$K_1(t) = k_1(t) - P_1(t)$$

$$f_1(t) = \sigma\, F(t)\, U(t)$$

$$U_1(t) = -\, f_1(t)$$

(Batty, 2005, 398-9)

Current assumptions include

(1) Available land is not a constraint

(2) Available land an new development are generated via diffusion

Including spatial coordinate (x,y) yield the notation for components as

U(t, x, y) – undeveloped, available land

K(t, x, y) – new (initiating) development

P(t, x, y) – established (mature) development

Spatial Epidemic

(α) and (γ) are growth parameters and (ξ^U) is a diffusion coefficient for undeveloped land.

$$K_1(t, x, y) = \alpha\, K(t, x, y)\, U(t, x, y) - \gamma\, K(t, x, y)$$

$$U_1(t, x, y) = \alpha\, K(t, x, y)\, U(t, x, y) + \xi^U \nabla^2 K(t, x, y)$$

$$P_1(t, x, y) = \gamma\, K(t, x, y)$$

(Batty, 2005, 402)

A computable form of the model is a discrete version based on CA, where

$$U(t, x, y) + K(t, x, y) + P(t, x, y) = 1$$

Cells made available for development if

$$K(t, x, y) = 1$$

Then

$$U(t+1, i, j) = 1$$

Where i = x +- 1, j = y +- 1

Unless

$$K(t, I, j) = 1 \text{ or } P(t, I, j) = 1$$

Available Land to New Development

Consider the following conditions,

$$U(t, x, y) = 1 \text{ then } K(t, x, y) =1 \text{ and } U(t+1, x, y) = 0$$

Thus, the term

$$\alpha\, K(t, x, y)\, U(t, x, y)$$

is translated into a simple un-weighted transfer. Introducing a random transfer, if

$$U(t, x, y) = 1$$

and the threshold of probability of development is exceeding

Then

$$K(t, x, y) = 1 \text{ and } U(t+1, x, y) = 0$$

A cell transfers from new to established development (matures) when it is no longer adjacent to available land (i.e., sum over Moore neighborhood is 8).

Then

$$P(t, x, y) = 1 \text{ and } K(t, x, y) = 0$$

(Batty, 2005, 404)

Life Cycle

Incorporating the life cycle of established development into the continuous spatial model yields

$$P_1(t, x, y) = \gamma\, K(t, x, y) - \lambda\, P(t, x, y)$$

And

$$U(t, x, y) = -\alpha\, K(t, x, y)\, U(t, x, y) + \xi^{U}\, \nabla^{2}\, N(t,x, y) + \lambda\, P(t, x,y)$$

Let a(i) = time of initiation of each cell at (x, y). a(i) = t is set at the time when a cell becomes newly developed or activity initiates. To account for the life cycle of development, replace (t) with t - τ(i) where τ(i) is an age limit.

If t – a(i) = τ(i), then

$$U(t, x, y) = 1$$

And

$$P(t+1, x, y) = 0$$

(Batty, 2005, 410)

Time Based States of Cells - Life Cycle of Development

Activities at location (i) and time (t) are defined as A(t,i) and can assume temporal states

A(k,t,i) – initiating

A(m,t,i) – mature

A(d,t,i) – declining

Such that

$$A(t,i) = A(k,t,i) + A(m,t,i) + A(d,t,i)$$

State change depends upon time factors affecting the exponential decay in likelihood of remaining in a given state. p(k,t,i), p(m,t,i) and p(d,t,i) are the respective probabilities of being in the above states.

Assume activity begins at time

$$\tau(i) = t$$

Once originated, the probability of the life cycle is

$$p(k,t,i) + p(m,t,i) + p(d,t,i) = 1$$

For $t > \tau(i)$

The probability that development remains in the initiating state is

$$p(k,t,i) = \gamma \exp(-\gamma (t - axy)$$

(γ) is a constant of proportionality that normalizes the distribution and (γ) is the parameter reflecting the rate of decay from initiating to mature activity.

$$p(m,t,i) = M (1 - p(k,t,i) * p(d,t,i)$$

M is a similar normalizing constant and

$$p(d,t,i) = \exp [- \mu(t - \tau(i))]$$

Where (μ) is a decay parameter and

$$p(p,t,i) = 1 - p(k,t,i) - p(m,t,i)$$

Essentially the state range is a random (or determined) number (ran) between [0,1] such that

If ran < p(k,t,i)

Then activity is initiating

If p(k,t,i) < ran < p(m,t,i)

Then the activity is mature

If p(m,t,i) < ran

Then the activity is declining

(Batty, 2005, 183)

Vacancy

Fractal holes can be represented via vacant land, where $V(t) = 1$, if a cell is not in the active development cycle for some reason, otherwise $V(t) = 0$.

So,

$$V(t, x, y) + U(t, x, y) + K(t, x, y) + P(t, x, y) = 0$$

I.e. begin simulation with initial conditions:

$$K(0, 0, 0) = 1$$

$$K(0, x, y) = 0, x/= 0, y/= 0$$

$$U(0, x, y) = 0$$

$$V(0, x, y) = 0$$

$$P(0, x,y) = 0, \text{ for all } x, y$$

1. Age threshold is checked and if development has reached age to be renewed, then it is returned to pool of available land

    If $t - a(i) \geq \tau(i)$

    Then $U(t, x, y) = 1$ and $P(t+1, x, y) = 0$

2. Determine whether or not available land comes onto the market.

    If $K(t, x, y) = 1$ and $K(t, I, j) \neq 1$ and $P(t, i, j) \neq 1$,

    Then $U(t+1, i, j) = 1$, where $i = x+- 1$, $j = x+- 1$

3. If land is available and or already vacant, it becomes or remains vacant if a probability threshold $\mathcal{G}$ is exceeded.

    If $U(t, x, y) = 1$ or $V(t, x, y) = 1$ and Ran > $\mathcal{G}$

    Then $V(t+1, x,y) = 1$

4. The neighborhood constraint, which determines if $K(t, x, y) \rightarrow P(t, x, y)$ is augmented with the vacant state.

    If Sum(Moore) = 8

    Then $P(t, x, y) =1$ and $K(t, x, y) = 0$

5. The transition from available to new development is made and the age of development is determined.

If $U(t, x, y) = 1$ and $Ran > \phi$

Then $K(t, x, y) = 1$, $a(i) = t$, $V(t, x, y) = 0$ and $U(t+1, x, y) = 0$

(ϕ) is the spread parameter and slows development as the parameter is increased.

($\mathcal{G}$) is the vacancy parameter and is the threshold for which land becomes or remains vacant (Batty, 2005, 415-6).

Cantor Extension

The Cantor Middle Third fractal serves as the basis for the area filling function, which simulates development. The segments are considered a radius, which sweeps an area centered at the "outside" endpoint.

Spatial scales or levels of urban development are defined as

$\Omega 1$ – region

$\Omega 2$ – field

$\Omega 3$ – neighborhood

The radius (r) of affected area is that contained in the particular level of point (i).

Influence declines via a negative exponential dependent upon r(i,j)

$$\Pi(t,i,j) = Q \exp(-\alpha\, r(i,j))$$

Where (Q) is a scaling coefficient and (α) a decay rate relative to (r).

A "free length" parameter can describe an activity's influence outside of its respective level.

Direction of Development

The directional probability reflects the choice of heading $\theta(z)$, z = 1, 2, …, 8 (Moore neighborhood) with weight w(z). So the probability of development growing in direction (x) from point (i) is

$$p(x|i) = w(z)\theta(z)/\sum_z[w(z)\,\theta(z)]$$

Thus, the probability of a particular cell (j) being chosen

$$p(t,i,j) = f\{\,\Pi(t,i,j),\, p(x|i)\}$$ (Batty, 2005, 184)

Tests of development

Function f(*) depends upon 3 conditions being met.

(1) Regional regulations are obeyed ($\Omega1$)

(2) Life cycle dynamics or vacancy statutes ($\Omega2$)

(3) Density thresholds regarding ($\Omega3$)

If the above tests are passed then the cell at location j may be added to the current stock give by

$$A(t+1) = \sum_j A[n,t,j] + \sum_j A[k,t,j] + \sum_j A[m,t,j] - \sum_j A[d,t,j]$$

This is a typical birth, death and survivorship equation. Change is given by

$$\Delta A(t+1) = A(t+1) - A(t) = \sum_i A[n,t,i] - \sum_i A[d,t,i]$$

With a growth rate

$$A'(t+1) = A(t+1)/A(t) = 1 - \Delta A(t+1)/A(t)$$

And growth of initiating development as

$$\lambda(t+1) = \sum_i A[k,t,i]/A(t+1)$$

Development Divided into Industry Sectors

Economies can be modeled as the interaction between basic (primary, secondary) and non-basic (tertiary) resources. For example, primary resources are a natural occurrence such as fossil fuels or agricultural advantage. Secondary resources are the industrial components resulting from the processing of the natural (primary) resource. And tertiary resources are the commercial goods and services resulting from the infrastructure that are necessary to support the population and production of primary and secondary resources. Basic economies (B) are a composite of the primary (B1) and secondary (B2) economic sources. The sum of these and non-basic (S) economic sources make up the total economic base of a region (E), where

$$E = B + S$$

The population (P) grows by a certain proportion (a) of the economy, or

$$P = aE$$

Which iterates to supplement the service based economy by a proportional factor (b).

$$S = bP$$

Manipulation yields

$$E = B(1\text{-}ba)^{-1} \text{ and } P = bP(1\text{-}ba)^{-1}$$

(Batty, 2005, 46)

This simple system can be expanded by incorporating multiple variables (location etc.) and implementing thresholds and check points.

The three potentials of development are

P(t,i) – population

E(t,i) – manufacturing industry

S(t,i) – services

And

$$P(t+1,i) = P(t,i) + \theta^{P} \nabla^{2} P(t,i) + \sum_{i} wA(p,t,i)\ oA(t,i) + \varepsilon(p,t,i)$$

$$E(t+1,i) = E(t,i) + \theta^{I} \nabla^{2} P(t,i) + \sum_{i} wA(e,t,i)\ oA(t,i) + \varepsilon(e,t,i)$$

$$S(t+1,i) = S(t,i) + \theta^{S} \nabla^{2} P(t,i) + \sum_{i} wA(s,t,i)\ oA(t,i) + \varepsilon(s,t,i)$$

The respective functions relate the following parameters

θ – diffusion coefficients

wA – weight on importance of addition of activity

oA – represents the new walker at location (i)

ε – error, noise function

(Batty, 2005, 248)

Adding Sectors or Zones

The above models consider a single activity type. However, more accurate versions include land use associated with location and activity. The (A) function will be reserved to represent Agents and the land use activities will be classified as:

P – housing or residential population

E – employment or industry

S – service or retail

T – transport, trails (change from author's notation)

V – vacant (Batty, 2005, 186)

Once initial tests are passed and thus qualifying a location for development, transition probabilities of given weights dictate the likelihood of a particular activity continues or transitions in another activity type (transition matrix). Transport rules prevent large numbers of adjacent cells devoted solely to transport.

For land uses for activity A at location (i) during time (t) are

	Population	Economic	Service	Transport	Vacant
	P(t,i)	E(t,i)	S(t,i)	T(t,i)	V(t,i)
initiating	P(k,t,i)	E(k,t,i)	S(k,t,i)	T(k,t,i)	
mature	P(m,t,i)	E(m,t,i)	S(m,t,i)	T(m,t,i)	
declining	P(d,t,i)	E(d,t,i)	S(d,t,i)	T(d,t,i)	

Growth rates of sectors q = {P, E, S, T}

$$P'(t+1) = P(t+1)/P(t)$$

$$A'(t+1) = E(t+1)/E(t)$$

$$S'(t+1) = S(t+1)/S(t)$$

$$T'(t+1) = T(t+1)/T(t)$$

Growth rates for initiating activity

$$\lambda P(k,t+1) = P(k,t+1)/P(t)$$

$$\lambda E(k,t+1) = E(k,t+1)/E(t)$$

$$\lambda S(k,t+1) = S(k,t+1)/S(t)$$

$$\lambda T(k,t+1) = T(k,t+1)/T(t)$$

(Batty, 2005, 190-2)

Agent/Landscape Interaction

The function A(*) represents an agent of type (k), during time (t) and at location (i). The use of agents implies the introduction agent-based models. The goal is to integrate these agents into the CA environment. Where cell states are determined by time based rules and diffusion from respective neighborhoods. Agents make decisions based upon environmental clues. The nature of the landscape indicates the likelihood of activities occurring at a given location. In this sense, the landscape influences agents' decision process. In doing so, the agent also writes to the landscape, thus providing information for future agents.

Agents: $A(k,t,i)$, $k = 1,2, \ldots, M$

M = total # of agents, index h

W = total # of landscapes, index w

N = total # of cells, index i

Locations: $L(w,t,i)$, $w = 1, 2, \ldots, W$, $i = 1, 2, \ldots, N$

The most simple relation between A <-> L

$$A(h,t+1,j) = f\{A(h,t,i), L(w,t,i)\}$$

$$L(w,t+1,j) = g\{L(w,t,i), A(k,t,i)\}$$

A typical hierarchy consisting of routine decisions (L -> A) has one-way influence where agent choices are based strictly on landscape options. There are 16 possible total systems, when considering the combination of all possible interactions from agent to agent, agent to landscape, landscape to agent and landscape to landscape. Also exhibiting influence are other agents and landscapes, AE and LE.

$$A(h,t+1,j) = f\{A(h,t,i), \textstyle\sum_{n \neq h} \circ\, [A(n,t,m)], L(w,t,i), AE\}$$

$$L(w,t+1,j) = g\{L(w,t,i), \sum_{\ell \neq w} \circ [L(\ell, t, m)], A(h,t,i), LE\}$$

Where (m) is a location which may be different than locations (i) and (j) and (∘) indicates the influence of agents and landscapes. The impact is not necessarily simple summation (Batty, 2005, 212).

A "landscaping rule" is the manner a walker changes the landscape.

The landscape altered at location (i) and time (t+l)

$$P(t+1, i) = P(t, i) + \mu[i, A(t, i)]$$

Where A(t,i) is the agent at position (j) at time (t). μ(t) is a function measuring the impact of the distance from (i) to (j).

A "stepping rule" is the manner a walker responds to the landscape.

Agent moves by defined rules for example in direction of minimum gradient ∇ P(t,i)

$$A(t+1,m) \leftarrow \min_j \nabla P(t,j)$$

Then

$$P(t+1,i) = P(t,i) + \sqrt{\alpha\, d(t, i, j)}$$

α is a scaling constant of the distance between (i) and (j).

Adding diffusion and a randomly distributed noise parameter ε(t, i) yields

$$P(t+1,i) = P(t,i) + \phi\nabla^2 P(t,i) + \sum_j \sqrt{\alpha\, d(t, i, j)} + \varepsilon(t, i)$$

$$A_m(t+1) \leftarrow \min_j \{\nabla P(t,j), \varepsilon(t, i)\}$$

(Batty, 2005, 216)

Interaction between Agents and Resources

Form and function are united via human movement through and interaction with the environment. Imagine that agents search the landscapes for resources. Doing so, they leave trails with contextual evidence that may be interpreted by others also moving through the same area with similar intentions. Their exploration begins at an origin. Depending upon their preferences and properties, they succeed or fail and consequences are rendered. For example, assume that the journey continues until resources are located and then the agent returns to its origin. Fixed origins are denoted O(i) with (i) representing the location at which an activity originates.

R(j) is the resource located at the coordinate associated with location (j). Agents are A(h,t,m) at location m of type (h).

$\overrightarrow{A}$(h,t,m) is the agent described above in the exploration mode.

$\overleftarrow{A}$(h,t,m) is the returning mode

Agents move through their environment by

$$\overrightarrow{A}(h,t,m) = \max_i\{\nabla\ T_i(t), \varepsilon(k,t,i)\}$$

and switches modes of search when

$$\overrightarrow{A}(h,t,m) \Leftrightarrow R_j$$

then, $\overleftarrow{A}$(h,t,m) return behavior is described as

$$\overleftarrow{A}(h,t,m) \leftarrow \mu(O(i), m)$$

where μ is the return function from m -> O(i) and m = j.

(Batty, 2005, 226)

T(t, i) is the trail landscape. Agent movement provides feedback to the trail landscape by function Φ, such that

$$T(t+1,m) = T(t,m) + \Phi \sum_h [\overleftarrow{A}(h,t,m)]$$

Agents are reset to exploratory mode once returning to the origin by

$$\overleftarrow{A}(h,t,m) \Leftrightarrow O(i)$$

then the respective agent switches back to exploratory mode $\overrightarrow{A}$(h,t,m).

(Batty, 2005, 228)

Probabilistic Interaction between Agents and Resources

A probabilistic association between resources and agent parameters can reflect agent choice. Mode ($\overrightarrow{A}$, $\overleftarrow{A}$), type (h), time (t), location (m) and all other applicable variables can be considered in the decision process depending upon sophistication. Decision mechanisms can be elaborated to reflect more accurate agent behavior. For example purposes, a simple version is introduced, by which consumption occurs with a certain probability once the resource is encountered. If the resource is not consumed (chosen), then the agent continues on its search. Consider (p), the probability of consumption and

$$p_1 = p < 1$$

While (q) is the probability of continuing to search after the resource is encountered and

$$q_1 = 1 - p_1$$

Assume resources are encountered in a space-time sequence defined as

$$z = 1, 2, 3, \ldots$$

The probability of consumption at z+1 is

$$p_{z+1} = q_z p$$

And continuing

$$q_{z+1} = q_z - p_{z+1} = q_z - q_z p = (1-p)^{z+1}$$

Or

$$p = (q_{z+1} - q_z)/q_z$$

And as z+1 -> Z

$$Ln(q_z) = -pz + G$$

Therefore, the probability of continuing is

$$q_z = G \exp(-pz)$$

And the probability of consuming

$$p_z = p\, G \exp(-pz)$$

If,

$$p^k \leq Ran \text{ and } \overrightarrow{A}(h,t,m) \Leftrightarrow R_m$$

Then $\overleftarrow{A}(h,t,m)$, where p^k is a type specific interaction probability and Ran is a random number 0 < Ran < 1 (Batty, 2005, 234).

If R(z,j) is a measure of size of the resource then, by substituting p R(z,j) for p_z, the above method can be used in entropy maximizing models, such as the intervening opportunities model in a Chicago Transportation study (Batty, 2005,235).

Reassign Origin

Origins can be reassigned to agents via an agglomeration algorithm where LL is a threshold for which if

$$\sum_{i \in \Omega m} A(h, t, i) \geq LL$$

Then ih -> mh, where (->) describes the change from origin (i) to (m).

If the above process is applied to agent interaction, it follows the theories of social agent models. Accuracy is improved by selection of origin. Initially, Origins can be uniformly distributed. Eventual models incorporate a selection based upon social mass. Attraction is a function of selecting a max of

$$P(t+1,i) = P(t,i) + \theta \nabla^2 P(t,i) + O(t,i) + \varepsilon(t,i)$$

Where θ is a diffusion coefficient O(t,i) is the location of origin for the new walker and

$$O(t+1,m) = \max_i \{P(t,i)\}$$

(Batty, 2005, 244-5)

Agent Models of Societies

If the assumption of infinite resources is relaxed, the agents have incentive to accumulate resources as wealth. Success depends upon the time frame within which resources are discovered and consumed. Thus, agents "survive" upon existing wealth until more resources are located or recycled. Regeneration of resources occurs at a rate that may or may not sustain behavior patterns of agents (Batty, 2005, 252).

The model below does not assume agents return to their origin. Rather they consume resources and accumulate wealth until a given conservation threshold (RR) is reached.

Let $R_{j^*}(t)$ be the resources located in the vicinity of location (j), such that

$$j^* \text{ is in neighborhood } \Omega(j)$$

Agent location is then modeled as

$$A(k,t,m) = \max_{j^* \epsilon \Omega j}\{\nabla R(t, j), \varepsilon(k, t, i)\}$$

The wealth variable of each agent is thus

$$W(h,t+1,i) = W(h,t,i) + \Phi[R(h,t,m)] - M[W(h,t,m)]$$

Where, $\Phi[R(h,t,m)]$ is the fraction of resource at (m) consumed by (k).

$M[W(h,t,m)]$ is a metabolic rate describing the fraction depleted – commuting cost – to survive the walk from (j) to (m).

State of resource with constant growth:

$$R(k,t+1,m) = R(k,t,m) - \sum_{h \in \Omega m} \Phi[R(h,t,m)] + 1 \text{ (growth increment)}$$

Conservation check point if

$$R(h,t,m) < RR \text{ then } R(h,t+1,m) = R(h,t,m) + 1$$

Economic death if

$$W(h,t+1,m) < 0, \text{ then } A(h,t+1,m) = 0$$

Intervention is possible via subsidies, bankruptcy, death birth or inheritance.

Pareto distribution applied to the concept of wealth where

$$F[W(h,t,m)] = W(h,t,m)^{-\alpha}$$

Is the parameter of the power law (~2).

Object-oriented approaches to social simulation support the perception that local behavior is fundamental to explain global patterns. These are consistent with concepts of complexity of systems emerging in global and structural from simple actions of autonomous agents acting with their own self-interest in mind.

Three interrelated problems of spatial dynamics

(1) Decline of central or core cities

(2) Emergence of edge cities that compete and complement

(3) Rapid suburbanization of cities at periphery

Growth dynamics are modeled from an aggregate basis focused upon geometric form. This generic form does not heed policy issues, taxation, the market or individual behaviors as reactions to these influences (Batty, 2005, 386-7). The growth/sprawl model of urbanization is analogous to the model of an epidemic in which an individual is infected and recovers. "Growth" manifests in four directions (in, up, down, out) mostly out, mostly. This tendency is due to geometrical and resource constraints, preference for newness, preference for lower densities, access to environmental amenities and the desire to avoid traffic congestion.

Networks

Network connections between nodes (i) and j = 1, 2, …, n. If a node has at least one connection, ai = 1, otherwise ai = 0. Aii is undefined and

$$aiaj = aij = 1 = aji$$

and

$$MAX = (1/2)\ n\ (n-1)$$

A direct link between two nodes i,j has distance

$$d(i,j) = aiaj = 1$$

So

$$0 < c < 1$$

aij = 0, for all i,j, c = 0 -> no connectivity

aij = 1, for all i,j, c = 1 -> all elements directly connected.

The mean distance

$$\overline{d} = \sum_{i,j} dij(c)/[(1/2)n(n-1)]$$

and the change in average distance

$$\Delta\overline{d}\ (c: cprevious) = \overline{d}(c) - \overline{d}\ (cprevious),\ c>0$$

The graph (network) is strongly connected at the critical point when each node has direct or indirect links to every other node in the network.

If,

$$aiaj = 0$$

Set

$$d(i,j) = n$$

N = # of elements in the system. (Batty, 2005, 482)

Local average distance for each cluster, the cluster distance

$$D(c,i) = \sum_{h\epsilon\Omega i} \sum_{m\epsilon\Omega i} ahm / [(1 + \sum_j aij)(\sum_j aij)/2]$$

$\Omega(i)$ is the set of nodes linked to i.

Overall clustering for the system

$$D(c) = \sum_i D(c, i)/N$$

N = number of clusters. (Batty, 2005, 488)

Krugmen Model

Centripetal force is the attraction towards location (i) at distance d(i,j). k1 is a size parameter and α is a scaling parameter.

$$V1(t,i) = k1 \sum_j P(t, j) \exp[-\alpha d(i, j)]$$

Centrifugal force, decentralization away from location (i)

$$V2(t,i) = k2 \sum_j P(t, j) \exp[-\beta d(i, j)]$$

k2 = size parameter, β = scale parameter

$$V(t,i) = V1(t,i) - V2(t,i)$$

$$\overline{V}(t) = \sum_j P(t, j) V(t, i)$$

$$P(0,i) = E(i)$$

Where

$$\sum_j P(0, j) = 1$$

$$\nabla P(t,i) = [V(t,i) - \overline{V}(t)]\ P(t,i)$$

$$P(t+1,i) = P(t,i) + \nabla P(t,i)$$

$$P(t+1,i) = P(t,i)[1 + V(t,i) - \overline{V}(t)] = P(t,i)\ \lambda(t+1,i)$$

Where $\lambda(t+1,i)$ is the growth rate (Batty, 2005, 60).

Cantor Set

Let the seed set, $I(\Omega) = [0,1]$.

Recursion: For each remaining closed interval remove the open middle interval of relative length 1/3 and keep the two outside closed intervals.

Define: C, Cantor's Middle Third, as the set which consists of all numbers x in [0,1] which are never removed.

Let us match the remaining intervals on level n with sequences of 0's and 1's of length n.

Let 0 be recorded in the sequence $E_1, \ldots, E_n$, where E = 0 or E = 1, if x in $[a(E_n), b(E_n)]$, where $a(E_n)$ and $b(E_n)$ respectfully are the left and right hand endpoints of the remaining disjoint closed interval immediately to the left of an open middle interval which was removed on level n.

Like wise, let 1 be recorded if x in $[c_n, d_n]$, where c_n and d_n are the end points of a remaining closed interval immediately left of an open middle third.

Claim: (*) Level n consists of 2^n intervals.

Show by induction, level n =1 has $2^1 = 2$ intervals.

Assume (*) holds for n.

Want to show (*) holds for n+1 levels. By assumption level n has 2^n intervals each of which are divided into three equal intervals with one removed. So, level n+1 has $2*2^n$ intervals. Since level n+1 has $2*2n = 2^{2n+1}$ intervals (*) holds for all n.

Claim: (#) Each interval has length $1/3^n$.

Show by induction, level n = 1 has length

$$1/3^1 = 1/3 = (1/3 - 0) = (1 - 2/3)$$

Assume (#) holds for n levels. Want to show (#) holds for n+1 levels.

By assumption, each interval on level n has length $1/3^n$. On the subsequent, n+1, level each interval has 1/3 the length of the previous interval or

$$1/3 * 1/3^n = 1/3^{n+1}$$

So, (#) holds for all n.

If

$$I(E_1, \dots , E_n) = [a(E_1, \dots , E_n), b(E_1, \dots , E_n)],$$

where $a(E_1, \dots , E_n)$ is the left end point of $I(E_1, \dots , E_n)$ a remaining interval, and $b(E_1, \dots , E_n)$ is the right end point, since each interval has length $1/3^n$,

$$b(E_1, \dots , E_n) = a(E_1, \dots , E_n) + 1/3^n.$$

If E = 0 on the n+1 level then

$$I(E_1, \dots , E_n, 0) = [a(E_1, \dots , E_n), a(E_1, \dots , E_n) + 1/3^n]$$

Also if E = 1, on level n+1

$$I(E_1, \dots , E_n, 0) = [a(E_1, \dots , E_n), a(E_1, \dots , E_n) + 1/3^n] =$$

$$[a (E_1, \dots , E_n) + 2/3^{n+1}, b(E_1, \dots , E_n)]$$

Claim: (@) C does not contain any interval.

Proof:

By way of contradiction, suppose $x \neq y$ and [x,y] is an element of C. Then x and y are in the same interval on level 1. So, $y - x \leq 1/3$.

Assume (@) for level n.

On level n+1 where if E = 0,

$$\hat{y} = y$$

$$\hat{x} = \hat{y} + 1/3^{n+1}$$

$$\hat{y} - \hat{x} = \hat{y} - \hat{y} + 1/3^{n+1} \leq 1/3^{n+1}$$

If, E = 1

$$\hat{x} = x$$

$$\hat{y} = \hat{x} - 1/3^{n+1}$$

$$\hat{y} - \hat{x} = \hat{x} - 1/3^{n+1} \leq 1/3^{n+1}$$

Since this is true for every n,

$$y - x \leq 1/3^{n}$$

For all n ->

$$x = y$$

#

Claim: For each n and sequence $E_1,\ldots, E_n$ of 0's and 1's such that

$$I(E_1, \ldots , E_n) = [a(E_1, \ldots , E_n), b(E_1, \ldots , E_n)],$$

And

$$b(E_1, \ldots , E_n) = a(E_1, \ldots , E_n) + 1/3^{n}$$

then, (+)

$$a(E_1, \ldots , E_n) = \sum_1^n 2E_i/3^i$$

Show by induction, for n = 1

Case 1: $E_i = 0$;

$$a(E_1, \ldots , E_n) = 2(0)/3^1 = 0$$

Case 2: $E_i = 1$;

$$a(E_, , \ldots , E_n) = 2(1)/3^1 = 2/3$$

Assume (+) holds for n. Want to show (+) holds for n+1 and sequence $E_1,\ldots, E_{n+1}$.

Case 1: $E_{n+1} = 0$

$$a(E_1,\ldots, E_{n+1}) = \sum_1^n 2E_i/3^i + 2(0)/3^i = \sum_1^n 2E_i/3^i = a(E_1, \ldots , E_n)$$

Case 2: $E_{n+1} = 1$

$$a(E_1,\ldots, E_{n+1}) = \sum_1^n 2E_i/3^i + 2(1)/3^i = \sum_1^{n+1} 2E_i/3^i$$

So, for each na and sequence $E_1, \ldots ,E_n$ the left end point of each remaining interval, $a(E_1, \ldots , E_n)$ can be expressed as

$$\sum_1^{\infty} 2E_i/3^i$$

And as n -> infinity

$$a(E_1, \ldots , E_n) = b(E_1, \ldots , E_n)$$

And C contains no interval; C contains only those numbers, which can be expressed as

$$\sum_1^{\infty} 2E_i/3^i$$

Theorem: x is in C if and only if there is an infinite sequence of 0's and 1's, $E_1, \ldots , E_n$ such that

$$\sum_1^{\infty} 2E_i/3^i$$

Proof

2 - > 1

Suppose

$$\sum_1^{\infty} 2E_i/3^i$$

Want to show x is an element of C.

Let

$$S_n = \sum_1^n 2E_i/3^i \leq x =$$

$$\sum_1^n 2E_i/3^i + \sum_{n+1}^\infty 2E_i/3^i \leq$$

$$\sum_1^n 2E_i/3^i + \sum_{n+1}^\infty 2E_i/3^i$$

$$\sum_1^n 2E_i/3^i \leq x \leq$$

$$\sum_1^n 2E_i/3^i + 2/3^n [\, 1 + 1/3 + 1/9 + \ldots] =$$

$$\sum_1^n 2E_i/3^i + 2/3^{n+1}(1/(1\text{-}1/3)) =$$

$$\sum_1^n 2E_i/3^i + (2/(3*3^n)) =$$

$$\sum_1^n 2E_i/3^i + 1/3^n$$

Therefore, for each n, x is an element of $I(E_1, \ldots, E_n)$.

1-> 2:

Suppose x is in C, want to show there exists an infinite sequence of 0's and 1's, $E_1, E_2, E_3, \ldots$ such that

$$x = \sum_1^\infty 2E_i/3^i$$

x is in C means x is never removed So, x is in one of the intervals on level 1. Let E_1, be such that x in $I(E_1)$. Suppose $E_1, \ldots, E_n$ has been determined such that x is in $I(E_1, \ldots, E_n)$. Since x is never removed x is in $I(E_1, \ldots, E_n, 0)$ or x is in $I(E_1, \ldots, E_n, 1)$. Let E_{n+1} be such that x is in $I(E_1, \ldots, E_{n+1})$.

Claim: $x = \sum_1^\infty 2E_i/3^i$

$S_n = \sum_1^\infty 2E_i/3^i$ is the left end point of $I(E_1, \ldots, E_n)$

So,

$$S_n = \sum_1^n 2E_i/3^i \leq x$$

Let {0,1} ℕ be the space of all infinite sequences of 0's and 1's.

Let

$$f(E_1, \ldots, E_n) = \sum_1^\infty 2E_i/3^i$$

f(∘) is a function from {0,1} ℕ onto C.

Claim: the function f(∘) is one-to-one.

Suppose

$$f(E_1, \dots, E_n) = f(E'_1, \dots, E'_n)$$

and

$$(E_1, \dots, E_n) \neq (E'_1, \dots, E'_n)$$

So, there is a smallest n, such that

$$E_n = E'_n$$

Then

$$f(E_1, \dots, E_n) \in I(E_1, \dots, E_{n-1}, E_n)$$

and

$$f(E_1, \dots, E_n) \in I(E_1, \dots, E_{n-1}, E'_n)$$

The arc length of f(∘) can be approximated as

$$\sum_1^\infty 2^{n-1}/3^n + \lim_{n\to\infty} 2^n \sqrt{(\tfrac{1}{2^n})^2 + (\tfrac{1}{3^n})^2}$$

$$(1/2) \sum_1^\infty 2/3^n + \lim_{n\to\infty} \sqrt{1 + (\tfrac{2}{3})^{2n}}$$

$$(1/2)\ (2/3)/\ (1-(2/3)) + \lim_{n\to\infty} \sqrt{1 + (\tfrac{4}{9})^n}$$

$$(1/2)\ (2) + \sqrt{(1+0)} = 1 + 1 = 2$$

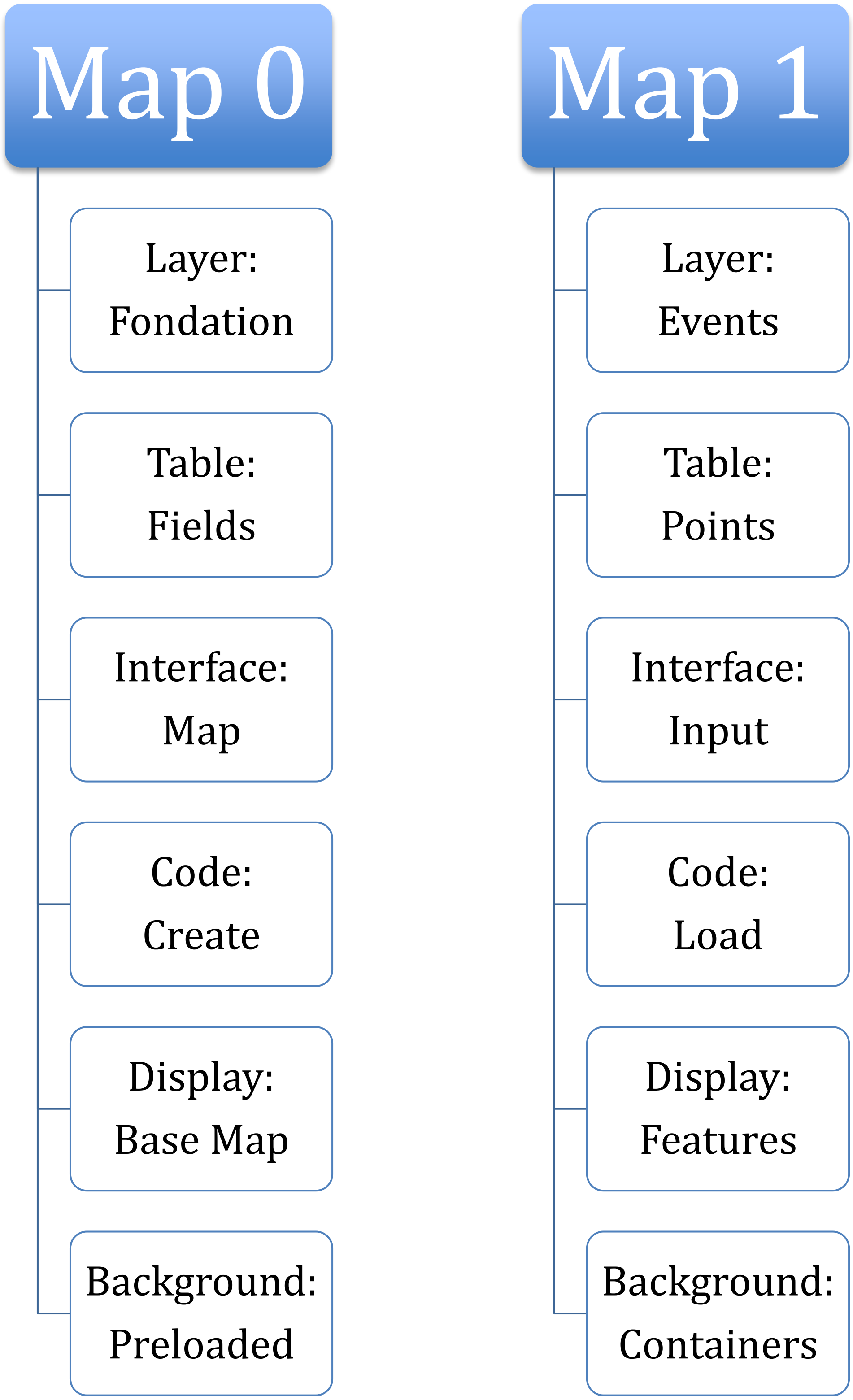

Figure 6. Map(0) and Map(1) Outline

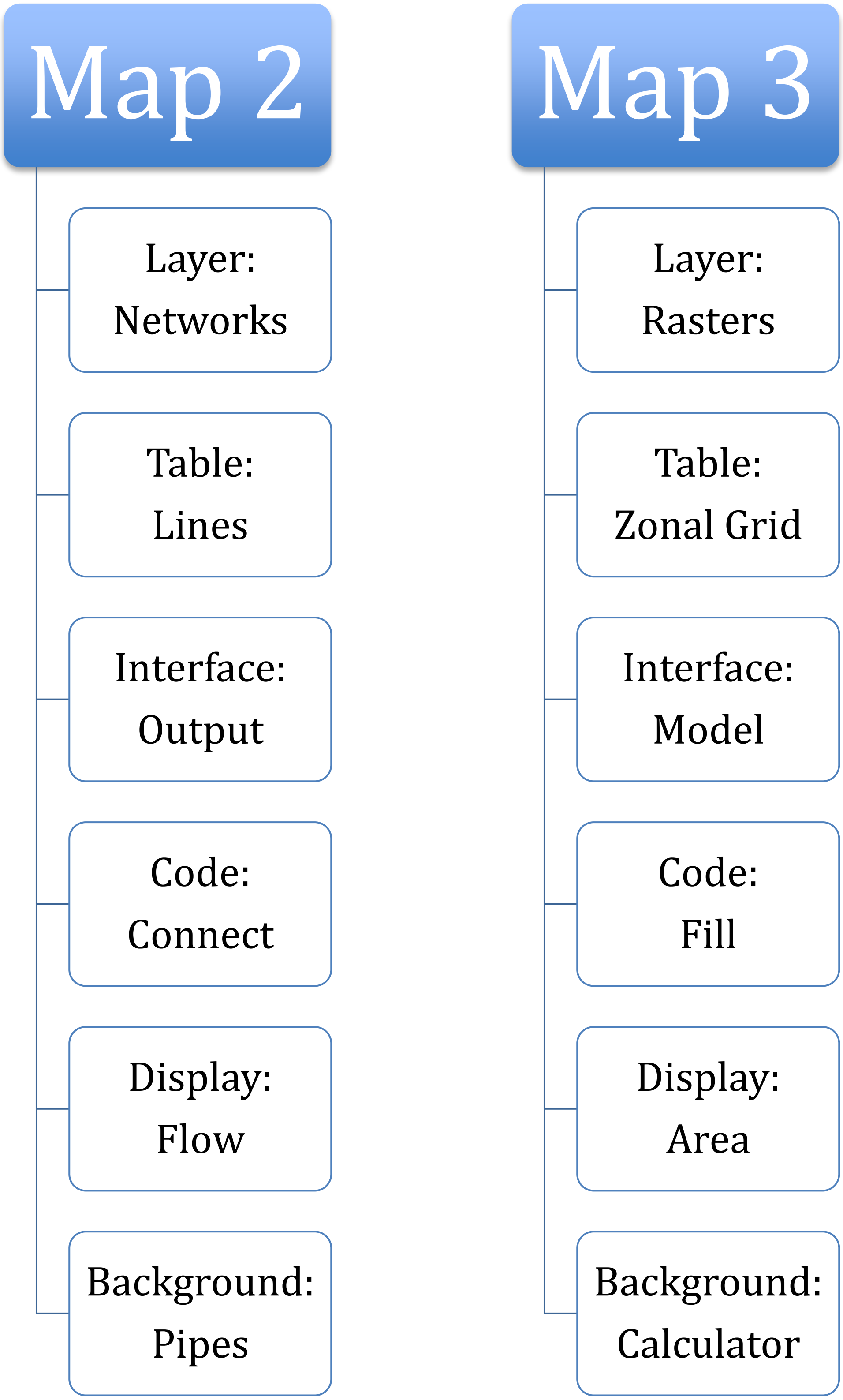

Figure 7. Map(2) and Map(3) Outline

Pseudo Code

MAP(0)

The first interface of the model design provides the background environment upon which user data is mapped and projections are made. Basic elements include political boundaries, transit systems and geographic features (ie. mountain ranges, water bodies etc.).

```
'If NewMap selected, then
'       txtNewMap Active
'       get name from txtbox
'       Activate Load_Click
' If CurrentMap is selected, then
'       Populate CurrentMap combo box
'       Populate CurrentLayer combo box
'       Select first map
'       Select first layer
'       Activate Save_Click
'       Activate Load_Click
'       Activate Remove_Click
' If NewLayer is selected, selected
'       CurrentMap boxes are assumed
'       Read path name are from txt box, maybe input box
'       Activate Load_Click
'BUTTONS:
```

```
'Load_Click
' If load Data is selected, then
'           data is read from Input file (ie, text, xls, shp, lyr, etc.)
'If NewMap selected, then
'           Create new data frame
'           Populate Current Map
'           Create three layers
'           Populate cmbCurrentLayer
'If CurrentMap selected, then
'           Load Map
'           Populate cmbCurrentLayer
'           Refresh active View
'If CurrentLayer selected, then
'           Load selected layer
'If NewLayer selected, then
'           Inputbox
'           Get layer type
'           Create NewLayer
'           Add NewLayer to current Map
'           Load NewLayer
 '          Add NewLayer to the combo box
'Remove_Click()
'If CurrentMap selected, then
```

```
‘           Remove map (data frame)

‘If current layer selected, then

‘           Remove layer

' Delete_Click()

' If CurrentMap selected, then

            Delete Map.mxd

‘If CurrentLayer selected, then

            Delete Layer.shp

' Save_Click()

' If CurrentMap selected, then

‘           Save Map.mxd

‘ If CurrentLayer selected, then

‘           Save Layer.shp
```

MAP(1)

The point-based application creates a GIS map from inputs or reading files (txt, shp, xls, etc.) representing the user's selected parameters. A variety of selection techniques should be available. Users can input data via individual entry, loading from a file or clicking on the map. Layers or shapefiles can be created for the selected data. Fields can be created or have values calculated based upon additional parameters. Functionality will provide for saving or exporting files in selected fashion.

The interface can be used as a spatial scheduler or log. The application provides the capacity to record and process individual surveys. Variables are aligned to the relevant subject. Upcoming events can be input and time coded to display where and when particular groups meet or events occur. Likewise, historical data can be input to display previous trends. Examples include dietary or financial logs, which convey when, and what calories were consumed or where particular purchases occurred. Eventual development will allow this code to be used as a survey devise.

```
' Event, Group, Location and Time boxes serve as basic parameters for user entry.
' They are populated by values corresponding to respective field names.
' Event loads data from combo boxes as an entry into the data file or from source
' data.
' Read events from source
' If LoadFrom is selected, then
'           get name from txtbox or inputbox
```

‘ Activate Create_Click

‘ Read location via tool

‘ Activate Create_Click

‘ Acticate Delete_Click

‘ Entry or Drop Down

‘ Activate cmbE_G_L_T

‘ Activate Create_Click

‘ Activate Delete_Click

‘BUTTONS:

‘ Create_Click

' An entry is created and written to the selected layer.

‘ Create Button loads user input form or data file.

‘ If LoadFrom selected, then

‘ Open Spreadsheet/Shapefile

‘ Read list of events

‘ Create features

‘ If PointNClick selected, then

‘ Start collection tool

‘ PopUp or Inputbox

‘ load files

‘ If Input selected

‘ Read cmboxes

‘ Add feature

‘ Delete_Click

‘If PointNClick selected, then

‘ Start removal tool

‘ PopUp or Inputbox, select feature

‘ Delete feature

‘If Input selected

‘ Read from cmbboxes

‘ Remove feature

' TOOLS:

' PointNClick_tool()

' Connections are created through interaction

' Point and click allows connections to be generated visually via tools.

‘ Tool can create points and connections within network, depending map1 or 2.

‘ Removal_tool()

MAP(2)

The linear feature application enables users to create conduits between selected features. Numerical comparisons and calculations can be processed according to the algorithms and equations of a particular environment. Connections are created based upon the feature selections via attributes, fields, etc. Census data will be input and evaluated according to commuting patterns and economic benefits. A correlation can be approached between individual residential and social behavior norms in the blocks in which they live. Initial efforts shall address functionality of selection and calculation of linear graphics will be added to represent connections.

<MAP> <LAYER> - are loaded from map page inputs

```
' FEATURES:  Coding Objectives
' Field boxes are populated with pre-coded entries or user input.
' Multiple inputs may be made via data file.
' Events or locations are connected by wire based upon user inputs.
‘ If PointNClick selected, then
‘       Read/Remove locations via tool
‘       Activate Add_Click
‘       if PipeCount > 0
‘               Activate Delete_Click
‘               Actovate Create_Click
‘If Points (From, To) selected, then
‘       Activate cmbFrom
‘       Activate cmbTO
```

‘ Load cmboxes with locations

‘ Activate

‘ Add_Click

‘ Delete_Click

‘ Create Click

‘ If Fields () selected, then

‘ Populate cmbFields

‘ Activate Create_Click

‘ Activate Delete_Click

' BUTTONS:

' Add_Click()

' Creates a link between From and To features of specified width and color

' Add link button generates the line instance.

‘ If PointNClick Selected, then

‘ Launch link tool

‘ Drag or click pipes

‘ If Points selected, then

‘ Add pipe from (a to b)

‘ Create linear feature

' Delete_Click()

' Allows user to delete an entry

' Edit button locates and provides entries to be amended.

‘ If PointNClick selected

```
'           Launch link tool
'           Select pipe from map
'           Popup/Inputbox
'           Delete feature
' If Points selected, then
'           Remove from (a to b)
' If Fields selected
'           Delete Network
' Create_Click
'           Build volume structure
'           Connect pipes
'           Activate flow
'           Pre-Raster
'           Save file.
' TOOLS:
' PointNClick_tool()
' Connections are created through interaction
' Point and click allows connections to be generated visually via tools.
' Tool can create points and connections within network, depending map1 or 2.
' Removal_tool()
```

MAP(3)

Raster data sets serve the benefit of being able to process areas, images and projections of fractal based models. Simulations and projections of urban development will be created via Raster based applications. The model design will be implemented to generate Raster projections. Initially, the code will be structured to model fractal equations (Mandlebrot, Cantor etc.) by iterating across a specified area. Eventually, the capacity to work both with Raster Data and fractals shall be used to create fractal renditions of specified areas and processes. Comparisons with historical data will enable the tool to provide projections of system based success and failure. When comparing demographics on social, physiological, economic scales, etc., principal concerns are an area to length ratio when describing the borders of particular "regions."

```
' RASTERS:  Coding Objectives
' User inputs values to originate a fractal.  Examples may be chosen from Mandlebrot
' or other families of equations, etc.
' Values can be read from a data file/raster to generate progression over time.
' Save button stores the raster.
' BoundX inputs the left/right edges of the area to be modeled
' BoundY inputs the upper and lower edges.
' Seed originates the fractal
' Load Raster allows for input from txt or xls
' Save Raster path for saving file
' BUTTONS:
```

' Save_Click()

' Store raster

' Create_Click()

' Creates raster based on inputs or loaded file

Output

' map is exported, saved and opened in chosen format.

REFERENCES

Batty, M., & Longley, P. (1994). *Fractal Cities: A geometry of form and function.* Toronto: Academic Press.

Batty, M., Xie, Y., (1999). Self-organized criticality and urban development. *Discrete Dynamics in Nature and Society*, Vol. 3, 109-124.

Batty, M., (2004). *Hierarchy in cities and city systems*. Centre for Advanced Spatial Analysis, University College London.

Batty, M. (2005). *Cities and complexity: Understanding cities with cellular automata, agent-based models, and fractals*. Michigan: MIT Press.

Batty, M. (2007) *Complexity in city systems: Understanding, evolution, and design*. Centre for Advanced Spatial Analysis, University College London.

Batty, M., (2007). *Model cities.* Centre for Advanced Spatial Analysis, University College London.

Batty, M. (2008). *Cities as complex systems: Scaling, interactions, networks, dynamics and urban morphologies.* Centre for Advanced Spatial Analysis, University College London.

Candia, J., Gonzalez, M., Wang, P.,Schoenharl, T., Madey, G., Barabasi, A-L., (2008). Uncovering individual and collective human dynamics from mobile phone records. *Journal of Physics A: Mathematical and Theoretical*.

Crooks, A., (2008). *Constructing and implementing an agent-based model of residential segregation through vector GIS*. Centre for Advanced Spatial Analysis, University College London.

Dodge, M., Kitchin, R., (2004). *Code, space and everyday life*. Centre for Advanced Spatial Analysis, University College London.

Gonzalez, M., Hidalgo, C., Barabasi, A-L., (2008). Understanding individual human mobility patterns. *Nature*, 453, 779 – 782.

Hudson-Smith, A., Batty, M., Crooks, A., Milton, M. (2008). *Mapping for the masses: Accessing Web 2.0 through crowdsourcing*. Centre for Advanced Spatial Analysis, University College London.

Hudson-Smith, A., Crooks, A. (2008). *The renaissance of geographic information: neogeography, gaming and Second Life*. Centre for Advanced Spatial Analysis, University College London.

Mackett, R., Banister, D., Batty, M., Einon, D., Brown, B., Gong., Y. Kitazawa, K. Marshall. S., Paskins, J. (2007). *Final report on 'Children's Activities, Perceptions and Behaviour in the Local Environment (CAPABLE).* Centre for Advanced Spatial Analysis, University College London.

Mackett, R., Gong, Y., Kitazawa, K., Paskins, J. (2007). *Children's local travel behavior – How the environment influences, controls and facilitates it*. Centre for Advanced Spatial Analysis, University College London.

Voudouris V, (2011), Towards a conceptual synthesis of dynamic and geospatial models: Fusing the agent-based and object-field models. *Environment and Planning B: Planning and Design, 38*(1), 95–114.

Ygnace, J., Drane, C., Yim. Y., De Lacvivier, R. (2000). *Travel time estimation on the San Francisco Bay Area Network using cellular phone as probes.* University of California, Berkeley.