SIMPAT 297 ARTICLE IN PRESS No. of Pages 17, DTD = 4.3.1
12 October 2002 Disk used SPS-N, Chennai

Available at www.MATHEMATICSWEB.ORG POWERED BY SCIENCE DIRECT

ELSEVIER Simulation Modelling Practice and Theory xxx (2002) xxx–xxx

SIMULATION MODELLING PRACTICE AND THEORY

www.elsevier.com/locate/simpat

Computational organizational science and organizational engineering

Kathleen M. Carley *

Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Received 30 May 2002; received in revised form 9 July 2002; accepted 24 July 2002

Abstract

The past decade has witnessed the emergence of a new scientific discipline—computational social and organizational science. Within organization science in particular, and social science more generally, scientists and practitioners are turning to computational analysis to address fundamental socio-technical problems that are so complex and dynamic that they cannot be fully addressed by traditional techniques. Consequently, there is an explosion of computational models, computationally generated findings, interest in doing simulation, and a dearth of support for this enterprise. This paper contains discussions of the underlying fundamental perspective, the relation of models to empirical data and characteristics of necessary infrastructure.

Keywords: Computational modeling; Simulation; Organization science; Organizational design; Organizational learning

Computational modeling is revolutionizing the social and organizational sciences. In part, this is a quite revolution. Traditional research questions are being addressed by computational models, new ideas developed through computation are often presented without the model being presented, models are being used within consulting firms to aid management by providing a new way of systematically thinking about complex processes, computational models are being used to design teams for routine tasks, and much of the work still appears in specialty journals or is spread across journals in the areas of organizations, social science, engineering, computer science

* Tel.: +1-412-268-3225; fax: +1-412-268-6938.
E-mail address: kathleen.carley@cmu.edu (K.M. Carley).

1569-190X/02/$ - see front matter © 2002 Published by Elsevier Science B.V.
PII: S1569-190X(02)00119-3

and management. The ubiquity of computing makes it second nature for many researchers to turn to computational modeling to understand the complexities of social and organizational systems. The ubiquity of computing and the increase in data available on line are leading mangers to ask whether computational techniques can be used to help manage knowledge, design teams, redesign processes and basically engineer the organization. New tools, such as simulation engines, platforms for building multi-agent systems, libraries of statistical functions, and so forth are making it easier to teach simulation and to design, build and analyze computational models. Further, the advance of multi-agent techniques provides social scientists, who are used to thinking about agency, the ability to reason in the terms of their theories. We can now address topics of historic and current concern such as coupling, managing change, strategic choice and evolution. As a result, the field is exploding as new tools, models and results appear constantly (e.g., see [28,32,38]).

Why are so many social and organizational scientists and practitioners turning to computational modeling and analysis as a way of developing theory and addressing policy issues? Essentially, the answer is that these models are, preparatory, cost effective, faster, appropriate, flexible and enable policy development in an ethical fashion. As to this last point, these models provide a virtual world in which policies, extreme conditions and new technologies can be pre-tested ethically. Note, for many issues, to test the basic theory requires examining behavior in extreme conditions (such as behavior in a space station, or under terrorist attack), or under conditions with diverse financial and health implications (such as different scenarios for downsizing). In such cases, it is not ethical for the scientists to set up the relevant natural or laboratory experiment. Using simulation models the researcher or policy maker can examine hypothetical conditions that do not exist, such as new legislation changing benefits packages or new types of immunizations. In this way computational analysis helps plan and prepare for real-world problems without human-in-the-loop testing. Creating new technologies, procedures and legislation for data collection in the field or lab is expensive. Running human laboratory experiments is time consuming and it is very expensive to run groups of more than five individuals. Using a computational model, the researcher can run a larger number of virtual experiments, more cases per cell, larger virtual experiments, than it is possible to collect in the field or lab in the same amount of time. Computational analysis is thus highly cost effective. Computational models can run faster than real time. As a consequence, errors in design can be detected sooner and process checks can occur more frequently thus leading to more rapid theory building. Social and organizational systems are complex non-linear dynamic system; hence, computational analysis is an appropriate technology as models can have these same features. Computational models are flexible. Response to novel situations requires rapid evaluation of previously unexamined alternatives. When there is a computational model of a system, such as an organization, if a new technology comes in to being it can be relatively rapidly added to the model and the implications examined.

Just as there are many reasons to turn to computational theorizing there are many ways in which computational models can be used. Computational models can be used as test beds for generating new ideas. They are “what–if” scenario generators

SIMPAT 297 ARTICLE IN PRESS No. of Pages 17, DTD = 4.3.1
12 October 2002 Disk used SPS-N, Chennai

that can be used to predict the impact of new technologies or policies. Many researchers used computational models to develop theories; for placing the idea in the model requires formalization and addressing many previously neglected issues. Computational models can be used to determine whether a mechanism that is posited in the literature to cause some outcome is necessary to cause it. This can be done, by examining whether other mechanisms can generate the same results. Models serve as decision aids helping researchers, policy makers and managers think through problems in a more systematic fashion. A traditional, and still important, use of computational models is to forecast the future. Computational models are beginning to be used as training tools; admittedly most of the early applications were in training to use a physical system, such as flight simulators. But social simulator tools are in the offing. Researchers also use the results from virtual experiments run through a computational model to generate hypotheses and suggest critical experiments, items for surveys and factors to be varied in experiments, and so on. Results from these virtual experiments are also used to suggest the relative impact of different variables (factors) on the outcome; e.g., what is the relative impact of education and age on performance. Computational models are being used to suggest limits to statistical tests for non-linear systems, often employing enumeration and Monte-Carlo routines to determine the shape of the underlying distributions. With verified systems, computational models can substitute for a person, a group, a tool, etc. in a lab experiment. This makes it possible to see how the human subject would behave in larger groups that are typically possible to bring together in a lab.

There are of course factors limiting the advancement of this field. Many of these stem from the difficulties in educating new members, sharing data, sharing programs and sharing tool kits. The wide range of expertise, the general lack of knowledge of who else is working in the field, the wide variety of computer platforms and languages, the relatively short length of time over which most people have worked in this field all contribute to less data, program and results sharing than is needed. Given the dearth of funding and the nature of academic promotion, most academics who develop simulation tools spend little time on developing good graphics, a flexible user-interface, or coding for portability. Few members of the computational social and organizational field are linked to industry in an effort to commercialize their tools—and there is relatively low commercial value for many of the simulations. A premium is generally places on developing new simulations and addressing new issues; rather, than on validation, model sharing, usability and portability.

1. The computational organization theory perspective

This new body of research in computational social and organizational systems has led to a number of major advances. The coupling of computational theorizing with the advances in social networks, cognitive sciences, computer science and organization theory has led to a new perspective on organizations that takes into account the computational nature of organizations and the underlying social and knowledge networks [12]. This basic perspective is now described. The fundamental tenets are sum-

Table 1
Basic tenets of computational social and organizational science

Tenet	Illustrative argument
Computational systems are organizational—multi-agent, require coordination and embedded social knowledge	[12]
Organizational design involves creating a match between tasks, resources, knowledge and agents, aka congruency, requisite variety	[6]
Search may either explore new parts of the knowledge and interaction space or exploit old competencies	[26,30]
Societies, organizations and individuals are complex, computational and adaptive	[12]
Agents are information processing systems	[12,36]
Synthetic adaptation: agents composed of intelligent, adaptive agents are intelligent and adaptive	[10]
Societies and organizations contain networks of multiple agents	[10]
Agents may be human, artificial or synthetic	[12]
An agent's actions are constrained and enabled by the networks in which they are embedded	[17]
Agents are structurally and boundedly rational	[36]
Agents have multiple categories of knowledge—task, social and transactive memory	[39,40]
Agents are heterogeneous	[16]
Knowledge as structural (ideas are connected)	[23]
Cognition, knowledge and memory is distributed	[20,21]
Transactive memory includes knowledge of who knows who and who knows what	[39,40]
Information technology as agent	[8,23]
All adaptation is constrained by networks and procedures	[10,24]
Organizational decision making is a product of learning through search	[30]
Search may be either local or global	[30]
Multiple levels of learning, adaptation, evolution	[10]
Multiple types of learning	[10]
Culture and social/organizational structure can co-evolve	[8]
Organizations exhibit liability of newness	[37]
Permeable boundaries around agents, resources and tasks	[10,22]
Agents, knowledge, resources and tasks form a meta-network	[25]
The meta-network is dynamic	[11,13]
Roles are derived as changes occur in the meta-network	[10]
Behavior emerges	[16,29]
The behavior of a synthetic agent is not a simple aggregation of the behavior of the members	[10]
History matters as does starting conditions	[33]
Societies and organizations as complex systems	[1]
Culture is a knowledge level phenomena	[18,19]

marized in Table 1. The computational social and organization science perspective is drawn from a large number of empirical studies. This perspective draws on the arguments of distributed cognition [20,21], transactive memory [31,39,40] and the social construction of knowledge [7,10]. Viewing organizations from this perspective makes it clear that multi-agent models can be meaningfully employed in the development and explication of organization theory.

At the heart of this perspective is the argument that organizations are complex, computational and adaptive synthetic information processing agents. This new perspective urges a formalization of the roles of networks, learning and agency in affecting social and organizational change. Organizations are composed of intelligent adaptive agents who are constrained and enabled by their positions in networks linking agents, knowledge, resources and tasks [25]. These networks are dynamic with learning and simple demographic processes such as mobility, birth and death, and organizational change processes such as hiring and downsizing being critical determinants of change in these networks. The principle of synthetic adaptation asserts that a composite agents (agents composed of other agents) composed of intelligent adaptive agents is also intelligent and adaptive [10]. A feature of synthetic adaptation is that a composite agent is not simply the aggregate of the component agents; behavior is not just the average or union of the behavior of the members. Rather, the composite agent is a network whose behavior is a function of complex processes for combining and generating collective outcomes.

The computational social and organizational science perspective extends the social network perspective to multiple networks by using the meta-matrix perspective (see Table 2). Social networks are typically seen as affecting a wide range of behaviors ranging from power to consensus to adaptability. According to the common formulation such networks are in terms of ties among personnel. However, networks are more ubiquitous than this and entities besides agents can be networked together [25]. Agents, knowledge, resources and tasks are linked in to a meta-network [11]. This meta-network representation effectively combines the knowledge level perspective, the social network perspective and the project management perspective. Models from many perspectives are moving to this representation. Once the organization or market is represented this way the entire host of social network measures become valuable, it is easier to link models to data, it is easier to contrast or dock models, and in principle, it is easier to link models to each other. Another advantage of this conceptualization is that complexity can be measured in terms of components and relationships—comparable to the way it has been measured in many fields.

In the computational social and organizational sciences, computational systems are often viewed as organizational. A typical computational system that is in effect an organization is a multi-agent system requiring coordination of the agents. Knowledge and "social" knowledge is typically embedded in the computational agents, and the incorporation of social knowledge enhance the ability of these agents to operate effectively. Computer science, at least a portion of it, is in this sense becoming a social computer science.

The computational social and organizational science perspective is distinct from many traditional social perspectives, particularly in the way "stability" is treated. In contrast to traditional economics, agents are seen as satisficers not optimizers. The collective behavior of agents, either real or artificial, either individual or composite, results in emergent phenomena. Consequently, organizations and political institutions take on a particular architecture or form because it happens to emerge from the complex dynamics and interactions of the agents in the organization and the networks in which they are embedded and not because they are trying to opti-

Table 2
Meta-matrix linking agents, knowledge, resources, tasks and organizations into a network

	Agents	Knowledge	Resources	Tasks	Organizations
Agents					
Tie	Interaction network	Knowledge network	Capabilities network	Assignment network	Employment network
Phenomenon	*Who knows who*	*Who knows what*	*Who has what*	*Who is assigned to what*	*Who works where*
Pattern	Structure	Culture	Capabilities	Jobs	Demography
Dynamic	Structuration Redesign	Learning (Socialization, acculturation, retraining)	Resource acquisition	Retasking, promotion	Hiring, firing
Knowledge					
Tie		Information network	Necessary expertise network	Knowledge requirements network	Knowledge competency network
Phenomenon		*What informs what*	*What must be known to use this resource*	*What knowledge is needed to do what*	*What knowledge is where*
Pattern		Knowledge	Training requirements	Knowledge needs	Inter-organizational culture
Dynamic		Innovation	Training	Project training	Selection, Organizational learning
Resources					
Tie			Substitution network	Resource requirements network	Resource competency network
Phenomenon			*What resources can be substituted for what*	*What resources are needed to do what*	*What resources are where*
Pattern			Training requirements	Resource needs	Corporate competencies
Dynamic			Training	Resource assignment, acquisition	Strategic choice, purchases
Tasks					
Tie				Precedence network	Industrial network
Phenomenon				*What needs to be done before what*	*What tasks are done where*
Pattern				Operations, task flow	Niche

Table 2 (*continued*)

	Agents	Knowledge	Resources	Tasks	Organizations
Dynamic				Re-engineering, Process planning	Strategic repositioning
Organizations					
Tie					Inter-organizational network
Phenomenon					*Which organizations work with which*
Pattern					Market or field structure
Dynamic					Formation of alliance, joint ventures, tiger teams, agreements

This extends the representation in Table 1 in [11,25].

mize relative to some decision criteria. In contrast to much of sociology or psychology where roles as viewed as given, here roles are derived as the way in which agents, tasks, resources and knowledge are connected change. This provides a very network definition of a role as a common pattern of relations in the meta-network shared by various agents. Another break with tradition is that computational social and organizational theorists treat agents as heterogeneous. The power and behavior of the models appears to derive precisely from the fact that not all agents are alike.

The computational social and organization science perspective is essentially a neo-information processing perspective. There are two key implications of this view—knowledge as network and information technology as agent. According to the knowledge as network perspective, knowledge resides in the minds of the participant agents and in the connections among them. Learning, in all its guises (such as mimicry, learning by being told, learning by doing, discovery) leads to changes in what is known and in the connections among agents [21]. To understand why information technologies can be thought of as agents, it is important to recognize that from this perspective any entity that has some information processing capability is viewed as an agent. These capabilities include the ability to initiate an interaction, to send, receive, learn, store, forget, discover, or manipulate information (which can include information about who knows what or does what or beliefs). Agents vary in which capabilities they possess and in the strength of that capability. For example, most databases cannot initiate interaction but can "send" information (when they are read) and can "learn" as information is added to the database. The capabilities of the agents in the organization or society will define what types of "social" behaviors emerge [15].

In many of the organizational models there is a focus on not just information but the type of information. Information varies in being either social, task related or transactive (knowledge of who knows who, who knows what, who does what). Relationships among agents enable and constrain access to information. Tasks impact what type of knowledge is currently salient and whether the information that agents have impacts their performance on the task. Differentiation of information appears to be key to improving our understanding of core processes such as the evolution of norms, negotiation and organizational learning.

A key feature of the computation social and organization science perspective is that agents learn and systems adapt and evolve. While traditional learning studies pointed to changes in market share, productivity or performance as indicators of learning [3]; more recent work looks deeper at the link between individuals, knowledge and organizational outcomes at both the micro and macro level [4,5,39]. As noted, learning and memory can exist at the individual, group, organizational and societal level. When learning at one level clashes with learning at another level, the organization may fail to adapt [10]. The dominant style of learning employed by an organization can be seen as a form of corporate strategy. For example, one type of learning is "search" or exploration. Another type of learning, based on reinforcement theory, is experiential or exploitation. Tradeoffs between exploration and exploitation are seen as fundamental to organizational strategic performance.

Unlike organization science where groups and organizations are often viewed as discrete entities, according to the computational perspective organizations and groups have permeable boundaries. The boundaries change as the set of agents, resources, knowledge and task that make up that group or organization change. Further, agents and tasks are not inevitably separate entities [22]. For example, imagine the task of painting a line on a car. For this task, for a human doing this task there are a large number of subtasks including picking up the paintbrush, getting paint on the brush, moving from one end of the car to another and checking that the line is straight. Now imagine a robot that happens to have a built in airbrush and paint with cars rolling in front on an assembly line. In this case the subtasks include starting the flow of paint, stopping the flow of paint. Organizational design thus becomes a process of trying to create congruence between the task-resource needs given the task precedence and the agent-resource capabilities and agent-task assignments. Such congruence can be found via optimization routines when the tasks are routine. This view of design may seem obvious to shop floor managers and industrial engineers. However, for organizational theorists and practitioners who assume that the only agents are humans, this approach to organizational design is rather novel.

Learning within organizations is ultimately tied to culture [18]. Operationalizations of culture are attempts to tangibly represent culture defined as the "... pattern of basic assumptions that the group learned as it solved its problems of external adaptation and internal integration that has worked well enough to be considered valid and therefore, to be taught to new members as the correct way to perceive, think and feel in relation to those problems'' [34] p. 13. Culture becomes the way in which the group responds to change in external and internal environments as well as a framework that guides the way individuals relate to each other. Since relationships provide the context for communicating basic assumptions they play an integral role in culture creation and maintenance [19].

2. Linking to data

As the field of computational social and organizational science matures researchers are linking their models to data. Within the field of simulation and modeling there are a number of ways in which such links are made. The set of techniques, as is evidenced in this issue, range from face validation to full-scale prediction. In general, it is important to recognize that the evaluation of a computational model's ability to predict and explain social and organizational behavior involves a number of steps looking at the internal integrity [1] of the model, running a series of virtual experiments, examining the model's sensitivity to various conditions and validation against external data. Data enters in to this process in a number of ways. Before discussing some of these ways it is worth noting that an advantage of the meta-matrix representation is that it provides a common conceptual device that can be used across models and real data; helping create a standard for thinking about the set of entities in these models and the relations among them. Putting the appropriate real data and simulated data in meta-matrix form increases the feasibility, likelihood and strength of validation possible for large classes of social and organizational science computational models. A final caveat is that the type, degree and need for validation depends on the purpose of the model, whether the uncertainties in the model have a critical impact on the model results, and whether the model itself will be used to set policy or make decisions where errors may be critical.

Virtual experiments: Many computational models in the social and organizational sciences are sufficiently complex that the entire response surface cannot be analyzed in a single paper. As a result, it is often critical to examine the models predictions relative to a specific set of questions. To that end, a virtual experiment is run. A virtual experiment is an experiment in which each cell in the experimental design is populated with data generated by running the computer simulation model or models whose behavior is being examined. Through this virtual experiment the computational model is used to generate hypotheses often by exploring a set of realistic or hypothetical conditions [9]. The statistical or graphical analysis of results from a virtual experiment conducted using the computational model leads to a series of hypotheses. Standard principles of experimental design should be followed in designing a virtual experiment; however, as is illustrated by several papers in this volume the amount of data that can be generated from a computational model is ex-

[1] Examining the internal integrity of a model involves simply making sure that the model does not contain programming errors.

SIMPAT 297
12 October 2002 Disk used
ARTICLE IN PRESS
No. of Pages 17, DTD = 4.3.1
SPS-N, Chennai

tremely high compared to that from a human experiment. Consequently, many statistical packages may even be overwhelmed by the amount of data that the computational experimenter generates. Nevertheless, the results from the virtual experiment should be analyzed statistically or graphically. The results of that analysis are the hypotheses that can be examined using data from human laboratory experiments, live simulations, games, field studies, or archival sources.

Sensitivity analysis: In order to test the strength of a computational model a series of sensitivity analyses are done on critical components and assumptions in the model. In a sensitivity analysis a variable that is critical to the model, such as the size of the group or population, is varied over a wide range, results are gathered and then statistically analyzed to see what impact change in size has on the conclusions drawn when using the computational model. In general, in computational models of complex systems, like those in this issue, there are two main determinants of model sensitivity. First, there may be strong non-linear interactions between key variables in the model. As a result, conclusions can be quite sensitive to the specific choices made. Second, there may be a high level of uncertainty resulting from use of random number generators to guesses about the underlying distributions or processes. Conclusions drawn from the model can be quite sensitive to the choices made to handle these uncertainties. For at least these two reasons, virtual experiments are needed to test the sensitivity of conclusions to initial conditions and the inherent uncertainties.

Face validation: The most pervasive approach to the evaluation of simulation tools is face validation. Here the authors argue that the critical characteristics of the process being explored are adequately modeled and that the model is able to generate results that capture known canonical behavior. In some fields the argument is made that there are a set of stylized facts about the real-world and the model is able to generate behavior that matches, at least in form, these facts. An example fact is the $1/F$ distribution of city sizes within the United States. Face validation should be done for all computational models and it is often the only type of validation relative to external data that is warranted for illustrative computational models like those in this volume.

Tuning: Tuning is the process of using a sample of existing data to set parameters in the simulation model so that the model's results match those observed in a sample case. Where possible, only a sample of the real data collected is used in this way; thus, reserving some for later steps in the validation process. Tuning is a critical procedure for demonstrating the capability of the model to match real-world data. Tuning is a critical step in the validation process. Tuning a model reduces the sources of uncertainty in a model by substituting known conditions and data for guesses and random choices from generic distributions. A special case of tuning is setting the initial conditions in a model to match real-world data. An example would be using census data on a particular city to set the base case for modeling a city of agents. Currently a huge amount of data that can be used to tune computational models exists on the web. However, most web interfaces are designed for human readability not program readability; hence it is difficult to link models to on-line data. The result is that researchers either ignore this data or code it manually in to a form their simulation can use.

General trend: A typical approach to validation is to compare the results of a tuned simulation with those of real data. Visualization and statistical tests are done to look for degree of fit between the real and simulated data. Depending on the focus, the emphasis may be on demonstrating that the pattern of data is the same, the statistical properties of the curves are the same (e.g., same moments), or point estimation (the exact values are the same). If the model is not deterministic then validation often involves comparing a specific set of data with a distribution of simulated data. Comparison against a single run of the simulation is inappropriate as the uncertainties produced by the underlying random number generator dominate behavior. Consequently, the research runs a series of runs, locates the average or distribution across those runs, and then compares that ensemble with the real data.

Docking: [2] An important goal in science is accumulation. One way of getting accumulation in the field of computational social and organizational science is to build models out of other models. Another is to determine whether models from different traditions, possibly with different underlying assumptions, nevertheless are capable of generating the same result. Docking can lend insight in to the fundamental nature of the underlying processes, the conditions needed to generate key results and factors that are irrelevant [2]. One way in which docking proceeds is to determine with which aspects of the meta-matrix the two models are concerned.

Social turing test: An important technique for the evaluation of simulated data is whether another algorithm behaves comparably on the simulated and real data [15]. If the algorithm does, then the computational model is generating data that is, at some level, indistinguishable from "real" data. As an example, imagine that the modeler has built a model of use of the Internet to search for health information and is using it to predict whether the behavior is different after news broadcasts that a new type of disease or medication exists. Now, if a detection algorithm is able to locate an upsurge in usage of the Internet and in particular access to web sites on the real data, and can detect the same pattern of behavior in the simulated data, then this test has been met. This test does not guarantee that the underlying processes in the simulation and in reality are the same; but it does suggest that they are functionally similar.

Policy/educational value: Simulation models are generally considered valuable if they provide insights into a problem. A measure of the value of the simulation is if real-world analysts, policy makers, managers or educators are interested in the simulation or the results from virtual experiments generated by using the simulation. With most simulation models of complex social systems it takes 5+ years (often over 10 people-years) to get to the point in the model's development where other analysts

[2] An important side note here is that the term docking as originally coined [2] refers to the process of determining whether two models generate the same or similar results and under what conditions. This is distinct from the way the term is used in transportation where docking refers to getting two items lined up so that things can move between them. Some researchers are beginning to use the term docking with computer models to mean aligning the models so that the output of one is the input to the other.

can use the model. Three factors are important in getting a model to this point: (1) portability (the model can run on diverse machines and under diverse operating systems), (2) validation (particularly in the policy realm, some level of validation has been achieved) and (3) transparent interface (usually this means good graphics, easy to use interface, documentation and sample cases of use).

3. Organizational engineering

The area of computational social and organization science has come a long way in the past 50 years. Theoretical and methodological results are leading to a new paradigm, a new field of intellectual endeavor that builds out of but extends beyond network analysis, computer science and the non-computational social sciences. The subject matter is the formal analysis and design of societies and organizational systems. These systems are complex and dynamic. Computational modeling and analysis is a cost effective, intellectually appropriate technique for understanding these systems. The need for change management, the ease of using even the simplest of computational models to do what if analysis and the increasing access to large-scale databases engenders and an intellectual environment where both theoretical and applied issues are strongly connected. Consequently, this new field has both a scientific and an engineering aspect.

Looking at the field now, one is tempted to conclude that there is a major gulf between theoretical or intellective models and the more applied emulation models. The intellective models often utilize extremely simple agents in highly stylized settings whereas the more emulative models increase either agent and/or context specificity. On the one hand, this gulf is real. Game theoretic models filled with rational agents deciding whether to cooperate or not have little to do with a detailed model of the interaction between humans and robots that take place to coordinate the robot in collecting data samples in extreme environments. On the other hand, there is no firm line of demarcation between emulative and intellective modeling. VDT [3] [27] which is a detailed model of organizational design for routine tasks has never the less been used at the intellective level to replicate basic portions of the garbage can model and many of the tenets of contingency theory. In contrast, ORGAHEAD [4] [14] which has more of the features of an intellective model has never the less been used to suggest adaptation policies within corporations and non-governmental units. Thus, what might first appear as a gulf is in reality a gradual slide from one type of model to another.

[3] The newest version is SimVision. Persons interested in a SW license for SimVision should contact Vicki Trafton vtrafton@vite.com at Vite Corporation www.vite.com. Both a commercial and academic version are available. Those interested in doing research with SimVision should first contact Ray Levitt ray.levitt@stanford.edu to discuss their research objectives, before talking to Vite about a license.
[4] ORGAHEAD is available from contacting Kathleen Carley kathleen.carley@cmu.edu. There is a PC and Unix version.

It is likely that in the future many models will slip back and forth along this continuum as they are used in different contexts. There are several reasons to expect this. First as more models are validated their use as policy and management tools increases. Modern programming techniques such as separating the IO of the program from the internal processes, utilization of agent based or object-oriented architectures and self-documenting code are increasing the utility and portability of new models. The increased sophistication of simulation scientists and the increased need to reduce the uncertainty in policy decisions is leading more computational social and organizational scientists to turn to virtual experiments to generate predictions from their models. The results from multiple virtual experiments are being compared statistically to answer "what–if" questions; questions about what would happen if the changes identified in the virtual experiments occurred. Being able to run more "what–if" analyses increases the policy value of these tools, their use as decision aids and their use as teaching devices. Further, computational advances and the increased placement of data on the web are increasing the ease of linking models to actual data feeds. For example, models that do organizational design will in the near future be capable of being linked to HR databases. Consequently, the utility of these models for social and organizational engineering will increase.

4. The path to the future

Computational social and organizational models are becoming more ubiquitous, in organizations and in policy analysis. We should expect to see an increasing number of models that become large-scale emulation models. This will increase the need to validate a wider range of models at a more levels. As the need for model in industry and government increases we should see a corresponding movement build models by building on or integrating previous models, to build and share model components, to create standard model and data interchange formats and increasingly standard representation schemes. We should also expect to see an increasing number of models that are linked to real time data feeds. For the field to flourish there should also be an increasing number of educational tools such as modeling platforms with a number of intellective models pre-built in that platform and available to students to manipulate and run virtual experiments. A key element here is increasing both the validity and generalizability of the models so that they can be used in situations not necessarily foreseen by the original developers.

What these predictions imply, is that a significant portion of this burgeoning field will be engaged in large-scale system building and/or the development of educational platforms. Consequently, social scientists are going to be faced with a change in the way they conduct science or the movement of social science problems in to computer and information science. One change is likely to be a movement from single person research to teams of researchers. It is not feasible for one scientist, or even a student–professor pair, to build, analyze, validate, increase portability, distribute and maintain a large-scale system let alone a set of systems. Rather teams of researchers with a range of skills are needed; with such skills including social networks, statistics, experimental design, organization theory, social science, computer modeling, optimization, operations research, machine learning, etc. In many universities, engineering and computer science departments often have a better infrastructure for supporting such system teams. Cohorts of faculty that work together in centers, a high ratio of graduate students per faculty member, the presence of research staff, high levels of funding, industrial collaborators, student internships, are all key to a supportive system building infrastructure for the computational social and organizational sciences. For most universities, this represents a major shift in the underlying culture of social science departments and business schools. The need for teaming and for infrastructures support for system level research is critical if the computational social and organizational sciences are to reach there full potential.

Fundamental changes in the way in which social science research is funded and the scale of funding are also needed. Creating portable, user-friendly, validated models in the social and organizational sciences will require the same level of sustained, team and center level funding more commonly available in the computational and physical sciences than in the social sciences. Labs, co-laboratories need to be established and maintained.

There are additional factors that are key to achieving the promise of this field. These factors include the development of open source program libraries, educational computational tools and data sharing. For major progress to occur researchers will need to stop reinventing the wheel and build on each other's code. This means that in writing a simulation in the first place the researcher should try to attend, as is feasible, to issues of documentation, portability and usability. Further, to reduce the cost of validating and testing model features shared datasets and system standards are needed. The field will achieve major gains by creating canonical datasets that are then shared across groups. Ideally, the field should develop a web-based international infrastructure for the computational social and organization science community. Key components of this infrastructure would be

1. intelligent agents for locating, describing and brokering between people, datasets, computational models, results and artifacts (research reports, articles, books);
2. socially intelligent agents for facilitating data collection, analysis, visualization and model building and validation;
3. educational outreach programs and educational tools including new textbooks and simulation platforms for the classroom;
4. formal interchange languages and programs so that data and model results can more easily move from one simulation or data analysis package to another;
5. a suite of models written to be as portable as feasible across multiple operating systems and platforms;
6. the development of standard web interfaces formats for public data that enable computer programs to download the data automatically without manual manipulation of the underlying html file.

Such an infrastructure should support computationally based research in social and organization science, with the aim of altering social science education and research on societies and organizations. A second aim should be to increase efficiency by decreasing the time-to-critical-insight; decreasing the time-to-application; improving the quality of research results; addressing qualitatively new problems and new scales of problems and decreasing the startup costs and time for new researchers.

The field of computational social and organizational science is growing rapidly. Changes such as those suggested here offer a new way of doing social science research; however, they require new tools and a new infrastructure. As we move toward the future, if we are to see the promise hinted at by the papers in this volume, these changes to the infrastructure and the associated cultural changes are needed.

5. Uncited reference

[35]

Acknowledgements

This work was supported in part by the Office of Naval Research (ONR), United States Navy Grant no. N00014-97-1-0037, the Army Research Labs, NSF IRI9633 662, NSF KDI IIS-9980109, NSF IGERT program in CASOS and the Pennsylvania Infrastructure Technology Alliance, a partnership of Carnegie Mellon, Lehigh University and the Commonwealth of Pennsylvania's Department of Economic and Community Development. Additional support was provided by ICES (the Institute for Complex Engineered Systems) and CASOS—the center for Computational Analysis of Social and Organizational Systems at Carnegie Mellon University (http://www.casos.ece.cmu.edu/casos). The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the Office of Naval Research, the National Science Foundation or the US government.

References

[1] R.M. Axelrod, M.D. Cohen, Harnessing Complexity: Organizational Implications of a Scientific Frontier, Free Press, New York, 1999.
[2] R. Axtell, R. Axelrod, J.M. Epstein, M.D. Cohen, Aligning simulation models: a case study and results, Computational and Mathematical Organization Theory 1 (2) (1996) 123–142.
[3] L. Argote, S. Beckman, D. Epple, The persistence and transfer of learning in industrial settings, Management Science 36 (1990) 140–154.
[4] L. Argote, Organizational learning curves: persistence, transfer and turnover, International Journal of Technology Management 11 (7&8) (1996) 759–769.
[5] L. Argote, Organizational Learning: Creating, Retaining and Transferring Knowledge, Kluwer, Boston, MA, 1999.

[6] R.M. Burton, B. Obel, Strategic Organizational Design: Developing Theory for Application, Kluwer, Norwell, MA, 1998.
[7] K.M. Carley, A theory of group stability, American Sociological Review 56 (3) (1991) 331–354.
[8] K.M. Carley, Communication technologies and their effect on cultural homogeneity, consensus, and the diffusion of new ideas, Sociological Perspectives 38 (4) (1995) 547–571.
[9] K.M. Carley, On generating hypotheses using computer simulations, Systems Engineering 2 (2) (1999) 69–77.
[10] K.M. Carley, On the evolution of social and organizational networks, in: S.B. Andrews, D. Knoke (Eds.), Special Issue of Research in the Sociology of Organizations on Networks in and around Organizations, vol. 16, JAI Press, Stamford, CT, 1999, pp. 3–30.
[11] K.M. Carley, Smart agents and organizations of the future, in: L. Lievrouw, S. Livingstone (Eds.), The Handbook of New Media, Sage, Thousand Oaks, 2002, pp. 206–220 (Ch. 12).
[12] K.M. Carley, L. Gasser, Computational organization theory, in: G. Weiss (Ed.), Distributed Artificial Intelligence, MIT Press, Cambridge, MA, 1999, pp. 206–220 (Ch. 7).
[13] K.M. Carley, V. Hill, Structural change and learning within organizations, in: A. Lomi, E.R. Larsen (Eds.), Dynamics of Organizations: Computational Modeling and Organizational Theories, MIT Press/AAAI Press/Live Oak, Cambridge, MA, 2001, pp. 63–92 (Ch. 2).
[14] K.M. Carley, D.M. Svoboda, Modeling organizational adaptation as a simulated annealing process, Sociological Methods and Research 25 (1) (1996) 138–168.
[15] K.M. Carley, A. Newell, The nature of the social agent, Journal of Mathematical Sociology 19 (4) (1994) 221–262.
[16] J. Epstein, R. Axtell, Growing Artificial Societies, MIT Press, Boston, MA, 1997.
[17] M. Granovetter, Economic action and social structure: the problem of embeddedness, American Journal of Sociology 91 (1985) 481–510.
[18] J.R. Harrison, G.R. Carrol, Keeping the faith: a model of cultural transmission in formal organizations, Administrative Science Quarterly 36 (1991) 552–582.
[19] V. Hill, An examination of culture creation and maintenance in temporary placement agencies, Ph.D. Thesis, GSIA, Carnegie Mellon University, 1998.
[20] E. Hutchins, The social organization of distributed cognition, in: L.B. Resnick, J.M. Levine, S.D. Teasley (Eds.), Perspectives on Socially Shared Cognition, American Psychological Association, Washington, DC, 1991.
[21] E. Hutchins, Cognition in the Wild, MIT Press, Cambridge, MA, 1995.
[22] D.J. Kaplan, The STAR system: a unified multi-agent simulation model of structure, task, agent, and resource, CASOS working paper, Carnegie Mellon University, 1999.
[23] D.S. Kaufer, K.M. Carley, Communication at a Distance: The Effect of Print on Socio-Cultural Organization and Change, Lawrence Erlbaum Associates, Hillsdale, NJ, 1993.
[24] S.A. Kauffman, At Home in the Universe: The Search for Laws of Self-Organization and Complexity, Oxford University Press, New York, NY, 1995.
[25] D. Krackhardt, K. Carley, A PCANS model of structure in organizations, in: Proceedings of the International Symposium on Command and Control Research and Technology, Evidence Based Research, Vienna, VA, 1997.
[26] D.A. Levinthal, J.G. March, A model of adaptive organizational search, Journal of Economic Behavior and Organization 2 (1981) 307–333.
[27] R.E. Levitt, G.P. Cohen, J.C. Kunz, C.I. Nass, T. Christiansen, Y. Jin, The 'virtual design team': simulating how organization structure and information processing tools affect team performance, in: K.M. Carley, M.J. Prietula (Eds.), Computational Organization Theory, Lawrence Erlbaum Associates, Hillsdale, NJ, 1994, pp. 1–18.
[28] A. Lomi, E.R. Larsen (Eds.), Dynamics of Organizations: Computational Modeling and Organizational Theories, MIT Press/AAAI Press/Live Oak, Boston, MA, 2001.
[29] M.W. Macy, Learning to cooperate: stochastic and tacit collusion in social exchange, American Journal of Sociology 97 (3) (1991) 808–843.
[30] J.G. March, Exploration and exploitation in organizational learning, in: M.D. Cohen, L.S. Sproull (Eds.), Organizational Learning, Sage, Thousand Oaks, CA, 1996.

[31] R.L. Moreland, L. Argote, R. Krishnan, Socially shared cognition at work: transactive memory and group performance, in: J.L. Nye, A.M. Brower (Eds.), What's Social about Social Cognition? Research on Socially Shared Cognition in Small Groups, Sage, Newbury Park, 1996, pp. 57–84.

[32] M.J. Prietula, K.M. Carley, L. Gasser (Eds.), Simulating Organizations: Computational Models of Institutions and Groups, AAAI Press/The MIT Press, Menlo Park, CA, 1998.

[33] A.M. Sastry, Problems and paradoxes in a model of punctuated organizational change, Administrative Science Quarterly 42 (2) (1997) 237–275.

[34] E. Schein, The role of founder in organizational culture, Organizational Dynamics 12 (1983) 5–23.

[35] H.A. Simon, A behavioral model of rational choice, Quarterly Journal of Economics 69 (1955) 99–118.

[36] H.A. Simon, Studying human intelligence by creating artificial intelligence, American Scientist 69 (3) (1981) 300–309.

[37] A.L. Stinchcombe, Social structure and organizations, in: J.G. March (Ed.), Handbook of Organizations, Rand McNally, Chicago, 1965, pp. 153–193.

[38] R. Suleiman, K.G. Troitzsch, N. Gilbert (Eds.), Tools and Techniques for Social Science Simulation, Physica-Verlag, Heidelberg, 2000.

[39] D.M. Wegner, Transactive memory: a contemporary analysis of the group mind, in: B. Mullen, G.R. Goethals (Eds.), Theories of Group Behavior, Springer-Verlag, New York, 1987, pp. 185–208.

[40] D.M. Wegner, A computer network model of human transactive memory, Social Cognition 13 (3) (1995) 319–339.