Simulating Information Growth & Diffusion in Agent Societies

Piotr Dollar
dollar@fas.harvard.edu
Paul Laskowski
plaskows@post.harvard.edu
Marshall Van Alstyne
mvanalst@umich.edu

The question of how to simultaneously promote growth and diffusion of ideas exhibits difficult yet pressing tradeoffs. To economists, it concerns the relationship between economic health and incentives for sharing information. When does the stimulus to innovation, founded on a profit motive, collide with access to information source material, which exhibits properties of a public good? To the legal profession, it influences what types of ideas should be intellectual property. Is society collectively improved when rights to information are broader or narrower, longer or shorter? To computer scientists, this tension is manifest as a debate in the production of software. Is an open source or proprietary model superior? To policy analysts, growth and diffusion are related to furnishing information access. Are improved communications technologies enough to bring quality information to those who seek it, and if not, why not?

Information flow systems depend on a distinctive set of factors including nonrivalry, incentives, agent rationality, topology, random interactions, access, productivity, and environments. Variations among these factors yield a diversity of emergent behaviors. In response to the complexity of practices and policies that explore these factors, we have developed an information growth and diffusion (IGD) simulator (http://si.umich.edu/~mvanalst/iShare).

IGD is intended to support research and teaching in three ways: First, software libraries that manage network models will free researchers from starting anew each time a mathematical model is proposed, while speeding the process of inquiry. Second, a simulation environment will help scholars explore robust policies in the context of changes in preferences, technologies, and network structure. Immediate visualization of various propositions is far more effective at conveying new results than static text; standardized simulation will promote faster assimilation of ideas by a broader research community. Finally, tutorials demonstrating known results will be an effective and direct way to help introduce new students to the field. These educational benefits can extend to scholars outside the original reference discipline where an insight is first developed.

Distinctiveness of Information Economies

All agent-based models share certain unifying characteristics, but the problem of information flow exhibits its own distinctive set of behaviors. IGD is particularly well suited to dealing with this field. We briefly consider three examples of constraints that specifically affect information flow. First, we argue that exponentially many subsets are required to accurately model information flow.

Consider two agents that share knowledge with each other. Their knowledge stores should not be allowed to grow without bound by re-sharing the same information back and forth. Hence, to describe who has what information, three quantities must be specified: what amount of information each agent has a monopoly over, and what amount of information is held in common. As the number of agents increases, the number of subsets required to describe knowledge ownership increases exponentially. This approach is motivated by a paper by Van Alstyne and Brynjolfsson [Van Alstyne & Brynjolfsson, 1997].

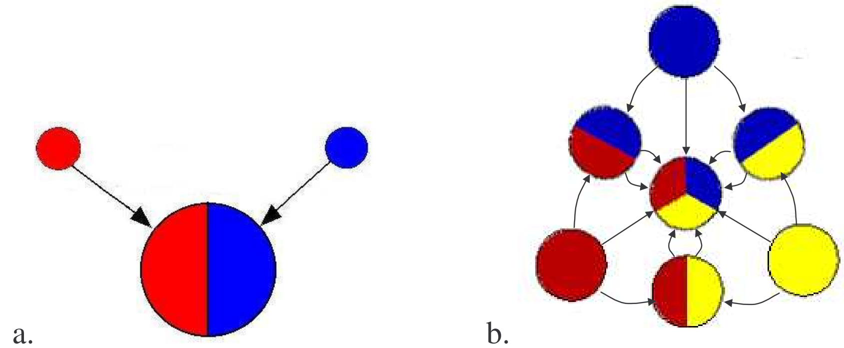

Figure 1: Knowledge sharing subsets for (a) two and (b) three agents

The IGD simulator can keep track of the knowledge unique to each subset of agents, including each subset that contains just one agent. That is, for every subset of agents, IGD records what amount of knowledge is owned by every agent in the subset, but by no agent outside the subset. This computation places no bound on the amount of knowledge that can be created by society. Alternatively, IGD can also treat knowledge as a collection of distinct facts and record which agents know which facts. This parallels the modeling of information as an agent genotype.

We next observe that in an information economy, the nature of interaction between people can have dramatic effects. Connections between people can differ in terms of their symmetry, their use for interacting with similar agents, what types of information flow across them, and how much information passes between agents. Connections can form due to deliberate actions of rational agents, or they can be exogenously determined, or they can indicate random interactions. Differences in willingness to share across connections can have a large impact on knowledge flow.

One specific instance of this compares Boston's route 128 to Silicon Valley, two high technology areas with different growth rates. Saxenian [Saxenian, 1994] presents a socio-economic explanation of why the California region, starting from the same size industrial base in 1975, had three times the employment growth and twenty five times the growth in market capitalization observed by 1990. The primary differences were attributed to several factors all related to greater information spillovers. Lower vertical integration allowed information to diffuse to buyers and sellers; less defense contracting reduced secrecy, and job-hopping and culture engendered a greater willingness to share with competitors.

The IGD simulator provides support for computing knowledge growth and flow in terms of connections between agents. Agents can connect to each other according to general societal structures, or they can choose their partners according to a suite of semi-rational strategies.

Finally, we consider the effect of geographic factors on information flow. Geographical, political, or language boundaries can limit what agents have access to what other agents. This can have dramatic effects on system behavior.

IGD agents can be placed on a 2-dimensional plane, in any one of several ordered structures or randomly. Users can then specify whom each agent has access to based on Euclidian distance. More precisely, setting access at some level, *a,* allows the average agent to connect to his *a* geographically nearest neighbors. As access increases, agents can connect to partners that are farther and farther away geographically. If access is set to the entire population size, access is universal and there is no geographical constraint at all. Increasing access simulates the effects of improved communication technologies on agent interaction. IGD allows us to explore the effects of access under a variety of agent preferences. We discover that the global behavior that results depends on the strategies that agents follow.

Utility of the IGD Simulator

The IGD simulator permits deep exploration of the relationship between local structure and global dynamics. At the agent level, users can specify agent connections manually, or as a function of the current system environment. These local connections determine how the system behaves on the global level. Furthermore, dynamic agent-choice strategies create a feedback mechanism that allows agents to alter their individual behaviors as the global environment changes through time.

The IGD simulator efficiently simulates a wide variety of agent-based models. This concept is one that Axtell, et al. have termed "alignment of computational models," or simply "model docking." [Axtell, et al., 1996] Put simply, model docking is the execution of one model with another model, or more generally, using two different models to simulate the same phenomenon in order to compare their results. The IGD simulator is designed to allow docking with a wide range of information models and to perform this task easily.

The simulator's docking ability has three main benefits. First, existing models can be replicated in IGD. We have completed this task successfully with a wide variety of models including [Carley '91; March '91; Van Alstyne & Brynjolfsson '97; and Watts & Strogatz '98]. Replication is an important part of scientific inquiry. It allows the robustness of results to be tested under a variety of inputs and assumptions.

Second, IGD allows researchers to efficiently implement new models for study. Implementing models in IGD is intuitive and requires no programming skill. Maximum flexibility is provided by the IGD simulator's system of production functions. Production functions are expressions that a researcher inputs into the IGD simulator. They describe how every component of every state vector in society changes in every time interval. Production functions are entered into the simulator in a format similar to LISP and can include mathematical operators, but also looping statements and variable declarations, giving them high flexibility. The production functions take all the knowledge components in society as inputs, as well as the connections that each agent has.

Once a model is written and ready to be run, the IGD simulator eases the process of debugging and analyzing the model. Many types of data are generated automatically and can be examined at any time, or can be graphed to show their history. These graphs are visual aids that should give the user a general idea of the behavior (or misbehavior) of the model. Production functions can be altered and new data can be gathered without the need for any recompiling of code.

Finally, IGD applies broadly to a comparative study of models. To motivate this discussion, note that while simulation is a powerful tool for conducting scientific inquiry, it has certain drawbacks. By definition, a model is an approximation, or simplification, of a real system, and is based on a collection of assumptions. It has been noted that model builders face a not-uncommon embarrassment when another modeler discovers widely different results under alternative reasonable assumptions [Sterman, 1988]. What is needed is a method for recognizing faulty models, and identifying what models are special cases or generalities of other models. In the IGD simulator, different models can be run side-by-side. Parameters can be easily altered and the changes in behavior can immediately be seen in the built in graphs. Relationships can be identified between different parameters and between different models. In this way, the IGD simulator makes it easier than ever to identify what assumptions dictate model behavior, reject fragile behaviors, and compare models to determine which is the most appropriate for answering different questions. Additionally, perhaps the largest obstacle to creating a real comparative study of models is the large amount of time required to implement each one. The IGD simulator aims to make this process easy, so that model behavior can be studied in the context of other models and their underlying assumptions.

IGD Target Audiences

We designed the IGD simulator both to be a foundation on which to study and compare particular models, and also a communications tool, with resources for demonstrating the properties of different models through interactive tutorials. These tutorials can deal with a wide array of topics and combine text with visual aids and hands-on experiments. The audiences that can use these features include students, teachers, and researchers. Each of these groups will use a different scope of the simulator's functions and capabilities.

Students will benefit from the simulator's ability to clearly demonstrate information phenomena. The interaction between social structure and information flow can be very complex and difficult to illustrate. The IGD simulator provides a visual and hands-on method of understanding these complex relationships.

Teachers may use the tool's scripting capability to develop new scenarios based on existing structure. In this mode, alternately configured and differentially endowed societies can be created and saved, then recalled easily for assigned exercises. This involves choices among elements of a general-purpose structure in order to present a coherent lesson.

Finally, researchers will make use of the full capabilities of the IGD simulator. They will also have access to the source code and will therefore be able to alter and develop new structure, adapting it to their own novel research questions. The IGD simulator is an open source project, and was constructed with future additions in mind. The source code is fully modular, and the basic structures are made to be as versatile as possible. Figure 2 contains a graphical representation of the program structure.

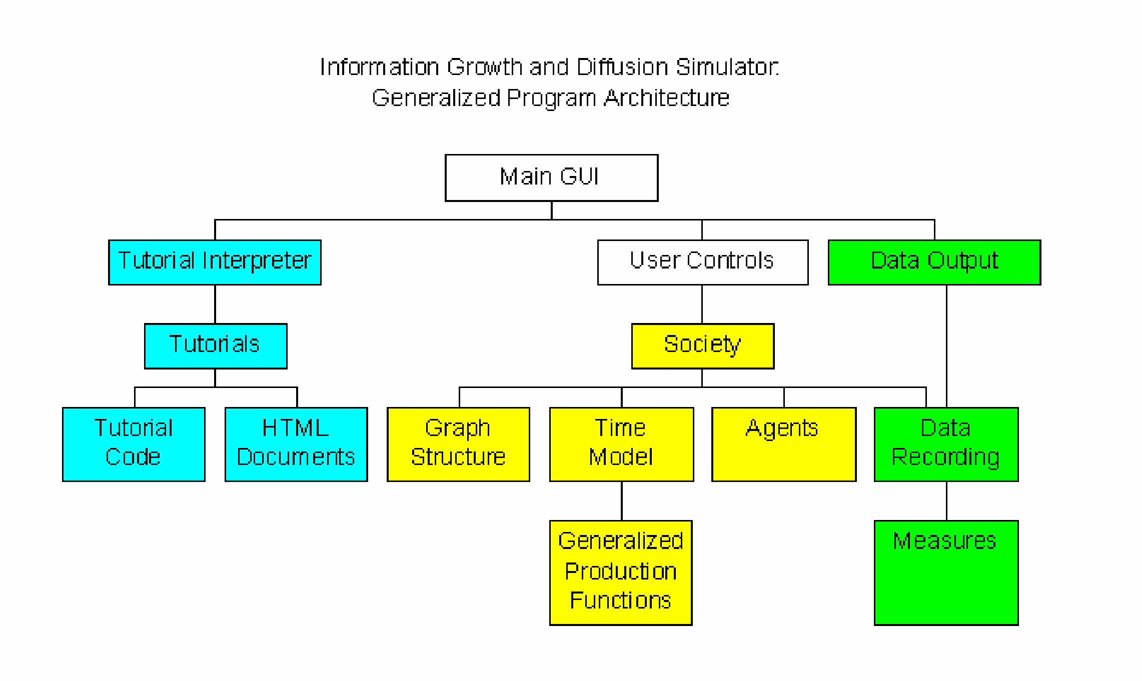

Figure 2: IGD Program Architecture

Often times, upgrading the program only involves adding a few modules and doesn't require an understanding of the entire program structure. The code is written in the popular Java (Sun Microsystems) programming language, which is particularly adept at creating clean, modular code. If a researcher has frequent modeling needs, the initial investment of basic familiarity with the code will enable easy upgrading.

To facilitate formation of an open source community, the simulator will be released under a modified or blended open source license. The rights of anyone submitting an extension are increased to allow for a brief proprietary or commercial interest of up to one year. Both firms and scholars are invited to submit their research, tutorials, and model extensions for peer review and subsequent inclusion. Upon inclusion, other scholars may build upon the tool set subject to a *temporary* commercial interest of the author who submits extensions.

References

[Arrow, 1962] Arrow, K. (1962). Economic Welfare and the Allocation of Resources for Invention. In R. Nelson (Ed.), *The Rate and Direction of Inventive Activity*, (pp. 609-625). New Jersey: Princeton University Press.

[Axtell, et al., 1996] Axtell, R., Axelrod, R., Epstein, J., & Cohen, M. (1996) "Aligning Simulation Models: A Case Study and Results." *Computational and Mathematical Organization Theory* 1(2): 123-141.

[Axelrod, 1997] Axelrod, R. (1997) "Advancing the Art of Simulation in the Social Sciences." In R. Hegselmann & P. Terna (Eds.), *Simulating Social Phenomena*, Berlin: Springer. (pp. 21-40).

[Carley, 1991] Carley, K. (1991) "A Theory of Group Stability," *American Sociological Review* v.56 June: 331-354.

[March, 1991] March, J. (1991) "Exploration and Exploitation in Organizational learning" Organization Science 2(1). Feb.

[Saxenian, 1994] Saxenian, AnnaLee (1994) Regional Advantage: Culture and Competition in Silicon Valley and Route 128, Harvard University Press.

[Sterman, 1988] Sterman, J. (1988) A Skeptic's Guide to Computer Models. In Barney, G. O. et al. (eds.), *Managing a Nation: The MicrocomputerSoftware Catalog*. Boulder, CO: Westview Press, (pp. 209-229)

[Van Alstyne & Brynjolfson, 1997] Van Alstyne, M. & Brynjolfsson, E. (1997) "Communications Networks and the Rise of an Information Elite -- Do Computers Help the Rich get Richer?" *Chapter 3, Managing Information: Networks, Value, Policy & Principles*. MIT Ph.D Thesis; Cambridge, MA.

[Watts & Strogatz, 1998] Watts, D. * Strogatz, S. (1998) "Collective Dynamics of Small World Networks?" *Nature*. 393:440-442