9

ENGINEERING THE PROCESS OF INSTITUTIONAL INNOVATION IN CONTESTED TERRITORY

Russell C. Thomas[1,2] and John S. Gero[1,2]
[1] *Department of Computational and Data Sciences, George Mason University, Fairfax, VA*
[2] *University of North Carolina Charlotte, Charlotte, NC*

9.1 Introduction

Typically, when we talk about an 'engineering approach' to a given social system, we are referring to the transformation of the current state of that social system to some future state that meets needs, requirements, or goals. The 'end product' of social systems engineering is the transformed social system. But what if the social system is an innovation community and the 'end product' is innovation or discovery of new knowledge? In this case, we are no longer applying an established body of knowledge and methods to solve a given problem, but instead the social system is trying to expand the state of knowledge and generate inventions that are genuinely novel and distinctly beneficial. Given the fundamental uncertainties of innovation and discovery, we can't reliably predict the end results or even the trajectory of innovation that will unfold.

Further, we consider settings that are *contested territory*—i.e. there are rival world views regarding the nature of the problems and the nature of innovations need to address those problems. With two or more rival views ('schools of thought'), there might be diverging innovation trajectories where each school of thought develops and exploits knowledge that fits that school but not rival schools. For an example of this in scientific knowledge, consider the rival theories of infectious disease in the early to mid 1800s: the Germ Theory versus the long-established Miasma Theory (i.e. noxious air due to rotting organic matter, etc.). In this case, there was very little synergy or mutual benefit in their knowledge and methods. Microscopes, microbe cultures, and statistical epidemiology were all tools and methods that advanced Germ Theory but had no relevance to people who were interested in advancing Miasma Theory. The same divergence can happen in institutional innovation.

Is it possible to apply an engineering approach to the *process* of institutional innovation? If so, how might it help promote innovation or influence the trajectory? Cyber security will be our motivating case, described in some detail in Section 16.2. Cyber security is one of a large class of complex socio-technical systems that are characterised by low probability/high cost loss events and interdependent risk. Policy makers have recognised since 2003 that we do not have adequate institutions to manage cyber security and mitigate risk, and therefore they have called for 'leap-ahead' innovation, especially in incentive institutions and quantitative measurement systems.

Unfortunately the pace of institutional innovation has been slow—much slower than the pace of innovation in information technology and by threat agents.

Institutions are norms and 'rules of the game' that support and enable social life (Scott 2007; North 1990). They can be explicit or tacit, formal or informal. Chisholm (1995) gives these examples: legislatures, bureaucracies, corporations, marriage, insurance, wage labor, the vacation, academic tenure, and elections. An *institutional innovation* is a significant change or improvement in an institution such that it has novel functional or performance characteristics. An example is consumer credit scoring (Ryan, Trumbull, and Tufano 2011; Marron 2007). The incumbent institution had been in place for hundreds of years: credit managers exercising subjective judgment of each consumer's 'character' and personal history. In its mature form, the institution of consumer credit scoring replaced personal judgment with algorithmic judgment, and replaced individual evaluation rules with population and portfolio evaluation rules. It was seen as legitimate, supported the values of consumerism and the information age, fit well with other institutions (e.g. credit reporting, financial risk management, etc.), and provided stability that enabled fast growth in both consumer credit and firms that depend on consumer credit.

Compared to technological innovation, much less is known about institutional innovation. Based on historical cases, we can say that institutional innovation can happen intentionally and by design, unintentionally through collective or accidental processes, or by a combination of the two. *Institutional entrepreneurs* are actors aim to achieve innovation intentionally, but they rarely control or determine the entire process, in contrast to technology inventors or entrepreneurs. To succeed, any institutional innovation must fulfill these desiderata (Battilana, Leca and Boxenbaum 2009; Weik 2011; Leca, Battilana and Boxenbaum 2008; Scott 2007; North 1990):

1. *Functional*: Does it work? Does it perform?
2. *Feasibility*: Is there a viable evolutionary path from 'here' to 'there'?
3. *Legitimacy*: Does it flow from legitimate authorities, as seen by social actors?
4. *Cultural Fit*: Does it fit and reinforce the society's values?
5. *Coherence*: Does it interrelate effectively with other institutions?
6. *Uncertainty Reduction*: Does it make social life more predictable or less risky?
7. *Stability*: Ultimately, does it make society more stable and resilient?

Given that institutional entrepreneurs are relatively 'blind' regarding most feasible or desirable paths of innovation, how do they take meaningful action? The thesis of this chapter is that institutional entrepreneurs *engineer the process* rather than the end result. By this we mean that they apply an engineering approach to design, build and use *knowledge artefacts* – e.g. dictionaries, taxonomies, conceptual frameworks, formal procedures, digital information systems, tools, instruments, etc. – as cognitive and social scaffolding to support iterative refinement and development of partially developed ideas. Our hypothesis is that the rate of innovation progress will critically depend on the quality and nature of the knowledge artefacts at their disposal.

The plan for the chapter is as follows. Section 16.2 describes the institutional innovation problem to be solved and describes the two schools of thought regarding how best to solve the problem. Section 16.3 presents a theoretical model of social processes of innovation, and explains why rival schools of thought arise in nascent fields where there is no established base of scientific knowledge or methods. Sections 16.4 and 16.4 present a computational model of innovation that will be used to illuminate the dynamics of social innovation aided by knowledge artefacts and to identify conditions under which one or the other rival school of thought is likely to prevail. The chapter closes with a discussion of the main findings and implications for institutional entrepreneurs.

9.2 Can Cyber Security and Risk Be Quantified?

Though there is no settled definition for 'cyber security', for our purposes we will define it as the confluence of information security, digital privacy, digital civil rights, digital (trusted) identity, digital (content) rights management, digital information protection, and the digital aspects of homeland and national security. Given the pervasive and vital role of information and communication technology (ICT) in modern life, cyber security affects every organization and government, plus a large and growing proportion of individuals worldwide.

Cyber security is a vexing problem. Many problematic aspects of cyber security are sociological, economic, political, and cultural. This has been well known for over a decade, leading to many policy reports and research funding solicitations that call for research and innovation in these domains (National Science and Technology Council 2011; Department of Homeland Security 2011). Many institutional innovations have been discussed or proposed, including some based on analogies from existing institutions in other domains[1]. Unfortunately, innovation progress to date has not been satisfactory.

9.2.1 Schools of Thought

With some oversimplification, we can identify two broad schools of thought regarding institutional innovation in cyber security: 1) the 'Quants' who believe that cyber security and risk can and should be quantified in ways similar to other domains involving socioeconomic-technical risk (Geer, Soo Hoo and Jaquith 2003); and 2) the 'Non-quants' who believe that cyber security and risk either cannot be quantified or that there is no net benefit compared to alternate methods of guiding or structuring action, decisions, rules, etc. Examples of non-quantitative methods include checklists, audit questions and procedures, policy and practice guidelines, and situational professional judgment (Langner and Pederson 2013). There is also third school of thought we might call 'Hybrid' in that they believe some degree of quantification can be usefully combined with non-quantitative approaches.

The degree of difference between these schools of thought varies by what is being quantified and how that quantification is used in analysis and decision-making. Where they differ least is in operational security – e.g. the uptime of a network firewall, the false positive rate for spam filters, etc. Where they differ most is on risk quantification, i.e. can we measure risk in economic units in a way that will guide investment decisions or serve as foundation for cyber insurance or other incentive contracts? We will focus our attention on risk quantification because it makes vivid the contest between these schools of thought. The quantification of cyber security and risk is an intellectual and social domain where control and influence is contested by interest groups. Focusing on risk quantification, examples of interest groups associated with the Quant school include the Society of Information Risk Analysts and companies who specialize in measuring or modelling risk. Examples of Non-quant interest groups include some security consultants (Langner and Pederson 2013) and most regulators[2] . Examples of 'Hybrid' interest groups include most large information security companies, the US National Institute of Standards and Technologies (NIST) in the Department of Commerce, and ISACA, a professional organization formerly known as Information Systems Audit and Control Association. Regarding evidence of innovation success and adoption, the Quants have been struggling for more than a decade. Verendel (2009) presents a comprehensive survey of academic research up to that time, and finds that quantified cyber security is still a weakly supported hypothesis. As of 2015, no one can yet say that cyber security risk can be effectively and efficiently quantified. Even so, there has been some progress in some areas and growth in the number of people and organizations actively working on new ways to quantify risk.

The Non-quants have frequently pointed to this lack of success as evidence that quantified risk is impossible in principle (i.e. in the same way that perpetual motion machines impossible) or, at least, too complicated and expensive to invest in. Additional negative arguments come from the Financial Crisis of 2008, where sophisticated/complicated risk models have been widely blamed as one of the aggravating factors. However, the most frequent and fundamental argument against quantified cyber security risk is based on the complicating factor mentioned above – intelligent, adaptive adversaries. Though taken from a report on risk analysis for physical security of nuclear weapons complexes and not cyber security, this quote nicely summarizes the argument against quantified risk in cyber security, too:

> "The committee concluded that the solution to balancing cost, security, and operations at facilities in the nuclear weapons complex is *not to assess security risks more quantitatively or more precisely*. This is primarily because *there is no comprehensive analytical basis for defining the attack strategies that a malicious, creative, and deliberate adversary might employ or the probabilities associated with them*." [emphasis added] (National Research Council 2011, p. 1)

The Quants respond to this negative argument in a variety of ways, including a claim that lack of progress or complete success over ten years is not sufficient evidence that it can't be successful given enough time and effort. To support this claim, references are made to historical cases of the development and adoption of quantitative methods in similar domains.

9.3 Social Processes of Innovation in Pre-paradigmatic Fields

---

[1] E.g. 'Cyber CDC': Center for Disease Control; 'Cyber UL': Underwriters Laboratory; and 'Cyber NTSB': National Transportation Safety Board.

[2] In the US, one example of a Non-quant regulator is the Federal Financial Institutions Examination Council. https://www.ffiec.gov/cybersecurity.htm

Generalizing from the cyber security case, we now turn our attention to social processes of innovation in nascent fields – those that Thomas Kuhn called ‘pre-paradigmatic’. In Kuhn’s model of scientific revolutions, an established field of science is characterised by a ‘paradigm’, which is an “entire constellation of beliefs, values, techniques, and so on shared by members of a given community” (Kuhn 1970, p. 175). Established paradigms feature exemplars that serve as ideal models or templates to be emulated by subsequent research. For example, Newton’s Laws of Motion served as exemplars in Physics until the early 20th century. In contrast, ‘pre-paradigmatic’ fields are those where there are no established, widely accepted paradigms and, therefore, lack of clarity over what constitutes ‘good’ or ‘normal’ scientific research. Without the normative influence of paradigms, the discourse and debate can be unproductive. Kuhn describes it this way: “the pre-paradigm period, in particular, is regularly marked by frequent and deep debates over legitimate methods, problems, and standards of solution, though these serve rather to define schools [of thought] than to produce agreement” (Kuhn 1970, p. 47-48). Even though quantified cyber security and risk is not purely about science or scientific research, we can characterize it as being in a ‘pre-paradigmatic’ state of development.

9.3.1 Epistemic and Ontological Rivalry

Recall that we said that institutional entrepreneurs aim to achieve innovation *on purpose*, not just by chance events or through collective processes of change. Furthermore, they aim to achieve innovation in a particular direction, not just anywhere. To achieve this, they need a way of thinking about problems and solutions that enables progress. They also need to have a model of reality that enables progress. In philosophical language, we can say that institutional entrepreneurs need both an epistemology and an ontology that are *beneficial* and *instrumental* to their teleological (i.e. goal-driven) approach to innovation. In pre-paradigmatic fields such as quantified cyber security and risk, the schools of thought often feature rival or even mutually exclusive epistemologies and ontologies.
In the case of quantified cyber security and risk, the two rivalrous schools of thought— Quants vs. Non-quants— differ sharply over the ontology of cyber security and risk, i.e. what is real and what is not real. For example, some Non-quants argue that quantifying cyber risk is impossible in principle because of the non-reality of hypothetical or counterfactual events: “How is it possible, they say, to quantify what didn’t happen?” (Borg 2009, p. 107). There is considerable disagreement over the ontological status of ‘intangible’ losses such as reputation. Also there is ontological debate carried over from Mathematics and Statistics concerning the reality or non-reality of subjectivist interpretations of probability. There is even dispute over the methods or possibility of ever resolving these ontological debates. For example, many Quants are in favor of computer simulations as tools to explore counterfactual or hypothetical situations, while many Non-quants argue against the validity of computer simulations as evidence in ontological arguments, given the intelligent, adaptive, creative, and malicious adversaries.

Likewise, there is sharp disagreement between Quants and Non-quants regarding the best way to think about cyber security and risk. Quants argue that quantification and quantitative analysis can be a powerful tool to make better decisions and achieve better outcomes, much in the same way that Statistical Process Control and Total Quality Management has helped revolutionize product and service quality across many industries from the 1980s to present. Some Non-quants counter with the argument that attempting to quantify abstract and non-real entities such as ‘risk’ is not only a waste of time and effort, but that it leads to *worse* outcomes through “analysis paralysis” or mistaken efforts to “manage” risk (Langner and Pederson 2013).

From the perspective of Sociology of Innovation, we are less concerned with the ultimate truth of any of these positions than we are with their functional and instrumental effects, i.e. *are they effective in helping the actors to achieve progress?* This leads to the next topic: what *knowledge artefacts* do institutional entrepreneurs develop and use in during the innovation process and how do they promote progress?

9.3.2 Knowledge Artefacts

A *knowledge artefact* is something created by actors informed by their knowledge and makes that knowledge useful or productive. A knowledge artefact can be a thing (i.e. a message, a book, a tool, a design, etc.) or a realizable process (i.e. a training process, a production process, a communication process, and information processing process, a utilization process, etc.). Thus, it is through knowledge artefacts that people create, transform, and use knowledge for practical aims. This is not meant to reify knowledge. Instead, knowledge artefacts can be seen as the tangible, observable instantiations of knowledge, much like a circle drawn on a sheet of paper is an instantiation of the Platonic idea of ‘circle’. Boisot (1999) uses a similar term—‘knowledge assets’—which he defines as “knowledge that yields an appropriable stream of benefits over time” (p 155). However, this definition presumes that we can point to, instantiate, define or specify the knowledge in question, which can be problematic. Instead, we prefer to use the term 'knowledge *artefact*' to highlight the point that knowledge artefacts are products of human intention and effort and that they can be observed and instantiated, at least in principle. We retain Boisot's notion of "appropriable stream of benefits over time" through the emphasis on instrumentality.

Boisot (1995) developed the Information Space (I-Space) framework for characterizing knowledge and knowledge artefacts along three dimensions: 1) Codification, 2) Abstraction, and 3) Diffusion. The main purpose of the I-Space framework is to study the transformation of knowledge through life cycles of discovery, learning, and diffusion. Our focus will be on the first two dimensions. The *Codification* dimension evaluates knowledge in terms its degree of compression or abbreviation within some coding scheme such as categories, taxonomies, variables, conditions, relations, and so on. For any given knowledge, expressing it in a highly codified will be very compressed and economical. In contrast, uncodified knowledge may take many more words to express, or may even be only learned through experience or example (i.e. 'tacit knowledge'). The Abstraction dimension evaluates knowledge in terms of the inferences you can draw from it, and the degree of generality regarding inferences, ranging from concrete (highly specific and contextual) to abstract (highly general and free of context).

9.3.3 Implications of Theory

Before moving to next section, we can summarize the implications of these theories in relation to our case and also to the general study of institutional innovation in pre-paradigmatic fields. First, institutional entrepreneurs in rival schools of thought are engaged in a contest between each other and also with Nature regarding who has the best way to think about problems and solutions (epistemology) and whose model of reality is most effective (ontology). While this contest plays out in many ways that are mostly or purely social, there is also a contest in the practical world of realizing inventions and innovation. It is not enough to talk a good game or convince many others. Eventually, some inventions work and others do not. Those that work and can be used and understood by the masses will get widely adopted. But institutional entrepreneurs often start in a fog of uncertainty and ignorance, even if some insights, intuitions, role models, or goals guide them. Therefore, they create and use knowledge artefacts that have several uses at once. They solve some immediate problem while providing some foundation or platform for further invention or knowledge creation/transformation. In this way, knowledge artefacts can serve as *cognitive scaffolding* (Lane and Maxfield 2005) to help the institutional entrepreneurs make progress in the face of ignorance and uncertainty. We can usefully characterize their knowledge artefacts along the dimensions of Codification and Abstraction. Boisot's Social Learning Cycle (SLC) theory predicts that insights that trigger innovation cycles start out as tacit, hard-to-explain and concrete/specific. SLC predicts that knowledge artefacts will be developed and unused in a specific sequence: *first* they will be increasingly codified (i.e. through formal definitions, taxonomies, measurement systems), and *then* they will be increasingly generalised thorough more abstract sign systems, relation systems, and inference systems.

In the case of quantified cyber security and risk, we can position specific knowledge artefacts in the I-Space, shown in this table:

Artefact	School of Thought	Codification	Abstraction
1) Ad hoc security metrics	Quant	Low	Low
2) NIST Cyber Security Framework (CSF)	Non-quant	Moderate	Moderate-Low
3) Risk analysis software	Quant	High	Moderate-High

Table 9.1 List of example knowledge artifacts and their position in the I-space

Here is the rationale, starting with 1) Ad hoc security metrics. Many medium-sized and large companies have a dedicated information security department and many of these collect and report 'security metrics' to company executives in regularly scheduled reports. Evaluated as artefacts for knowledge relating to quantified security and risk, most of these reports are low in Codification because they do not follow any well-defined taxonomy for what should be measured and reported. Also they are relatively low in Abstraction because the rules or logic as to how the different metrics might be combined or interpreted together is mostly in the form of heuristics. Regarding 2) US National Institute of Standards and Technology (NIST) Cyber Security Framework (CSF), we can locate it as moderate in Codification because it does attempt to define key phenomena and conditions in cyber security and risk, but in itself it does not attempt to quantify security or risk. It is moderately low in Abstraction since it mostly points to classes of phenomena and does not embody any specific theory or knowledge as to how security is achieved or risk reduced through the implementation of the 'best practices'. Finally, regarding 3) Risk analysis software, there are several commercial software products and services that quantify some aspect of cyber security and risk. In terms of the I-Space we can locate them as high in Codification, certainly much higher than the NIST-CSF, and moderately high in Abstraction, since they embody formalised knowledge regarding how quantitative inferences are to be drawn from evidence (i.e. 'ground-truth data').

From the point of view of the SLC, we can see that there are (at least) two innovation and learning cycles at work. The Non-quant learning cycle is represented by the NISTCSF, and is explicitly aimed to be a viable stage of refinement, performance improvement, and usability that can support wide spread diffusion and adoption. The Quant learning cycle is aiming for a much higher goal in terms of Codification and Abstraction, and thus wide spread diffusion is not yet happening, or maybe it is just beginning.

In summary, the contest between rival schools of thought in a pre-paradigmatic field such as cyber security can be viewed as different navigational strategies through I-Space. The Non-quant school is aiming for a lower region in I-Space (i.e. less Codification, less Abstraction), betting that this will be more feasible and will achieve practical success and wider adoption, compared to the higher road of the Quants. Conversely, the Quants are betting that the high road (i.e. more Codification, more Abstraction), though more difficult to traverse, will ultimately lead to more compelling results – better security, lower risk, and better use of societies resources. Note that the I-Space framework and SLC theory do not represent the space of possible inventions because they do not account for the specific traits, characteristics, or dependencies of each invention. Therefore, I-Space and SLC not facilitate analysis of how difficult it may be to go from any point 'A' to any other point 'B' in the space of possible inventions. We address this in a computational model, presented in the next section.

9.4 A Computational Model of Innovation

In this section our goal is to demonstrate how computational modelling can be used to investigate institutional innovation in contested territory, and the effects of knowledge artefacts. In the specific case of quantified cyber security, we don't yet know who will win: the Quants or the Non-quants. Given the time span of institutional innovation, we may not know for many years. By using computer simulation, we can examine a generalised abstract model of innovation and perhaps learn more about the conditions under which one or the other rival school of thought is likely to prevail.

To model the phenomena of interest, we need a way to model the space of possible inventions. We also need a way to model the relative difficulty of achieving each invention, both with respect to making the final discoveries or solving the final problems, but also to the inventions that came before (i.e. precursors and dependencies). Finally, we need a way to model the relative effects of knowledge artefacts as characterised by I-Space.

9.4.1 Base Model: Innovation as Percolation

To meet these requirements in a parsimonious way, we chose to develop a *percolation model of innovation* based on the model presented in Silverberg and Verspagen (2005) ("S&V"). 'Percolation' is the phenomena of fluid moving or filtering through porous materials. Percolation modelling originates in the fields of Physics, Chemistry and Materials Science, and has been abstracted in Mathematics as Percolation Theory. In the model of Silverberg and Verspagen (2005) , the 'porous material' is taken to represent the space of possible inventions, the 'fluid' is taken to represent the advancing front of innovation ('best practice frontier') in that space, and the local dynamics of percolation are taken to represent the local dynamics of innovation. We adopt the S&V's term 'technology' to mean any solution, method, process, procedure, tool, or machine, and also the term 'R&D' means inventive activity, whether formal or informal. S&V use the term 'firms' but we prefer the more general and abstract term 'agents' to refer to localised bundles of inventive activity, be it a person, a team of people, a firm, or some mixture.

For readers not familiar with agent-based modelling (ABM), here is a basic overview. An ABM consists of an environment and a set of agents that operate and interact within that environment. An 'agent' is a simple program that has a set of behaviour rules, a memory (i.e. internal state), a position within the environment, and runs once each time step. Generally, all agents run the same program, with the only difference being the agent's internal state, it's location, and the state of the local environment. The following is pseudo-code for a generic agent (single step):

```
input: Internal_state, Location, Environment
output: Internal_state, Location, Environment
begin
  local_state ← Sense_local_environment (Environment, Location)
  foreach Behavior_rules
    if (match(local_state, rule_condition)
       then rule_behavior(Internal_state, Location, Environment)
  end
end
```

Figure 9.1 Generic agent algorithm for a single time step.

In our model, the 'agents' represent R&D effort in a particular technology type (column). There is one agent per technology type. Each possible technology is a cells connected in a discrete two dimensional lattice[3] (i.e. grid, see Figure 9.2). Each cell in the lattice has horizontal neighbors that are very similar and interrelated, and vertical neighbors that are slightly more or less sophisticated. Overall, the neighborhood structure reflects technological interrelatedness. Considering the horizontal dimension, each column represents a 'technology type', all sorted to the most similar types are next to each other. Considering the vertical dimension, each row represents a degree of sophistication, from the minimal 'baseline' at the bottom, rising monotonically without bound in principle, but limited to a maximum size to fit constraints of computer processing. The lattice is connected with a periodic boundary the horizontal dimension so that it has a cylindrical topology. This allows every technology type (column) to have exactly two neighbors and eliminates horizontal boundary effects in the model.

Formally, we define a lattice A with h columns and periodic boundary in the horizontal dimension (i.e. cylinder topology), v rows, and $h \times v = N$ cells indexed by i and j, $0 < i < h$ and $0 < i < v$. Parameters $h > 0$, $v > 0$ are set by the experimenter to be large enough so that the boundaries of the lattice do not influence the results involving rates of innovation and distribution of sizes of innovation. Each lattice cell $a_{i,j}$ can be in one of four states: 0 = impossible (black); 1 = possible but not yet discovered (white); 2 = discovered but not yet viable (light grey); and 3 = discovered and viable (mid-grey).

Through R&D activity of agents, some cells near the baseline are discovered and therefore move from state 1 to state 2. These newly discovered cells become the 'adjacent possible' (Kauffman 1996) meaning they are the next candidates for becoming viable technologies (state 3). Any discovered cell (state 2) becomes viable (state 3) when there is a contiguous Manhattan[4] path from it to the baseline. Sites initialised as impossible (state 0) can never be converted into any other state.

[3] A 2D lattice is chosen for simplicity, but also Silverberg & Verspagen (2005) say that they believe their main results will hold for more general topologies.

[4] A 'Manhattan' path is a series of up-or-down and left-or-right steps. This implies a Von Neumann neighborhood for each cell, i.e. only the cells reachable with Manhattan distance of n steps. An alternative definition for path is 'chessboard distance', which include any combination of up-or-down, left-or-right, or diagonal steps. This would imply a Moore neighborhood for each cell. If we switch to chessboard distance and Moore neighbor- hood, this provides more connectivity between cells and increases the number of possible-but-not yet-discovered technologies (state = 1, white) to explore. The main effect is to increase the overall rate of innovation because 'impossible' regions are more easily traversed. However, the qualitative results are not different from current model with Manhattan path and Von Neumann neighborhood.

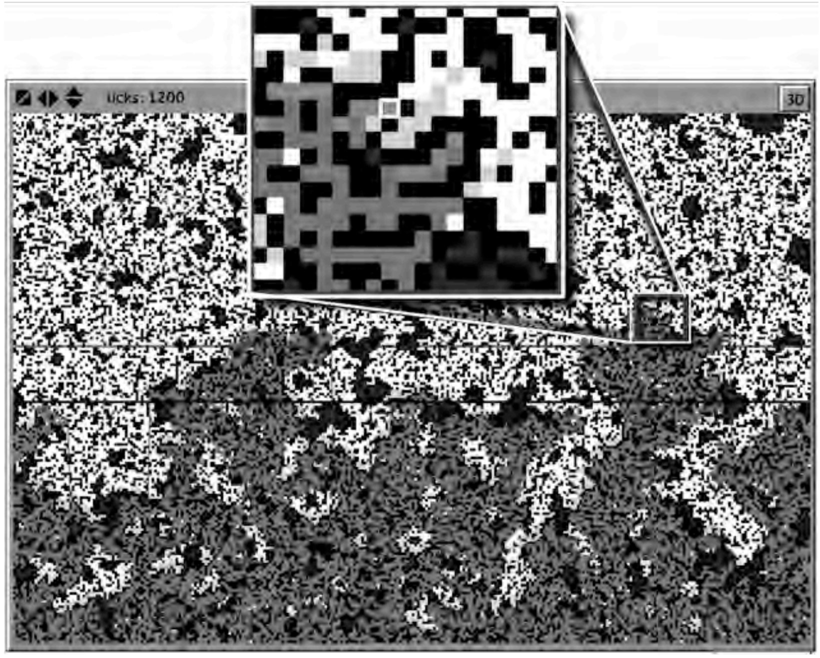

Figure 9.2 The 2D lattice of 'technologies' in our percolation model of innovation after 1,200 simulation steps. One region in the lattice around the best-practice frontier (BPF) is magnified. See Figure 9.3 for details. Grey scale code: black = impossible, white = possible but not yet discovered; very dark grey = possible, but not reachable; lightest grey = discovered but not yet viable; medium-light grey = discovered and viable; medium-dark grey = discovered, viable, and on best-practice frontier (BPF). Medium-dark grey cells are the loci of R&D. The two horizontal lines show the "average" level of innovation (i.e. BPF) across all 'technologies' (columns). The dark line is the average (i.e. mean) BPF, while the dark grey line is the mean + standard deviation of the BPF.

The probability that any cell will be initialised in state = 1 (possible) rather than state = 0 (impossible) is given by parameter p. If all paths from a given possible cell to the baseline are blocked by impossible cells, then we say that these cells are *not accessible*. There is a critical value for $p \approx 0.6$. Much above 0.6 and nearly every possible cell becomes accessible. Much below 0.6 and nearly every possible cell is not accessible. When $0.6 \leq p \leq 0.65$, each random realization produces a different complex pattern of accessibility. For all of our simulation runs, we set $p = 0.62$.

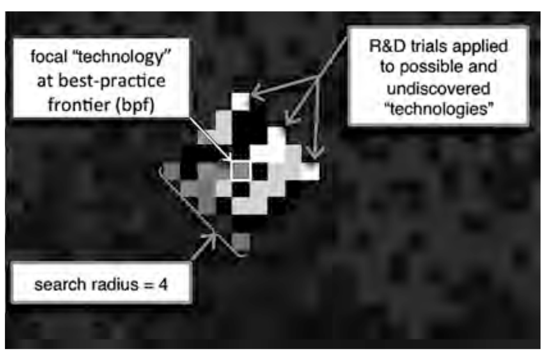

Figure 9.3 Detailed view of the magnified region in Figure 9.2.

Every technology type (column) has an individual technology (cell) that is the most advanced (highest in the column), and this cell has *BPF* = true. All other cells have *BPF* = false. All R&D activity take place in technologies (cells) around the best practice frontier (*BPF*), i.e. *BPF* = true. BPF moves vertically in each column as R&D is successful. All R&D takes place within the search radius around the technologies (cells) at the BPF (Figure 9.5). R&D activity consists of each agent expending a budgeted amount of effort to 'discover' cells that were previously 'undiscovered' within their search radius. This is realised through probability of discovery that is the search effort e divided by the size of the search area. (In our simulation runs, $e = 0.5$ and radius = 4, which yields a search area of 40 technologies (cells) and a probability of discovery $p = 0.0125$ per R&D attempt.) Thus, in the base model of S&V, an ‘ R&D attempt’ is realised as a random draw from (0, 1), and if this is less than the probability of discovery, the technology is ‘discovered’. Here is pseudo-code for this algorithm:

```
input: Location, Environment
output: Environment
begin
  local_cells ← get_neighboring_cells (Environment, Location)
  possible_cells ← { }
  foreach cell in local_cells
    if (get_state(cell) = 1) then append(possible_cells, cell)
  end
  foreach cell in possible_cells
     draw ← random-uniform(0,1)
     if (draw ≤ 0.0125) then set_state(cell, 2)
end
```

Figure 9.4 Algorithm for R&D activity in a single time step

After all agents have executed this algorithm, an Environment program runs to identify newly discovered technologies (cells with state = 2) that are now viable. If any are found, those cells have their state is set to 3. The best practice frontier (BPF) is adjusted by setting BPF = true for highest viable cell in each column. Agents are moved to the BPF cell in their column.

The ‘size’ of any innovation is the number of rows between the newly discovered technology and the previous best-practice frontier technology in that column. In other words, large innovations make a big ‘leap’ in the vertical dimension, while small innovations might only move up one cell.

9.4.2 Full Model: Innovation with Knowledge Artefacts

We made several extensions to the base model to simulate the effects[5] of knowledge artefacts and learning. The effect of knowledge artefacts is to improve the effectiveness of R&D, but only if the knowledge appropriately matches the domain, i.e. effectiveness is proportional to their *fidelity* relative to the Nature and also relative to the social and technical context of innovation.

In our extension, there is only one knowledge artefact available to all agents with parameters set at initialization time corresponding to their characteristics in I-Space: Codification $c \in (0 \ldots 5)$ and Abstraction $a \in (0 \ldots 5)$. For simplicity, these parameters take integer values. We model the effects of knowledge artefacts by through a sub-model of the R&D activity: a random draw from multiple balls-in-urns instead of the uniform random number draw in the base model. Each undiscovered technology (cell) starts the simulation with a large but finite number of ‘balls’ ($N = 1,000$), which are possible solutions that are either ‘successful’ (13 balls) or ‘unsuccessful’ (987 balls). These balls are distributed in a number of urns determined by the Codification parameter c. If $c = 0$, then all the balls are in a single urn, and this is equivalent of blind ‘trial and error’ search. If $c = 5$ (maximum value), then the balls are allocated to 32 urns, with all the ‘successful’ balls in one urn (13 out of 32 balls in that ‘lucky urn’). The effect of Abstraction is that it increases the probability that the agent will select the ‘lucky urn’. If a given agent is aided by a high fidelity knowledge artefact with high abstraction, then they will most likely select the ‘lucky’ urn containing all the ‘successful’ balls. But if either the Fidelity parameter $f \in (0 \ldots 1)$ is zero or Abstraction $a = 0$, then the agent will be choosing among the 32 urns with uniform probability, which again is equivalent to blind ‘trial and error’ search. While both c and a are fixed for the duration of a simulation run, f can change if the Learning rate parameter $l \in (-1 \ldots 1)$ is not zero. If $l > 0$, then Fidelity f increases during the run as a function of the height of the BPF, and conversely if $l < 0$, then Fidelity f decreases during the run as a function of the height of the BPF. This allows us to simulate scenarios where a knowledge artefact is initially appropriate and effective in guiding innovation but decreases in appropriateness and effectiveness as technologies get more advanced/sophisticated.

[5] For simplicity, we are only simulating the *effects* of knowledge artefacts with different I-Space characteristics rather than the specific contents or traits of the knowledge artefacts themselves.

9.4.3 Experiment

We design three experimental treatments that, in an abstract way, represent the different schools of thought (Quant vs. Non-quant) and the differences in the knowledge artefacts they are attempting to create and use. Recall that the Non-quants are building and using knowledge artefacts with less Codification and less Abstraction, believing that this will be more feasible and will achieve practical success soon and therefore wide adoption soon, too. Conversely, the Quants are developing and using knowledge artefacts with higher Codification and higher Abstraction will be more successful to promote innovation, though progress may be more difficult to achieve initially. We do not know whether the Quants will learn rapidly or slowly (i.e. adapt and refine their knowledge artefacts), and therefore we will define separate experimental treatments for each scenario. We add 'control' treatment with no knowledge artefact, resulting in four treatments total:

1. *Trial-and-error* with no knowledge artefact
2. *Non-quant* with initial knowledge artefact parameters: Codification = 2, Abstraction = 2, Fidelity = 1.0, and Learning Rate = 0.0
3. *Quant – slow learning* with initial knowledge artefact parameters: Codification = 4, Abstraction = 4, Fidelity = 0.2, and Learning Rate = 0.2
4. *Quant – fast learning* with initial knowledge artefact parameters: Codification = 4, Abstraction = 4, Fidelity = 0.0, and Learning Rate = 1.0

Figure 9.5 shows a single run and series of screen shots at different time steps, along with a time series chart of the innovation rate (i.e. change in BPF per time step). Figure 9.6 compares two experimental treatments on the same initial lattice configuration.

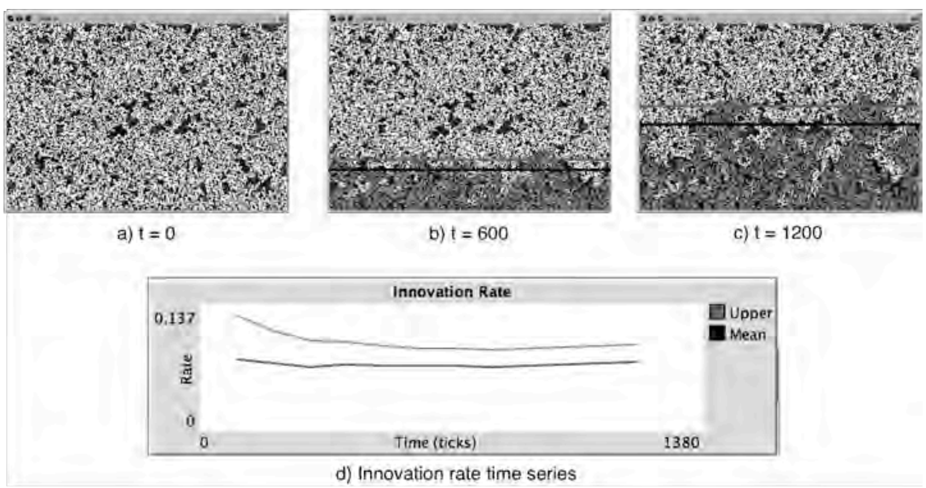

Figure 9.5 A single run at different time steps, with a time series chart showing Innovation Rate for Mean BPF (black horizontal line in b) and c)) and Upper BPF (mean + standard deviation, dark grey horizontal line in b) and c)).

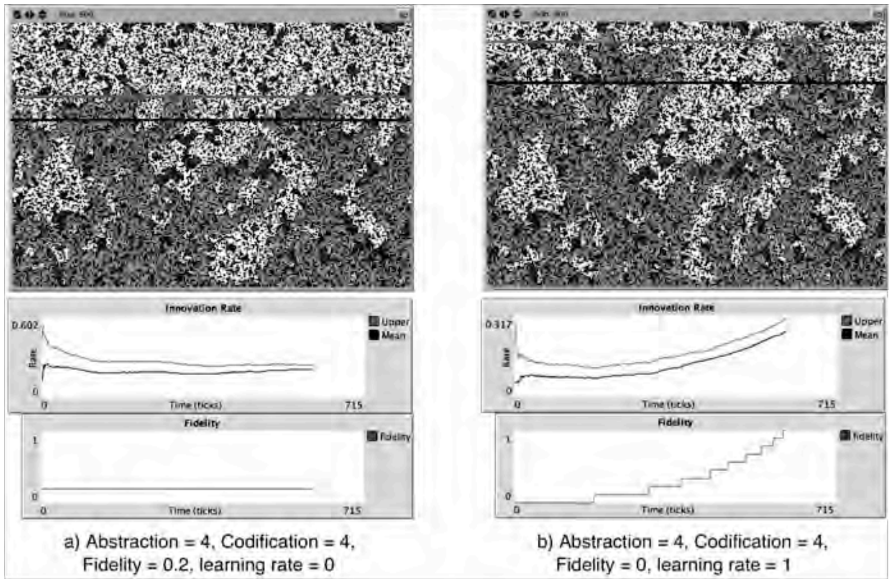

Figure 9.6 Screen shots of two treatments tested on the same lattice configuration: a) Quant–slow learning vs. b) Quant–fast learning. In b), the Innovation Rate increases as Fidelity increases due to learning. Therefore, even though it started out slower, the Quant– fast learning treatment wins this innovation race. (The black horizontal lines are mean BPF. The dark grey horizontal lines are Upper BPF = mean + standard deviation.)

Each of the four experimental treatments were tested with the same random initial lattices, ten in total, with twenty runs for each lattice condition using different random seeds for each run. Each run ended when the best-practice frontier (BPF) reached the highest row in the lattice. Figures 9.7a and b show violin plots for the experiment results for two dependent variables: 1) Innovation Rate at the end of the run and 2) Time to Complete a run (i.e. the BPF reaches the top of the lattice).

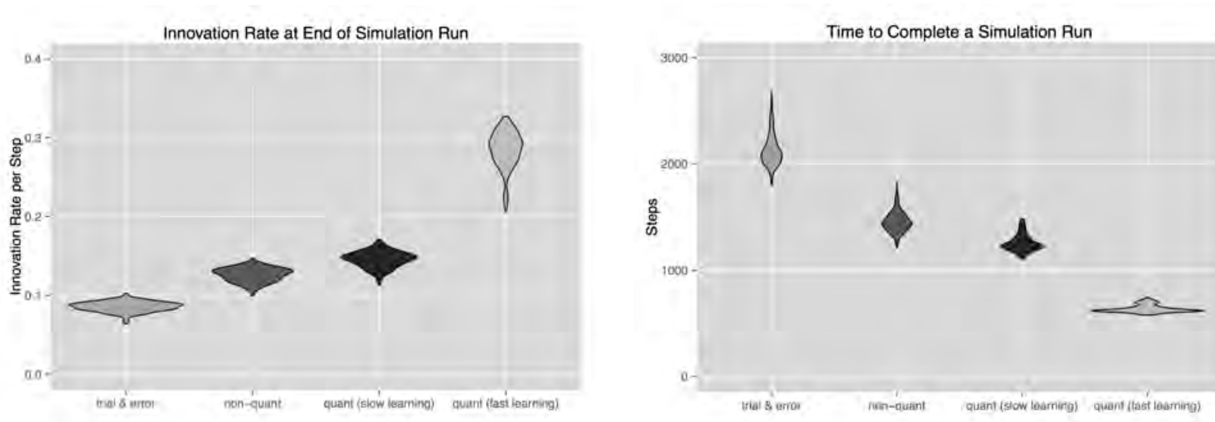

Figure 9.7 Violin plot of experiment results. Experiment results for four treatments, each tested with 10 random lattice initial conditions, 20 runs per lattice condition. a) Innovation rate at the end of simulation run; b) Time to complete a simulation run.

Even without statistical hypothesis testing, we can make several inferences from the results shown in Figures 9.7a and b. As we might expect, the 'control' treatment of Trial-and-error R&D had the lowest innovation rate at the end of each run and the slowest time to complete a run (i.e. the BPF reaches the top row in the lattice). Notice that the distribution of Time to Complete for Trial-and-error is not symmetrical; instead it is skewed with some runs taking much longer than average. In comparison, the distributions for the other treatments are much less skewed and more symmetrical. From this we can infer that blind Trial-and-error R&D is more prone to getting 'stuck' on difficult landscapes, compared to the other treatments that have the benefit of knowledge artefacts to improve their success rate.

The second result is that the Non-quant treatment is noticeably better than Trial-and-error, both in Innovation Rate and Time to Complete, but not by a large margin. Thus, even though the Fidelity = 1.0 (i.e. the knowledge artefact was an ideal match to Nature), the relatively low Codification and Abstraction characteristics provide only modest improvement in innovation rates.

The third result is that the relative success of the Quant approach depends critically on the learning rate. With ‘slow learning’ (initial Fidelity = 0.2, Learning Rate = 0.2), the Quant innovation results are only slightly better than the Non-quant results, both in final Innovation Rate and in Time to Complete. However, with ‘fast learning’ (initial Fidelity = 0, Learning Rate = 1.0) the Quant treatment wins the race by a wide margin, both in final Innovation Rate and in Time to Complete.

9.5 Discussion

If we frame the contest between the Quant and Non-quant schools as an innovation race (perhaps on a slippery surface) then we might draw an analogy to the parable of the Tortoise and the Hare. The Non-quants are adopting a Tortoise strategy toward knowledge – slow and steady – mostly because they believe that no more aggressive strategy is feasible given the state of Nature. The Quants are adopting a Hare strategy toward knowledge – start very slow and then rapidly accelerating to the finish – mostly because they believe that this will ultimately achieve cyber security outcomes that are much better than less ambitious methods. Though quite abstract and stylised, controlled experiments with our computational model have allowed us to explore the circumstances where either school will win the innovation race, if any.

The experiment results show that the likelihood of the Quant school winning is critically dependent on the learning rate, i.e. the rate of improvement in how well its knowledge artefacts fit Nature and are therefore effective in facilitating innovation success. If the learning rate is slow, then even if the Quant school achieves slightly higher innovation rates, the Non-quant school might still win the race due to substantial social advantages. It appears that viability Quant school can only be assured if it achieves a high learning rate and becomes demonstrably effective at achieving innovation.

Therefore, it is imperative that institutional entrepreneurs within the Quant school should adopt practices that accelerate learning regarding their knowledge artefacts. While this advice could apply to any professional community, the risky approach of the Quant school means they have more to gain and more to lose, compared to the more conservative approach of the Non-quant school.

Of course, there are limitations to our approach. Our computational model is both abstract and simplified. Therefore it excludes many important factors and dynamics that might ultimately decide who wins, or if there is a winner at all. In a more complete analysis, we would like to assess the scientific merit of each of the schools of thought, i.e. their explanatory coherence (Thagard 1992). We would also want to analyse social dynamics such as legitimization (Nicholls 2010), power struggles (Aronowitz 1988), rivalry over discourse frames (Werner and Cornelissen 2014; Torgersen and Schmidt 2013; Hoffman and Ventresca 1999), and structuration (Giddens 1984). Finally, it would be important to analyse the institutional structure of R&D associated with each school of thought. This holistic analysis would give us a rich picture of the dynamics of institutional innovation in a contested field like quantified cyber security. It would shine light on the challenges and opportunities faced by institutional entrepreneurs who are trying to accelerate innovation in particular directions. As illustrated in this chapter, computationally modelling can complement other methods of analysis and can make unique contributions to research.

Acknowledgements

This chapter is based upon work supported by the National Science Foundation under Grant No. CMMI-1400466. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. Aronowitz, S 1988, *Science As Power: Discourse and Ideology in Modern Society*, University of Minnesota Press, Minneapolis, MN.
2. Battilana, J, Leca, B & Boxenbaum, E 2009, 'How actors change institutions: Towards a theory of institutional entrepreneurship', *The Academy of Management Annals*, vol. 3, no. 1, p. 65–107.
3. Boisot, M 1995, *Information Space: A Framework For Learning in Organizations, Institutions and Culture*, Routledge, London, UK.
4. Boisot, MH 1999, *Knowledge Assets: Securing Competitive Advantage in the Information Economy*, Oxford University Press, Oxford, UK.
5. Borg, S 2009, 'The Economics of Loss', *in* C. W Axelrod, J. L Bayuk & D Schutzer (eds), *Enterprise Information Security and Privacy*, Artech House, Norwood, MA, p. 103–114.
6. Chisholm, D 1995, 'Problem solving and institutional design', *Journal of Public Administration Research and Theory*, vol. 5, no. 4, p. 451–492.
7. Department of Homeland Security 2011, 'Cyber Security Research and Development Broad Agency Annoucement (BAA) 11-02', *Solicitation*, US Department of Homeland Security.
8. Geer, D, Soo Hoo, K & Jaquith, A 2003, 'Information security: why the future belongs to the quants', *IEEE Security Privacy*, vol. 1, no. 4, p. 24 – 32.
9. Giddens, A 1984, *The Constitution of Society: Outline of the Theory of Structuration*, University of California Press, Oakland, CA.
10. Hoffman, AJ & Ventresca, MJ 1999, 'The Institutional Framing of Policy Debates Economics Versus the Environment', *American Behavioral Scientist*, vol. 42, no. 8, p. 1368–1392.
11. Kauffman, S 1996, *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity*, Oxford University Press, Oxford, UK.
12. Kuhn, TS 1970, *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago, IL.
13. Lane, DA & Maxfield, RR 2005, 'Ontological uncertainty and innovation', *Journal of Evolutionary Economics*, vol. 15, no. 1, p. 3–50.
14. Langner, R & Pederson, P 2013, 'Bound to Fail: Why Cyber Security Risk Cannot Be "Managed" Away', *Paper*, The Brookings Institution, Washington, D.C.
15. Leca, B, Battilana, J, Boxenbaum, E & School, HB 2008, 'Agency and institutions: A review of institutional entrepreneurship', *Working Paper Volume 8, Issue 96*, Harvard Business School, Cambridge, MA.
16. Marron, D 2007, "Lending by numbers': credit scoring and the constitution of risk within American consumer credit', *Economy and Society*, vol. 36, no. 1, p. 103–133.
17. National Research Council 2011, *Understanding and Managing Risk in Security Systems for the DOE Nuclear Weapons Complex*, National Academies Press, Washington, D.C.
18. National Science and Technology Council 2011, 'Trustworthy Cyberspace: Strategic Plan for The Federal Cybersecurity Research and Development Program', *Official policy*, United States Government, Executive Office of the President National Science and Technology Council.
19. Nicholls, A 2010, 'The Legitimacy of Social Entrepreneurship: Reflexive Isomorphism in a Pre-Paradigmatic Field', *Entrepreneurship Theory and Practice*, vol. 34, no. 4, p. 611–633.
20. North, DC 1990, *Institutions, Institutional Change and Economic Performance*, Cambridge University Press, Cambridge, UK.
21. Ryan, A, Trumbull, G & Tufano, P 2011, 'A Brief Postwar History of U.S. Consumer Finance', *Business History Review*, vol. 85, no. 03, p. 461–498.
22. Scott, WR 2007, *Institutions and Organizations: Ideas and Interests*, 3rd edn, Sage Publications, Inc., Thousand Oaks, CA.

23. Silverberg, G & Verspagen, B 2005, 'A percolation model of innovation in complex technology spaces', *Journal of Economic Dynamics and Control*, vol. 29, no. 1–2, p. 225–244.
24. Thagard, P 1992, *Conceptual Revolutions*, Princeton University Press, Princeton, NJ.
25. Torgersen, H & Schmidt, M 2013, 'Frames and comparators: How might a debate on synthetic biology evolve?', *Futures*, vol. 48, p. 44–54.
26. Verendel, V 2009, 'Quantified security is a weak hypothesis: a critical survey of results and assumptions', *Proceedings of the 2009 New Security Paradigms Workshop*, NSPW '09, ACM, New York, NY, p. 37 – 50.
27. Weik, E 2011, 'Institutional Entrepreneurship and Agency', *Journal for the Theory of Social Behaviour*, vol. 41, no. 4, p. 466–481.
28. Werner, MD & Cornelissen, JP 2014, 'Framing the change: Switching and blending frames and their role in instigating institutional change', *Organization Studies*, vol. 35, no. 10, p. 1449–1472.