Table of Contents

Go Back

SOFTWARE TECHNOLOGY TRANSITION MODEL DYNAMICAL SYSTEM MODEL

Michael S. Saboe

US Army Next Gen Software Technology Area
Attn: Associate Director AMSTA_TR_R/265
Warren MI, 48397-5000
SaboeM@TACOM.Army.mil

Abstract

This paper considers an information theoretic and dynamical systems technology transfer model. The model enables decision-makers to "engineer" resource and risk programmatic issues. Analysis with the model enables prediction and prescriptive action for a research or program manager. The model deals with entropy as defined in information theory. The TechTx Basic Entropy model addresses macro level trends of a technology at the community level. The dynamical systems TechTx Entropy Feedback model, is based on non-linear control theory. The paper develops the state quantities to develop state functions for analysis of an evolutionary technology process. This summary paper focuses on the elements required to model the technology transfer process. Specifically, this paper develops the fundamentals for a rigorous software technology transfer model as required by the TechTx Entropy Feedback model. The relationship of entropy (S_H) as defined for information by Shannon, and the eigenvalue (l), the norm of a dynamical system is explored. The Lyapunov number is a natural measure developed from the eigenvalue of a dynamical system, e.g. related to entropy. The significance of the eigenvalue for a communications software technology transfer model is discussed. The model shows that the TechTx Basic Entropy and TechTx Entropy Feedback models converge at the same rate. Empirical data with tens of thousands of data points provides a tight confidence interval and demonstrate that the messages conform to a Boltzmann distribution. The paper describes a method to measure the temperature of an evolutionary information theoretic process. This temperature is measured in degrees (^{o}Saboe ã) based in information units. The result is a rich set of analytical tools previously unavailable to be used for policy and programmatic issues to reduce risk, increase efficiency and accelerate maturation of a desired technology.

1 INTRODUCTION

This paper discusses the elements of a new technology transfer model that can be represented mathematically. This provides a method for analysis for both predictive and prescriptive activities. All of the existing work in software technology transfer is lacking mathematical models. The models addressed are 1) *TechTx Basic Entropy*, and 2) *TechTx Entropy Feedback*. Each of these models represents a macroscopic and a discrete representation of an evolutionary communicating system. The mathematical implications of the dynamical system model are developed. Both models represent an extension to the state-of-the-art for predicting the arrival of technologies. The models are validated with over 100,000 data points, collected in one instance over 21 years.

The most significant contribution is the development of a set of state equations, similar to those used in statistical mechanics. A temperature is defined for an evolutionary communications process. Other intensive and extensive variables are defined. A method is shown to determine a conserved quantity (information terms in message), entropy, pressure, and volume.

Finally, a method to represent an engine is provided to show the application of the approach as an engineering model.

The key underlying communication diffusion research of Rogers (Rogers 1983, 1995) is pervasive in Buxton 1991, Raghavan 1988, 1989, Fichman 1993, 1994, Jaakkola 1995, Fowler 1994, Pfleeger 1999, and many more. The paper suggests preliminary analysis of the basic elemental tools required for a software technology transition cycle analysis approach. This work is motivated (Section 2) by the need for an acquirer, or research program manager, to assess risk related to the maturity date of a technology. Data and charts that summarize relevant aspects of this work are presented. A sample data set for "software engineering" is illustrated using current methods. Section 3 addresses the implementation details required to address a software technology model.

This paper presents a general set of state variable for software technology transition. These models provide a method to analyze and later prescribe the size and probability of a technology maturing at a given time. The elements of the analytical model proposed promise to permit analysis of various alternatives for policy and investment trades. Tools that build on this analysis approach can help identify leverage points and opportunities to accelerate progress in a repeatable and rigorous process enabling quantification of maturity at a given date and confidence in a subject technologies stability.

2 MOTIVATION

At the *International Conference on Software Engineering 2000,* the keynote speech (Shaw 2001) illustrated the trends in maturation of software technology. The model cited was one from 1985 (Redwine 1985). It seems apparent that a good model for technology maturation and transition is lacking for software engineering. There are no references in the software technology transition literature that indicates that this was a good model for analyzing, predicting and prescribing exits. There is a clear need based on the researcher's extensive personal experience (nearly 30 years at every level of industry and the Department of Defense). Discussions with the software technology transition program at the Software Engineering Institute (SEI) consistently indicated that there is a critical lack of and need for an analytical model of the type proposed. The elements of an analytical model proposed promise to permit analysis of various alternatives for policy and investment trades. Tools that build on this analysis approach can help identify leverage points and opportunities to accelerate progress in a repeatable and rigorous process.

With such tools, a decision-maker can determine the confidence with which a technology or group of technologies will stabilize and converge in a given time frame. For example (see Figure 2-1), a program might expect a portfolio of technologies to arrive by year 06 with an 80% certainty, but the model might show that in 06, there is only 60% certainty of being available using the current trends. The desired 80% certainty would not be available until 08. If the technology is not predicted to arrive as required, the model will point to the areas for remedy with a prescriptive solution to *organize, train and equip* an organization in order to change the confidence of arrival of the technology for the program's required schedule.

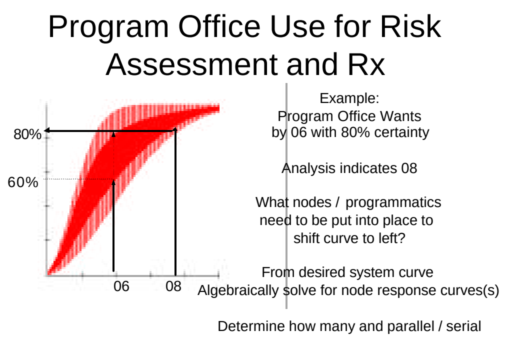


Figure 2-1 Program Office Use of Objective Model

3 TECHNOLOGY TRANSITION MODELS

The research validation follows the strategy shown in Figure 3-1. The proposed *TechTx Basic Entropy* model "X" asks the question, "Can we do better?" in assessing the maturity of a technology, "Y". Validation compares it to the existing methods. The *TechTx Entropy Learning Curve* model, while not discussed in this paper, builds off the *TechTx Basic Entropy* model. In this model, the technology transfer maturation process is characterized by learning curves. The *TechTx Entropy Feedback* model is the most difficult challenge and the subject of the paper for the Systems Dynamical Society. Here the research is asking, "Can it be done at all?" The *TechTx Entropy Feedback* model was developed. The model is exercised with data from the *TechTx Basic Entropy* model. The validation is of the form, matching the model's stabilization equation form of the information theoretic and the dynamical systems model. The data is real world data, and the research suggests that the models are representative of the behavior seen in the real world for this class of processes.

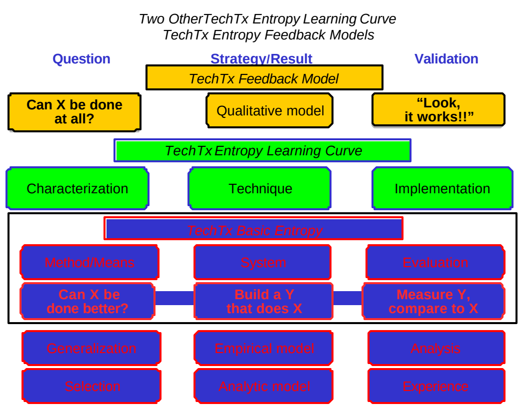


Figure 3-1 Validation Strategy (Source: After Shaw 2001)

The proposed model was compared with the traditional diffusion of innovations communication model to predict trends and the maturation of a technology. The traditional model is the baseline model and uses the *message-counting* method. The first proposed model is the *TechTx Basic Entropy* model. This is the first improvement over the traditional model and uses the content of the message, measured in the information dimension of entropy. Entropy is represented in information units - bits. Section 4 includes a brief review of entropy in information theory will be provided.

The dynamical systems model is the *TechTx Entropy*

Model In Tech Tx Literature	Model Feature	Proposed Information/Control Theory Model Contribution
Theory of Human Needs Model	Complexity factor framework facts, perceptions, actions	Learning Curve Actions on messages (tasks)
Structure Changes Model–	Internal and External Relationship	Shannon Entropy of Messages Joint entropy Information In, Information Out
Technology Model	Goodness of Technology Alone causes Diffusion	Identifies Minimum number of nodes (senders and receptors) extensions may address vacuum and pressure
Institution Building Model	External Influences affect the human behavior to assimilate a technology	Identifies Entropy as a factor that can influence the acceptance of a technology
Equilibrium vs Conflict Model	Equilibrium is an Instrument for Balance Conflict Is a Instrument to apply Pressure	Convergence in Entropy yields balance Large Differences in Entropy yields Pressure
Communication Model	A Technology is Delivered to Adopters Through a Channel, If Understood It is acted upon.	Quantifies the Information Being Acted on, and Quantifies notion of "Understood " in terms of Entropy and Learning Curve
Problem solving Model	Present Hypothesis Test Hypothesis with Data and Logic	Hypothesizes a Mathematical Model and Explains based on Data Analysis

Table 1 Technology Transfer Models, Features, and Relation to Proposed Model

Feedback model. The feedback model is designed to operate at the organizational node level. This model can accommodate a producer-advocate creation of new information. The information is provided in the form of a work product as a "message" consisting of terms. The model includes a feedback request transition and a clarification transition of the messages. Learning is factored into the amount of feedback requested and permits tuning of the dynamical system model. The characteristics of the *TechTx Entropy Feedback* model are discussed. Introduction of the model lays the groundwork for future research.

1. Context and Overview

Let's set a context. Induction[1] is a process of inferring a general law or principle from the observations of particular instances. This is inductive inference. Inductive reasoning is a more general concept than inductive inference. It is a process of assigning a probability (or credibility) to a law or proposition from observation of particular instances. Inductive inference draws conclusions on rejecting or accepting a proposition, possibly without total justification. Inductive reasoning only changes the degree of our belief in proposition. Deductive reasoning of inference derives the absolute truth or false hood of a proposition. This is a case of inductive reasoning.

This approach to explaining things around us dates back at least to Epicurus (342?-270?BC) (Li 1993, p. 274). Let's consider theory formulation in science as the process of obtaining a compact description of past observations together with future ones. Let us suggest that the: preliminary data of an investigator, the hypothesis proposed, the experimental design and setups, the trials performed, the outcomes obtained, the new hypothesis formulated, etc., can be encoded as an initial segment of an infinite binary sequence. The investigator obtains increasingly longer initial segments of an infinite binary sequence by performing more and more experiments. To describe the underlying regularity in the sequence, the investigator tries to formulate a theory that governs the sequence based on the outcome of past experiments. Candidate theories or hypotheses are identified from the sequences starting with the observation of the initial segment.

There are many different possible infinite sequences or histories on which the investigator can embark. The phenomenon the investigator is trying to understand or the strategy used can be stochastic. In this type of view, a phenomenon can be identified with a measure, i.e. probability distribution, on a continuous sample space.

This research attempts to express the task of learning a certain concept as in terms of sequences over a basic alphabet. We express what we know as a finite sequence over the alphabet. An experiment to acquire more knowledge is encoded as a sequence over the alphabet, the outcome is encoded over the alphabet, and new experiments are encoded over the alphabet, and so on. This way we can view a concept as a probability distribution (measure) over a sample space of all one way infinite binary sequences. Each sequence corresponds to one never ending sequential history of conjectures, refutations, and confirmations. The distribution can be said to be the concept of phenomenon involved. We can predict what is likely to turn up next with an initial segment. Using Bayesian analysis (Bayes 1763) to compute the conditional probability, we can predict and extrapolate future outcomes. This is the general thrust of this research.

Let's develop an analogy of the flow of communication to a physical model to illustrate the concept. When two people meet, they converse, and consequently modify their thinking to some extent. These modifications are brought to subsequent meetings and modified further. The word for this is dissemination or diffusion. There is a flow of communication in society, just as there is a flow of correlations in matter. Let's explore this idea of correlations using the analogy of a physical system and look at what happens in terms of distribution functions.

Consider a glass of water[2]. We may visualize the interactions as leading to collisions between the molecules. We can describe the water containing them in terms of a statistical ensemble. The water is not aging if we were to consider the individual molecules over geologic time[3]. Yet, there is a natural time order in the system from a statistical point of view. Aging is a property of populations, as in the biological theory of evolution as developed by Darwin. It is a statistical distribution that approaches the equilibriumdistribution.

Consider a probability distribution $p(x_1, x_2)$ dependent on the two variables x_1, x_2. If x_1 and x_2 are independent, we can factor $p(x_1, x_2)= p_1(x_1)\ p_2(x_2)$. The probability $p(x_1, x_2)$ is the product of the two probabilities. On the other hand, if $p(x_1, x_2)$ cannot be factored, x_1 and x_2 are correlated (Bayes 1763 p299). Return to the glass of water molecules. The collisions between the molecules have two effects: they make the velocity distribution more symmetrical, and they

[1] The Oxford Dictionary defines induction this way.

[2] This discussion follows from Prigogine 1997.

[3] Newton scholium differentiates time this way. "Time, space, place, and motion … quantities are popularly conceived solely with reference to the objects of sense perception. … 1. Absolute, true, mathematical time, in and of itself and of its own nature, with out reference to anything external flows uniformly and by another name it is called duration. Relative, apparent, and common time is any sensible and external measure (precise or imprecise) of duration by means of motion; such a measure – for example, a month, a year – is commonly used instead of true time." (Newton 1726 p408)

produce correlations (see Figure 3-2). However, two correlated particles will eventually collide with a third one (see Figure 3-3). Binary correlations are then transformed into tertiary ones, etc. Prigogine illustrated this molecular model, and it has been verified (Prigogine 1997 p79).

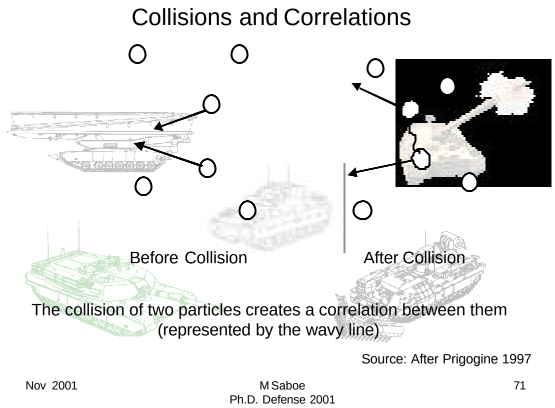


Figure 3-2 Collisions and Correlations (Source: After Prigogine 1997)

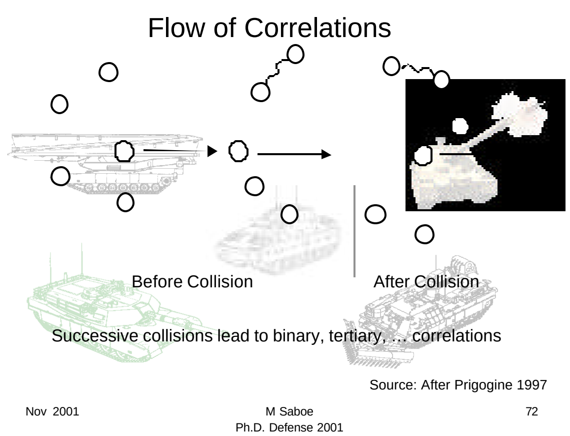


Figure 3-3 Flow of Correlations

We could conceive of inverse processes that make the velocity distribution less symmetrical by destroying correlations. Processes that invert the velocity of particles for a physical world as in Figure 3-3 have been reproduced. However, this inverted flow of correlations can only be achieved for a short time, with limited numbers of particles. Then we again have a directed flow of correlations involving an ever-increasing number of particles leading the system to equilibrium

We now have a flow of correlations that are ordered in time just as there is a flow of communication in society. There is a method to describe this irreversibility. This statistical description is dynamics of correlations leading to the equilibrium solution.

In this research, we use messages instead of particles. This turns out to be a conserved quantity (conserved quantities shared between two systems need not be restricted to energy[4], or mass, or volume, the conserved quantity could be a number of measures, even money) (Yakavenko 2000) (Farmer 1999). We are concerned with a deterministic dynamical system as an especially simple type of dynamical system, corresponding to dynamical system maps. Contrary to what occurs in ordinary dynamics, time in maps acts only at discrete intervals. Maps represent a simplified form of dynamics that make it easy to compare the individual level of descriptions (trajectories) with the statistical description.

2. Communication, Continuity

Communication is a process in which participants create and share information with one another in order to reach a mutual understanding. This definition implies that communication is a process of convergence (or divergence) as two or more individuals exchange information in order to move toward each other (or apart) in the meanings they ascribe to entities (objects, acts, events, etc) (Rogers 1983). Rogers and Kinkaid represent this communication in the general case as a two-way process of convergence rather than a one-way linear act in which one individual seeks to transfer a message to another. (Rogers Kinkaid 1981).

This simple concept of human (or machine) communication seems to accurately describe certain communication acts or events involved in technology diffusion.

3. Diffusion

Diffusion is the process by which an innovation is communicated through certain channels over time among the members of a social system. It is

[4] Energy is an interesting term. It is a primitive term. It is a mathematical abstraction that has no existence apart from its functional relationship to variables or coordinates that do have a physical interpretation and that can be measured (Abbott 1989 p1). The 1st law of thermodynamics is merely a formal statement asserting that energy is conserved. This represents a primitive statement about a primitive concept. Moreover, both are linked. The 1st law depends on the concept of energy, and it is equally true that energy is the *essential* thermodynamic function precisely because it allows the formulation of the 1st law.

a special type of communication, in that the *messages* are concerned with new ideas (Rogers 1983). For example, when a change agent seeks to persuade a client to adopt an innovation. Examining what occurs in the time step prior to an event and after an event, it is clear the event is only a part of a process of exchange between individuals (or machines). Rogers asserts that it is the newness of the message content of the communication that gives diffusion a special character. The newness implies that some degree of uncertainty is involved.

4. Uncertainty and Confidence

Let's set the context. How do we make choices in the face of uncertainty? We know that a reasonable person having some historical experience with a true coin A, would assign a degree of belief (subjective probability) of about .5 probability for heads. Based on the history with the coin, we would be rather *confident in that belief.* Now imagine a coin B, and we know absolutely nothing about this coin. We don't know whether it has two heads or two tails or if it is a fair coin. Yet, if we had to pick, we would be compelled to assign a single .5 probability, since we lack any information to indicate a greater or lesser belief in heads vs. tails. But, our *confidence* in .5 for coin B would surely be less.

On the one hand, it is not the psychological sensation of confidence that we are interested in. Rather, as an engineer or decision-maker, the consequences of the decisions are the driving issue. When we have the option of acquiring information through an *informational* action, we are likely to invest energy (money, effort) before making a decision that results in a terminal action. We would be willing to invest this additional effort in acquiring information about coin B vs. A. So we see that one's *informational* actions, though not one's *terminal* actions, do depend on one's *confidence* in beliefs.

This notion of confidence plays an important role in this discourse's assessment of a software technology.

5. Chance, Aggregation through Mixing

Today we tend to regard knowledge as a process more than a state. This stems partly from the epistemologies of the philosophies of science: Cournot's probablism and his comparative studies of various types of notions set the stage for such an understanding. Critical reviews of historical works, which reveal the oppositions among the various types of scientific thought, clearly promote such a development. Even after the victory of Newton, physics believed for hundreds of years in the absolute character of its principles. So, the arguments developed in this research very much depend on the state and maturity of the knowledge process for software engineering.

Another probabilistic feature of software technology transition is *chance*. *Chance* is a curious notion which is defined by Cournot as an interference of independent causal series and which generally can be designated under the term *"mixture"*. (Piaget 1977, p. 19) This is an important concept to expose. *Mixture is irreversible and grows with an increasingly weaker probability of return to the initial state*. This starts to address the aggregation typical of composition of terms and integrating domains and technologies.

Topical Threads of Research to Date

The following sections review key efforts and models that have been identified in the literature. This review is meant to illustrate the state of the practice for technology transfer models. Further, it helps establish desirable aspects that should be addressed by a model. Technology transfer (TechTx) or transition is referred to as *diffusion* in the literature. This section reviews the basics of technology transition through the current state of software technology transfer as seen in the literature through 2000. Various theories and principles felt to be underlying human behavior and learning are presented first. The technology transition models basics identified in the literature are then summarized. Seven models researched identify a facet or feature of technology transfer. These models are shown in Table 1. Table 1 shows the model, a key feature of the model, and indication that the model proposed in this research address that feature. Each of these models in Table 1 are summarized in the next few paragraphs.

1. The Theory of Human Needs (Leagans 1979)

The theory of human needs (Leagans 1979, p. 15) has a number of components. These are as follows: the facts, the perception of the facts, human attitudes or value judgements about the facts, and human actions related to the facts as they perceive them. Leagans establishes a framework addressing the complexity factors that affect behavior with respect to technology transfer. The model elements suggested in this paper had to be general enough to permit lower level detailed elaboration that address these details. This requirement for generality is driven by the need to address refined implementation aspects of technology transfer. The proposed model addresses this through the mechanism of the learning curve.

2. *Structure Changes – Internal and External Relationship (Piaget)*

While Piaget's work was not focused on technology transfer, his work is fundamental to learning schemes and an accommodation of these schemes to the environmental situation (Piaget 1963, p. 103). He develops the relationship between the genotype (internal) and phenotype (external) information influences. Yet, neither internal nor external factors can individually explain human development of skills. We can think of this learning in terms of the acquisition of technology. During human knowledge and skill development, it seems to tend toward the establishment of equilibrium of the internal and external factors. (Piaget 1967, p. 113) The proposed model explored in this paper addresses this in several ways. First, the Shannon entropy approach, which takes a vocabulary as input and a vocabulary as output, and from the joint entropy (Bayesian) relationships, yields a grammar. In the *TechTx Entropy Feedback* models, the vocabulary-grammar relationship between internal and external factors is incorporated using Shannon's statistical approach to entropy. The *TechTx Entropy Feedback* model adds mixing. It also accommodates structural changes (more explicitly addressing the external factor) due to feedback from external nodes.

3. *Technology Model*

The technology model (Leagans 1979) deals with potential. This model suggests that the attractiveness of a new technology alone is sufficiently strong to induce wide diffusion, acceptance and adoption by users. It tends to assume that users would use the new technology and attendant parts of the technology successfully without the persuasions of an organized education system. This model has proven highly inadequate when trying to introduce technology to large masses of users, rather than the elite self-motivated few (Leagans 1979, p. 17). This inadequacy is also consistent with the small percentage of innovators and early adopters identified by Rogers. (Rogers 1983 p. 247) However, it does imply that a pressure or a vacuum may have some influence. The model currently being explored seems to be able to be extended to see the effects of a vacuum, e.g. the growth of the internet creates a requirement and hence a vacuum, and intelligent agents move in to fill the void. This is analogous to the saying, "necessity is the mother of invention."

4. *Institution-Building Model*

The laws of maximum and minimum are often referred to as the "limitation factors". These factors are used to explain the forces influencing biological, e.g. plant growth. Briefly the laws say, "If one of the participating nutritional constituents of the soil or atmosphere are deficient, or wanting or lacking in assimilability, either the plant does not grow or its organs develop only imperfectly." (AAAS 1972). This has been applied to human behavior with the following rationale (Leagans 1979, p. 13): human behavior is the dependent variable. The assumption is that man can influence the economic, biological, and other forms of change to the extent that he controls the forces (nutrients) that influence change and the status quo. There is the implication here that there is a vector of forces that can be added up. In this context, Leagans argues that people see one or more inhibitors (limiting factors) and one or more incentives to innovation, simultaneously in any situation. These variables contain and exert varying force on the dependent variable - human behavior, and that when the deficiencies (inhibitors) are weakened or removed, the balance or equilibrium of opposing forces will be altered. Changes in human behavior are expected to be proportionate to the amount of cumulative influence exerted by the change incentives present. These changes are the net sum of the counteracting influences or change inhibitors operating in the situation.

The model in this paper uses information theory to quantify the information entropy in terms of mutual information, joint and conditional entropy and to address the relationship of these forces. In the model, the information previously published is persistent and influences the output. The proposed *TechTx Entropy Feedback* model builds on the contribution of feedback. The feedback is seen as proportional to the cumulative influence of the change incentives (information) present. The feedback control model used herein is non-linear. This addresses the comment by Leagans (Leagans 1979, p. 14) that "the input-output function is not always linear." He states that the influencing factors vary by situation. The result is probabilistic. This derives from the fact of variation in the nature of each interaction. For the research herein, we address this by means of an ensemble of very probabilistic, primitive communication interactions using both information and control theory.

5. *Equilibrium vs. Conflict Model*

In the equilibrium vs. conflict model, equilibrium is regarded as an instrument for achieving balance, while conflict is an instrument for applying pressure. Some combination of these divergent approaches does in fact operate in most models as a force for motivating people to adopt new patterns of behavior. This is consistent with Piaget and the tendency toward the establishment of an equilibrium of these factors. In developing the mathematical model of this study, it was interesting to discover that the communication control model used can settle down into equilibrium (oscillating), repellor or attractor stable state. Oscillation is seen under some conditions of the feedback model. When there is a vacuum, or pressure is applied to a node, learning is more rapid, up to a point. This can be seen in the proposed model.

Prigogine (Prigogine 1980, 1984), who won the Nobel Prize in 1977, says that living (read this as evolving) systems are rarely static, and if they are, they are likely to atrophy and die from stagnation. Living organisms do not thrive in a state of balanced equilibrium, but usually in fluctuating restlessness. The data presented exhibits exactly the behavior described by Prigogine. Consumers, organizations, and the technology evolution system itself seem to act as a living organism. The model developed herein addresses these concerns.

6. *Communication Model*

The communication model is considered the classical model for diffusion of technology. It is well developed and documented by Rogers (Rogers 1983, 1995). This consists of making a new technology discovery, delivering it to potential adopters through various communication channels, and then being understood and acted upon by the consumer. The communications model is generally seen as a macro model.

Almost every well-researched technology transfer model addresses the communication model. Leagans (Leagans 1979, p. 19) cites Rogers (Rogers 1975), who identified several shortcomings of the model. These include the need to address greater process orientation, greater attention to causality, and recognition that the adoption requires a physical or overt act. This paper addresses these shortcomings in the formulation of the mathematical model in sections 4 and 5. The process aspect is in the message and feedback loops in the control model. Causality and overt act are built into the transforming function $f(x_k)$ in a time step in Sections 4 and 5.

7. *Problem Solving Model*

This model presents a hypothesis of an explanation of a troubled situation. It tests the hypothesis with data and logic developed putting those specific results into a model. The hypothesis for solving the problem is formulated. Implementing programs and evaluations to assess the consequences tests the proposed solutions. This evaluation/implementation includes the means and the ends. Boehm and Basili (Boehm 1999, 2000) essentially are espousing that the Department of Defense institute a National effort with Centers for Empirically Based Software Engineering (CeBase) to address transition, using essentially this model.

The study develops a model at a macro, or strategic, level to predict and plan the technology portfolio of a National Technology Transition effort. The current model efforts and elements are reflected in the Department of Defense Software Engineering Science and Technology Summit findings (Boehm 2001).

Classic Diffusion Tech Tx Models (Rogers 1983, 1995)

The Diffusion of Innovation (Rogers 1983, 1995) is one of the most valuable readings on technology transition in general. The approaches of virtually all aspects of technology diffusion are covered. Rogers discusses a communication model that depicts the classic business school "***S***" curve (Rogers 1983, p. 47). This is a cumulative plot of publications covering a given topic over time. Further, he categorizes the four main elements of diffusion of innovations as follows:

- The Innovation
- Communication Channels
- Time
- A Social System

He lays out clear definitions that are commonly accepted in the literature of technology transition and diffusion. Rogers' lexicon can also be seen in the software engineering technology transfer literature. (see Moore 1991, Redwine 1984, Fowler 1994, Fichman 1993, Zelkowitz 1995, and Pfleeger 1999).

Looking at Rogers’ work, you can see all of the elements of a communication system. He classifies and distributes the types of adopters (see Figure 3-2) as Innovators, Early Adopters, Early Majority, Late Majority, and Laggards. He stresses the uncertainty-reduction aspect of technology. He, as do many, use the terms “innovation” and “technology” as synonyms.

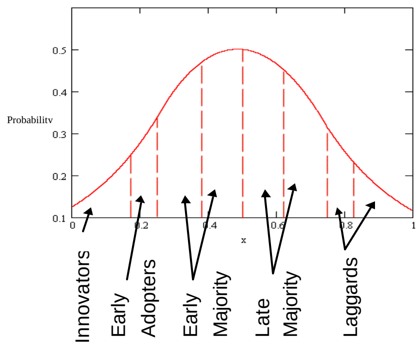


Figure 3-4. Distribution of Adopters.

(Source: Rogers 1983, p. 11).

Rogers identifies *technology* as a design for actions that reduce the uncertainty in the cause and effect relationship involved in achieving a desired outcome. (Rogers 1983, p. 12). The technology developed in the case of this study is itself the technology transfer model. The proposed model in this paper provides a method to analyze options for instrumental actions in order to reduce uncertainty in the arrival of a given set of software technologies.

1. The Innovation

In the literature, technology generally is seen as having two components, hardware and software. Rogers is speaking of hardware and software in the most general sense, not limited to computers. 1) *hardware* consists of the tool that embodies the technology as material or physical objects. 2) *Software* consists of the information base of the tool.

Technological innovation creates one type of uncertainty in the minds of potential adopters (about its expected consequences), as well as representing an opportunity for reduced uncertainty in another sense (that of the information base of the technology itself). The latter is the potential uncertainty reduction representing the possible efficacy of the innovation in solving an adopter's need or perceived problem.

Once information-seeking activities have reduced the uncertainty about the innovation's consequences to a tolerable level, a decision to use innovation will be made. The models in this paper address the innovation-decision process, which is essentially an information seeking, information sending, and information processing process. While this is not visible in the *TechTx Basic Entropy* model, the effects of the learning curve are found in the *TechTx Entropy Learning* model. The *TechTx Entropy Feedback* model factors in the request for clarification and feedback in order to reduce the uncertainty about the advantages and disadvantages of the innovation.

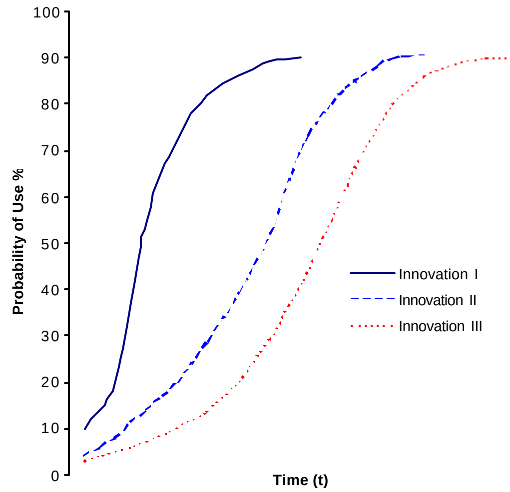


Figure 3-5. Diffusion. (Source: Rogers 1983, p. 11)

2. Communication

The primary model in Rogers 1983 is a communication model. While Rogers lays out the communication channel element as component critical to diffusion, he performs and references an enormous amount of empirical data without addressing the model in terms of a communications system. Applying communication and information theory methods to this observation is indeed an area that could benefit the study of software technology transfer. The benefit of an information theory and communication model approach has not been addressed to date. The model developed in this paper suggests a quantitative method to address the communication model using Shannon's entropy, the bakers' transformation (an entropy) of the control model and learning curves.

3. Time

Time is an important element of the diffusion process. Time does not exist independently of events. It is an aspect of every activity. We think in terms of astronomical time, or time differences similar to asking a person on the street for the time and they look at their watch. Rogers and all of the technology transition literature address this type of time. This is time as described in classical physics. We in western scientific tradition take this for granted since the writings of the philosopher Aristotle, in which *time* is closely related to motion and therefore to space. This is a classical interpretation of time in which the present separates the past from the future.

In the basic work *Process and Reality*, Whitehead emphasizes that the simple location in space-time cannot be sufficient and that the embedding of matter in stream of influence is essential (Prigogine 1983). Whitehead emphasizes that no entities, no states can be defined without activity. No passive matter can lead to a creative universe.

It is only recently that *time* can be expressed in a precise mathematical form. Since we are faced with Planck's Paradox, with the absence of a physical reality, this study moves toward the mathematical notion of time as taken with the use of the bakers' transformation in time steps and presented by Prigogine (Prigogine 1983).

The bakers' transformation is essentially the folding and stretching that results in mixing. To better understand the function, let's examine two examples normally given to describe the process. Imagine Rome, when we observe the city, we see architecture and buildings from many time periods. They are all interspersed and mixed into the city. These areas and remnants, which are interspersed, are the result of mixing at a number of iterations. The other example, and the one where the bakers' transformation gets its name, is folding and stretching of dough horizontally and vertically. Take a piece of dough, and place a spot of sauce on the dough. Fold the dough. Stretch the dough to be the original area again. Then successively repeat the iteration action. We can let X be the function that represents the value corresponding to the application of n bakers' transformations.

$$X_{n+1} = F(X_n) \qquad (3.1)$$

The various functions X_n are functions of internal time. The internal time is an *operator* like the one used in quantum mechanics. The age of partition X_n is the number n of iterations i that are to be performed to go from X_o to X_n. Whenever the internal time exists, it is an *operator*, and not a number. Further discussion can be found in Prigogine (Prigogine 1983), Farmer, York Ott, (Farmer 1983), McCauley, (McCauley 1993), and Baker (Baker 1990). This is the form of the finite difference equations used in the models.

4. *Social Structure*

The social structure provides the network and media to transmit the messages in the communication-diffusion model. Rogers (Rogers 1983 p. 25) quoted Katz, "It is unthinkable to study diffusion without some knowledge of the social structure in which potential adopters are located as it is to study blood circulation without knowledge of the structure of the veins and arteries." The social system is a set of interrelated units that are engaged in joint problem solving to accomplish a common goal (Rogers 1983 p.24). In other words, the model is a kind of graph.

There is more to it than interrelated units when establishing the network of individuals and organizations. Hargadon (Hargadon 1997) provides an interesting insight via an ethnography on these network mechanisms, for technology brokering and innovation in a development firm that produces one of a kind products. He identifies the mixing mechanisms and the feedback process, building on historical data and experience. The experience is held in informal networks and is communicated in terms that are aggregations and abstractions of terms that were used in prior internal efforts. Typical of the communication were short hand descriptions that would sound like, "We can build this with a *X* like a *Y* from the *Z* project." In this dialog, *Y* is an abstract chunk of a previous project. Correlations are established in the participant's mind's eye.

Allen (Allen 1977, 1983) emphasizes the importance of the "messages" from outside organizations. He indicated that as many as 80 percent of the messages come from sources outside the organization. This is interesting since the model proposed will draw on external sources of information providing "messages".

There is a method to determine effective – efficient network size and diversity, referred to as optimizing *structural holes* of *social capital* (Burt 1992). Essentially social capital is found in relationships – whom you know. It is managed, and it aggregates from people to organizations and can be orchestrated to build an effective social structure and network. The model proposed in this paper addresses the node linkages of authors and corporate sources by using the joint entropy of Shannon. While the models herein do not develop these details, the models have been developed to accommodate a structural hole analysis. The approach chosen enables later refinements as detailed node relationships are developed for lower level models, e.g. references cited or actual studies of message traffic of a receiver node.

In competitiveness, or survival, social capital is organized naturally around the human behavior and the principle of least effort. In simple terms, this principle of least effort says that a person solving the immediate problems will be viewed against the background of the person's future problems, as estimated by the person. Moreover, the person will strive to solve the person's problems in such a way as to minimize the *total work* that must be expended in solving *both* the person's immediate problems *and* the person's future problems. That in turn means that the person will strive to minimize the *probable average rate of his work-expenditure* (over time). In addition, in so doing he will be minimizing his *effort* (Zipf 1965, p. 1).

In the area of software engineering, Boehm (Boehm 1989) developed a *Theory W* to help individuals and organizations to negotiate win-win conditions, given constraints and alternatives. *Theory W* is a management theory and approach which says that making winners of the key stakeholders is a necessary and sufficient condition for an effort's success. (Boehm 1998) First-hand experience by the Army (Saboe 2001a) over the last 10 years with the WinWin process model and tool, indicates that *Theory W* does provide a method for a group of individuals (and by extension this could be seen as representative of organizations) to analyze and act over a larger visible decision space when acquiring a software engineering process technology. This does enable the principle of least effort to be used in a group setting in a quantitative fashion.

The current research addresses minimum effort through the study of joint entropies in the model. Minimizing the rate of change of entropy, i.e. watching a technology mature, is something that can be observed in the model. On the prescriptive side, actions can be taken to get the technology to stabilize quicker, by investing in refinements, redundancy of the message set, propagation of the messages, and analyzing the effect on the entropy, and hence the principle of least effort.

With the foregoing, we are armed with the basics that influence technology transfer.

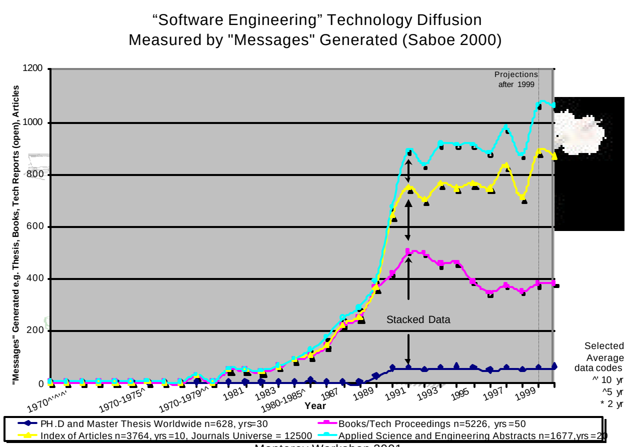


Figure 3-6 **"Software Engineering" Messages Initial** Data.

(Source: Saboe 2000, 2001)

The first experiment, which we refer to as experiment 0, starts to quantify this for software engineering, and is seen in **Figure 3-6**. **Figure 3-6** illustrates the message-counting approach of Rogers for the technology. We have the number of messages published in a given year on the *Y*-axis, and time in years on the *X-axis*.

.

5. Experiment 0 "Count Every Message - Everywhere"

The initial study, called experiment 0, evaluated the technology "Software Engineering" to determine if indeed there was a better way to get a handle on maturation of technology. During this experiment, the effort looked at all print messages available. Software engineering "messages" were counted starting in 1968. The leading indicator messages appeared to be graduate programs that performed research and published messages in the form of Master's theses, and Ph.D. dissertations. Searching Dissertation Abstracts, 628 of these messages were found over a 30-year period. Messages showed up on software engineering technology in the form of books and technical proceedings. 5226 of these book/technical proceeding messages were found from a source going back 50 years. Messages in the form of articles in abstracted journals had a yield of 3764 messages, over a 10-year period, from a journal universe of 12,500 journal titles. Messages similar to these were searched in another source, the Applied Science and Engineering Abstracts. The result was 1677 messages over a 20-year period. This yielded the data shown in **Figure 3-6**. The data for this chart is found in the appendix. This is a typical message-counting approach. Even when the data is not cumulative, we can see that there are general trends.

We can make a few qualitative observations from the message-count data for software engineering. Looking at the messages published *each year* in **Figure 3-6**, we get a sense of capacity. The research messages from the research institutions seem to be one of the limiting factors. Books and technical proceedings top out as well, also giving an indication of steady state capacity. Articles seem to be still growing. Articles are shorter and therefore more of a gloss than the high-end messages in the form of a book or a technical report, thesis or dissertation. These high-end messages are where one would expect the new ideas to come from. It is easy to see that the capacity to produce this type of high-end messages has stabilized. The academic research infrastructure is only capable of producing on the order of 100 "new idea" messages per year. Producers of books and technical reports add another 300 messages per year at capacity. While researchers producing high-end messages containing new information are not the only source of new information, we see they have a capacity limit in the number of messages produced. In order to build a nationally competitive infrastructure, these are the types of leverage points to which research managers and government policy makers need to have access.

While this is interesting, the message-counting approach is limited in its analytical value. It is a very labor-intensive effort with minimum quantitative yield that would enable better-informed decisions for proactive actions. The idea to find a representative sample of messages for the technology under examination pointed to professional societies. While their databases would not cover every message, they would yield a rich enough source to potentially bear meaningful fruit.

6. Crossing the Chasm

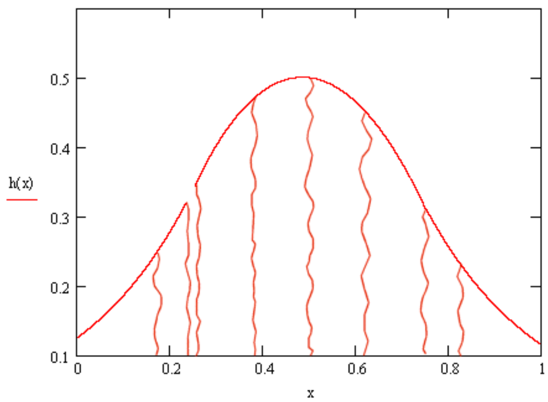


Figure 3-5. Chasm and Fissures in Technology Transition.

(Source: After Moore 1991)

Moore (Moore 1991) identified a chasm between the early adopters and early majority. Fissures were identified between the other adopter segments of the communiy. At least two factors contribute. First, the communication channel between the segments of the community may be non-existent or spotty. Second, if the communicaiton channel existed and was established, there is an impedence mismatch between advocates and receptors in different communitiy segments.

7. States of Software Technology

Redwine et. al. (Redwine 1984) studied 14 different cases in considerable detail. They identified five major phases, and two sub phases, 4a and 4b, that a technology passes through as it matures. Figure 3-7 shows the states. While the analysis is extremely good for the cases studied, there is a bit of imprecision in states 4a and 4b, e.g. popularization throughout 40% and 70% of the community respectively. It is extremely difficult to determine based on their methods, or any other reasonable approach, how to determine what the total community is.

During the validation of this research, data has been gathered on five of the fourteen technologies in addition to more current technologies. This data is found in other research (Saboe 2002).

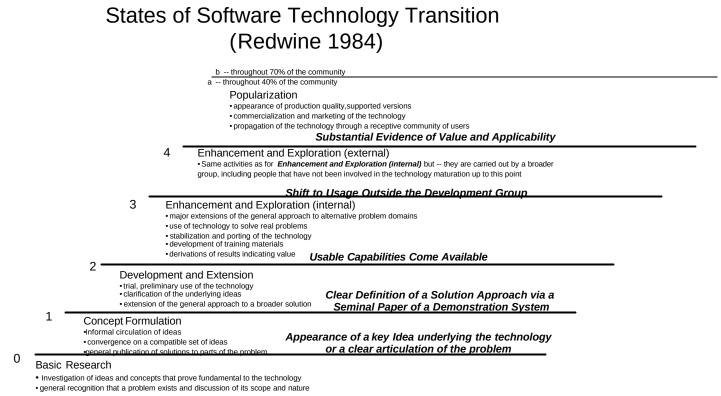


Figure 3-7. States of Software Technology Transition. (Source: Saboe 2001, Redwine 1984)

Software Technology Transition Framework, Advocate/Receptor

The Software Engineering Institute has been the single most prolific source on the subject of software engineering technology transfer. This is readily understood since this Federally Funded Research and Development Center was established with a primary mission to establish transfer of software engineering technology to the Department of Defense. Fowler (Fowler 1994) developed a framework for technology transfer identifying advocates and receptors (change agents) mediating between producers and consumers (see Figure 3-8). In this work, three life cycles of technology transition are presented: research and development, new product development, and implementation. Emphasis on the need for common terms between receptors, consumers, and researchers is identified as an important aspect of the SEI studies. This paper's model accounts for this finding. A clear signal, with minimum noise and need for requests for feedback, between a sender and receiver improves technology transfer.

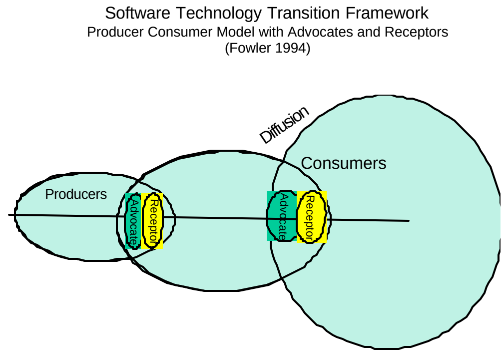


Figure 3-8. Software Technology Transition Framework.

(Source: Fowler 1991)

This paper does not address the lower level implementation details of that framework; rather it builds an analytical framework useful to determine probability of success and quantity and redundancy of messages that need to be sent as a clear signal.

Significant additional work (Forrester 2000, Fowler 1992, Fowler 1992a, Fowler 1990) has been developed at the SEI. This work primarily focuses on the lower level implementation details of the framework, e.g. methods on how to plan and effectively communicate technology to an organization.

Saboe (Saboe 2001) has mapped the framework of Forrester into the phases and state transition points of Redwine (See Figure 3-9).

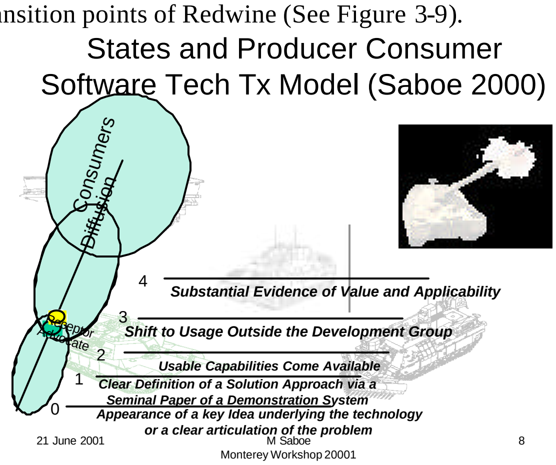


Figure 3-9. Mapping of the SEI Transition Framework and Redwine's Stages.

(Source: Saboe 2002)

1. Extension to Address Standardization Effects (Fichman 1993)

Fichman and Kemerer (Fichman 1993) focused on organizational and community-wide technology adoption. They develop a two dimensional framework based on theories relating to organization and communities. They particularly bring the economics of standardization to the literature for the first time in the software engineering process technology literature. This work points out the economic factors affecting technology adoption. These are summarized as follows:

Prior Technology Drag

A prior technology provides significant benefits because there is a large and mature installed base. The research model of this paper addresses "pushback" by measures of the entropy. e.g. the terms of the technology show up more and more in the community lexicon. The more familiar, the less likely the technology will be pushed back and the less requests for clarification will be required. In the *TechTx Basic Entropy*, the measure of entropy as input gives synthetic metric for the technology drag.

Irreversibility of Investments

Adoption of the technology requires irreversible investments in areas such as products, training, and accumulated project experience. For example, once the money is spent on a technology, it is gone. It cannot be spent again. Another example is closer to the thermodynamic aspect of irreversibility. Once the community or a node in the community is exposed to a technology, you can not unexpose them. The future is influenced by that exposure to a product, training and prior experience. This paper addresses this resisting organizational force of prior experience, training and exposure through the entropy aspect of the model. In the control theory part of the model, the requests for feedback become less if the input messages represent well-understood messages by the resources and assets in the node.

Sponsorship

Strong sponsorship seems be beneficial in moving a technology to standardization when a single entity (person, organization, consortium) exists to define the technology, set standards, subsidize early adopters, and otherwise promote adoption of the new technology. This paper suggests that setting standards reduces the noise in the producer-(advocate–receptor)-consumer lexicon, thereby reducing the rate of change of the entropy. In addition, large quantities with a limited amount of new terms introduced published each year, would reflect sponsorship. Even if there were not a single entity with resources focused to promote the technology the models would suggest that the technology is approaching stability, and converging. While the model does not address resources explicitly, the result of concentrated, with the same vocabulary, messages with new information either reduces entropy, moving the vocabulary toward stability, or retards the movement toward stability and convergence. The change in entropy, as the result of new messages in the result of effort, implies resource consumption to produce the messages. The stability and convergence (i.e. decrease in the rate of change of entropy) suggest the lexicon is becoming standardized. This may be defacto. The vocabulary, communication network approach and the change agent (sender – receiver) aspect of the model address this factor which was seen as desirable and identified by Fichman and Kemerer (Fichman 1993).

Expectations

Technology benefits from an extended period of widespread expectations that it will be pervasively adopted in the future. This paper sets up the ability to further analyze the notion of expectations. However, this is the topic for further research as identified in the final sections. Work addressing mathematical concepts of momentum and potential can be developed based on the elements of the initial model.

The work by Fichman and Kemerer also identifies attributes of innovations. Although Rogers addressed and identified five generic attributes of innovations (1) relative advantage, (2) compatibility, (3) complexity, (4) trialability, and (5) observability, his work is based mostly on study of individuals. Van de Ven (Van de Ven 1991) argues that these same innovation attributes play an important role in adoptions by organizations. The Rogers' attributes have been generally adopted by the community due to the familiarity to the diffusion of innovations community. Others (Moore 1987), (Kwon 1987) use these as well. Alternate taxonomies show up in Leonard—Barton (Leonard-Barton 1988), who identify transferability, organizational complexity, and divisibility; Pennings (Pennings 1987) identifies concreteness, divisibility and cost; and Eveland and Toratzky (Eveland 1990) identify trialability, lumpiness, adaptability, degree of packaging, and the "hardness" of the underlying science. Zelkowitz (Zelkowitz 1998) relates different styles to Rogers' attributes and characteristics of the adopter type. In most cases, all of these can be mapped back to Rogers' original attributes.

This research was constructed to address Rogers' compatibility, complexity, and trialability in terms of the entropy metric. Relative advantage is addressed only indirectly, but the mechanism is there to compare two or more competing technology entropy metric curves and to determine the rate of change, crossover, and probability of arrival of a technology's maturity. Studies that spot Redwine's observable (first four) state transition points in data are made for five of the fourteen technologies they studied. It is premature to say that we can make any predictions by spotting observable points alone. However, future research could spot the observable events and attempt to correlate probability of success with the entropy metric. .

2. Diffusion/Infusion Issues (Zelkowitz 1995)

Zelkowitz (Zelkowitz 1995, 1998) has extensive experience with infusing technology into organizations. Infusing is differentiated from diffusion as it relates to internal adoption by a particular target organization, while diffusion generally refers to movement of the technology to the broader user community in a macro sense. His study within NASA builds on the "experience factory" work with NASA's Software Engineering Lab and the experimental approaches of Basili (Basili 1994, 1994a). He studies the differences in the industry-wide phenomenon of a technology specifically focusing on the infusion process which actually make the changes in the current state of technology. This paper addresses infusion process in $f(x)$ and the interactions as successfully retransmitted messages from a change agent (receptor) to a consumer. The fraction of messages that need clarification (b) address the efficiency of the infusion process. If all of the material is well understood into highly encrypted, and without a lot of noise, the technology is passed directly to the consumer. At the macro diffusion level, looking at the entropy rate of change for the ensemble of nodes, we see the associated clarification (b's) which give us the average rate for the request for feedback (lack of understanding) of a technology. This in turn can be fed to infusion, where the technology program manager and adopter organization can further study the details of the infusion process. Individual b values for an organization and a given technology can be measured, if it is so desired.

3. Technology Transfer and the Learning Curve (Nishiyama 2000), (Hanakawa 1998)

During infusion, there is evidence that the learning curve is in play. The skill level and the improvement in productivity due to the technology, productivity loss during transfer, and the combined effects, net gain (Nishiyama 2000). The learning curve impacts on assimilating a new technology into a project were seen by the number of tasks performed over a study of several projects (Hanakawa 1998). This study in software development and others suggest the learning curve of Newell and Rosenbloom (Newell 1981) for power law *chunking* is appropriate for the various types of learning that need to be handled. This paper refines the basic control model $f(x)$ with the power law learning curve chunking model equations. While this is not important for the development of the basic model in this research, it provides the linkage to all manner of studies of organizational learning and ultimately, the breakeven and return on investment curves (Nishiyama 2000). This can be developed to make resource decisions, both for the infrastructure and for a specific research program or organization.

There is a broad base of literature on learning curves. During the study for this research, a large number of papers were reviewed. (Anderson 1981, Guiliksen 1934, Knecht 1974, Langley 1981, Lewis 1981, Mazur 1978, Newell 1981, Nembhard 2000, Miller 1956, Vigil 1994, Yelle 1979) and many more. Several of these are discussed elsewhere (Saboe 2002) in the learning curve section of that research.

These papers developed the basic relationships from learning curves, through relevance to software engineering. Anderson (Anderson 1981) is from Carnegie Mellon University, and the book he compiled under NSF and DARPA funding has a strong bent to showing the relevance to software development. (Langley 1981), (Lewis 1981), (Newell 1981). Linkage to distributions of terms and statistics of language and Zipf's law for the principle of least effort, are connected through (Mandlebrot 1953), (Simon 1955), (Snoddy 1926), and (Zipf 1949, 1965).

Mapping of Motives of Actors (Pfleeger 1999)

While the work by Pfleeger (Pfleeger 1999) never explicitly defines technology transfer, it provides the most comprehensive literature summary of the essential software technology literature. While not addressing all of the transfer field literature, or even all of the software technology studied in this area, the paper is an excellent review, a great overview and starting point. There are several important contributions beyond the survey of the field. She describes the process and roles involved in order to move technology in a transition from idea (technology creation) to adoption (technology diffusion). The generation of evidence, packaging, support and attention to the audience are identified as essential elements in the process of transfer. In this paper, these characteristics are primarily addressed in the clarification (b) in the control model. The clarification (b) values are driven by the commonality of terms to the audience measures in terms of the frequencies and entropy metric.

Pflegger also maps the motivations of the adopters to the category of adopter (innovators, early adopters, etc. per Rogers 1983) (Table 1). Also identified are the effects of rules imposed on an organization, a standards committee or a customer. These rules can encourage the success of a technology (this push or pull) when other models fail. For instance, she cites the effect of the Department of Defense's endorsements of products, recommendations for process improvement, or mandatory rules about tools as a positive influence to encourage "laggards" to take risks and try new technologies. The successful technology requires not only a new idea, she claims, but also a receptive audience with a particular adoption style. The various models (people mover, communications, on the shelf, vendor and rule as introduced by Pflegger) are mapped to the level of risk the adopter community is willing to take.

Adopter Category	Level of Risk	Adopter Model
Innovators	Very High	People-mover model
Early adopter	High	Communication model
Early Majority	Moderate	On-the-shelf model
Late Majority	Low	Vendor model
Laggards	Very Low	Rule model

Table 1 Relationships among Adopters, Risk and likely Transfer Model.
(Source: Pflegger 1999)

So to reduce the impedance mismatch between researcher and the method of moving the technology, "message" has to be matched with the audience. While Pflegger cites Zelkowitz and other studies that look at the actual implementation details of the transfer process, it is useful to note the factors that affect clarification requests (b) in this research. Another way to view the stream of messages is to suggest all that does not move to the consumer is in the feedback-entropy streams. Pflegger, Zelkowitz, the SEI and others generally are looking at the implementation details of technology transfer. All of the research to date generally looks at technology transfer from this perspective. This paper addresses a macro process, useful to the research manager and program managers, to assess the risk of the technology maturing at a given time. Implementation in a specific program of a technology should try to minimize the clarification requests (b). This is done by insuring that the mismatch is minimized, using messages that are matched for the audience. The message is packaging of the evidence . Pflegger (Pflegger 1999) and Schum (Schum 1994) describe evidence.

Types of Evidence	Characteristics
Tangible	Objects Documents Images Measurements Charts Relationships
Testimonial (unequivocal)	Direct Observations Second-hand Opinion
Testimonial (equivocal)	Complete equivocation Probabilistic argument
Missing tangibles or testimony	Contradictory data Partial data
Authoritative records or facts	Legal documents Census data

Table 2. Messages in Forms of Evidence.
(Source: After Schum 1994, Pflegger 1999)

Schum presents the categories of evidence seen in Table 2. The specific observational sense, objectivity and veracity of the message enable decisions to adopt or not adopt. In terms of this paper, if message is clear, unambiguous, and well understood, the advocate can pass on the message to the receptor with little to no requests for feedback. Schum and Pfleeger argue for this packaging of the message. This research supports those observations with the Shannon entropy component where noise and non-signal are minimized, e.g. the vocabulary converges between advocate and receptor.

4. INFORMATION THEORY APPLIED TO TECHNOLOGY TRANSITION

Informally, information measurement can be understood as anything that increases the variance also increases the information. Generally, variance is usually stated in units of measure, e.g. meters, volts, etc. The amount of information is a dimensionless quantity. When we have a large variance, we are very ignorant about what is going to happen. If we are very ignorant, then when we make an observation, it gives us a lot of information. On the other hand, if the variance is small, we know in advance of our observation how the result is likely to come out; hence, we get little information from making the observation.

1. *Information, Uncertainty*

Information is a difference in matter-energy [change of status – i.e. state] that affects the uncertainty in a situation where a choice exists among a set of alternatives (Rogers Kincaid 1981). "Information is something which reduces uncertainty. Communication is exchange of information." (Wiio 1980, p. 18) Information is the ability to choose between alternatives reliably. Before you send me an email, I cannot reliably, guess your message. After I receive it, I can do so. I have gained information (www.aip.org).

Uncertainty is the degree to which a number of alternatives, the multiplicity of options, are perceived with respect to the occurrence of the event and the relative probability of the outcomes. Uncertainty implies a lack of predictability, of structure and /or information. This multiplicity of option states can be quantified in terms of *entropy*.

Entropy and uncertainty can be considered synonymous (Jaynes 1957). Jaynes made the linkage between statistical mechanics as we know it from (Gibbs 1903), and entropy as we know it is thermodynamics, by relating a common concept to both – maximum entropy. Mathematically, maximum entropy has the important property that no possibility is ignored. It assigns positive weight to every possible situation that is not absolutely excluded from the information. It is the state where we can deal with equilibrium properties. According to Jaynes, this is quite similar to an ergodic property.

The macro equilibriumstate of a system (this is what we see in classical thermodynamics), is the macro equilibrium entropy, *S*. From Boltzmann, we get

$$S = k\mathcal{P}\ (\{p_i\}) \qquad (4.1)$$

This is when the maximum value $\mathcal{P}$ of the statistical entropy functional $\mathcal{P}$ ($\{p_i\}$) through the Boltzmann constant[5] k. Where $\mathcal{P}$ ($\{p_i\}$)= ln Ω is the uncertainty. Where k for {nats, bits, bytes, or Joules/ ^{o} Kelvin} is $\{1, \frac{1}{\ln 2}, \frac{1}{\ln 256}, 1.38X10^{-23}\}$ respectively. We can convert the natural log, ln, to $\log_2$ easily.

$$\log_2 x = \frac{\ln x}{\ln 2} \qquad (4.2)$$

The probability distribution $\{p\}$ is on the set of available microstates Ω={i} or multiplicity. The functional S=$k\mathcal{P}$ ($\{p_i\}$) needs to satisfy two general properties. (i) $\mathcal{P}$ must be positive, taking the value zero only in the case of absolute certainty ($p_i = 0$ for all states, except for a given state j for which $p_i = 1$).

5 Shannon (1948) quickly points out that k is just a convenient constant to relate to our physical world.

(ii) P must increase monotonically with increasing uncertainty. In addition, a third condition is required. (iii) The P is additive for independent sources of uncertainty (Bayes 1763), (Planes 2002). Because of this, we have the property of extensibility. This means if you add or subtract these quantities which contribute to uncertainty, the system size – the extent -- changes. Adding these quantities requires a product of the probabilities.

We can compose a system like this, with a system composed of two subsystems which are independent, *A* and *B*, so that the set of microstates is $W_{A+B} = W_A ´ \ W_B$. Each microstate (*i*,*j*) can be specified by fixing a state $i\,Î\,W_A$ of subsystem *A* and a state $j\,Î\,W_B$ of subsystem *B*. If a probability density, $p_{(i,j)}^{A+B} = p_i^A\, p_j^B$, then $P^{\ A+B} = P^{\ A} + P^{\ B}$. (Planes 2002), (Munster 1969).

$$P\ (\{p_i\}) = -\sum_{i\in\Omega} p_i \log_2(p_i) \quad (4.3)$$

2. *Extensive and intensive properties*

Extensive properties in the physical world are volume, mass, particles, energy, money, messages, records, etc. *Intensive* properties (e.g. pressure and temperature) on the other hand are independent of the size of the system. A method to determine whether a property is extensive or intensive is to divide the system into two equal parts with a partition. Each part will have the same value for the intensive properties, but half for the extensive properties.

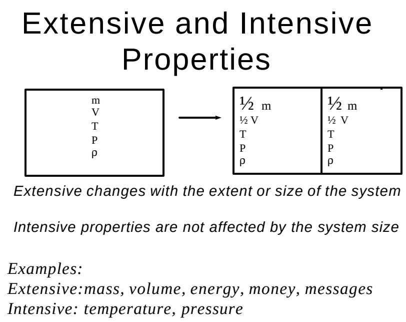


Figure 4-1 Extensive and Intensive properties

It would be valuable to identify analogous extensive and intensive properties in the technology transition model, or in or general terms.

Property	***Ext*ensive**	***Int*ensive**	**Thermodynamics/ Physical**	**Tech Transfer/ Information/ Communication System**
Particle Mass	X		• *N* particles per mole	• Unit of entities, e.g. Term per some standard message length
Volume	X		• L^3 *(length^3)* or • *AL (Area * length)*	• *v * s* nodes consisting of authors * state change
Energy	X		• eV, Joules, BTU's	• Some conserved property • Messages, information
Temperature		X	• ^{o}K degrees kelvin	• Some measure of change is cardinal related to two variables *ext* and or *int*
Entropy		X	• S≥0 • $S=kP(\{p_i\})$ • *S = k ln W* • Always increases • Additive for Independent Identical Distributions	• Similarly defined for information (Shannon 1948) • $S=kP(\{p_i\})$ • $S=- S\ p_i \log_2 p_i$ • Maximum entropy – uniformly distributed probabilities, same as thermodynamics
Pressure		X	• Force per Area	• Messages per node
Density		X	• Extensive property per volume	• Messages per node (author)

Table 4-1 Property Relationships

Particles are analogous to sets of terms in a message in this model. A message is made up of sets of terms. Counting all of the sets of terms is the same as determining the number of entities, particles. Just like in molecules, some entities have more weight than others. If all null and single term sets have the same weight, the analogy is a set of sets of terms e.g. {}, {A}, {B}, {C}, {AB}, {AC}, {BC}, {ABC}. {A} is "lighter" than {AC} which is a composite of two if a term is made up of {A}+{C}. There should be some relationship between changing the status of a term and analogous principles in the physical world. e. g. Newton's laws (see the next section).

Volume in the physical world, is in three dimensions measured in some length units. We can get a volume with units of l^3 by measuring the volume. Integration over small *dl* is used in continuous space. For a discrete system, we count the points defined in phase space. For the models, this volume is defined in only two dimensions, nodes (a publisher) and state points.

In a classical thermodynamics model, energy is measured in Joules, or BTU. It is often convenient to measure energy units in electron volts, which is the kinetic energy of an electron that has been accelerated through a voltage difference of one volt. This is moving an electron from its status at point A to point B. This is directly related to the conservation principle, the 1st law of thermodynamics, and Newton's 3rd law. The first law of thermodynamics says that *energy is conserved and transformed*. Energy is a primitive and essential thermodynamic function. It is a mathematical abstraction. (Abbott 1989, p1). Newton's 2nd and 3rd laws similarly constructed using the principle of conservation.

Law 1 *"Every body preserves in its state of being at rest or moving uniformly straight forward except insofar as it compelled to change its state by forces impressed."*

Law 2 *"A change in motion is proportional to the motive force impressed and takes place along a straight line in which a force is impressed."*

Law 3 *"To any action [change of state] there is always an opposite and equal reaction; in other words, the actions of two bodies upon each other are always equal and always opposite in direction"*[6]. (Newton 1726, p417).

Newton says in definition 3 of law 1, "because of inertia of matter, it is only with difficulty put out of its state either of resting or of moving." In Newton's interleaved copy of edition 2, he adds the following which was never printed: "I do not mean Kepler's force of inertia, by which bodies are moved toward rest, but a force of remaining in the same state either of resting or moving." (Newton 1726 p404). Change of state, status, must overcome some inertia. E.g. changing v_0 *to* v_1 meaning to change from an initial state, say a velocity, to a new velocity. Even to change one orientation of one atom, or one bit, such a change of state, takes some force or stimulus. Something must happen to change the state of information otherwise it stays in its current state.

Below we show the relationship using a Venn diagram, that shows the probability of two sets can represent this conservation through correlations of extensive properties at the intersection consisting of mutual information. The left hand subsystem *A* is composed of the sum of the uncorrelated part P $(A|B)$, plus the correlated part $I(A;B)$ still equal to the total and the P (A), where $I(A;B)$ is the shared mutual information. This is the equal and opposite amount required by the 2nd and 3rd laws of Newton. Similarly, the right hand subsystem *B* is composed of the sum of the uncorrelated part P $(B|A)$, plus the correlated part $I(A;B)$ which is still equal to the total and the P (B). Looking at relation 4, $I(A;B)=I(B;A)$ and other relations in **Figure 4-3**, we see how the conservation principle is realized. The key is not conservation of energy in this research, but rather the conservation of the correlated components of extensive properties in two interacting subsystems (Planes 2002). What one subset looses, the other gains.

3. *Entropy Review*

Entropy, as a concept, can readily be seen as logical entropy (think of it as a measure of uncertainty, noise, non-signal, process inefficiencies, the percentage of work resulting in defects and requiring rework, etc.) and physical or thermodynamic entropy (i.e. mixed-up-ed-ness, disorder, disorganization, etc), which is the quantity of energy not available to do work. Logical entropy is Shannon's entropy (S_H) as defined by Shannon on his treatise on communication theory (Shannon 1948). Shannon's theory says that the entropy of an information source measures how well its behavior (e.g. the next symbol in a sequence it produced) can be predicted.

Mixing entropy can be represented by the eigenvalue of a bakers' transformation function. This bakers' transformation in state space represents entropy in terms of folding, stretching, translation and rotation (Speigel 1998 p292). This transformation is the representation of a *dissipative structure.* These are structures with an innate capacity to dissipate anything

[6] This is the exact statement taken from Newton's original work. Modern texts have often changed the wording slightly on each of his laws, but the original statements give us the closer intent of the law to this

that comes in to disturb the system. The term "dissipate" is somewhat unfortunate, because what really occurs is *integration* not *dissipation* (O'Murchu 1997 p.168). The entropy is the quantity of information not available to help us work, yet is valuable to understand if the objective is propagation and diffusion. The relationships are developed below.

Recently a number of undergraduate texts are illustrating entropy as the accessible state multiplicity for quantities that must be conserved -- e.g. volume, and particles. The notion of conservation of a quantity is important to this research, as this could be momentum or more importantly information. This is understood from the logical-mathematical interpretation of the equations vs. physical interpretations. It requires us to step back and look at conserved quantities in the mathematical sense, then map those to our problem. Further, entropy, temperature or coldness (*1/T*) and heat capacity have been developed on the basis of information units alone (Fraundorf 2000).

The definition of information entropy here is related to the definition of entropy in thermodynamics. What follows is a basic review of entropy in information theory after Shannon, Jaynes, Kolmogorov, Uspenski, and others as found in Li, (Li 1993) and Cover (Cover 1991). This review section is drawn from Cover (Cover 1991 p13).

Let X be a discrete random variable with alphabet X and a probability mass function $p(x)=\Pr\{X=x\}$, $x \in X$. $p(x)$ and $p(y)$ refer to two different random variables and are in fact two different probability mass functions $p_x(x)$ and $p_y(y)$.

The definition of information entropy is:

$$S_H(X) = -\sum_{x\in \Xi} p(x)\log_2 p(x) \qquad (4.4)$$

S_H is the entropy measured in bits, and the log is base 2. For example, the entropy of a fair coin toss is 1 bit. The convention of 0 log 0 =0 is used, which comes from continuity since x lox x $\to 0$, as x $\to 0$. The base of the log is two for the natural units as developed by Shannon (Shannon 1948). The entropy is a function of the distribution of X. It does not depend on the actual values taken by the random variable X, but only on the probabilities.

If $X \sim p(x)$, then the expected value E of a random variable $g(X)$ is denoted

$$E_p g(X) = \sum_{x\in \Xi} g(x)p(x) \qquad (4.5)$$

The entropy of X can be interpreted as the expected value of $\log \dfrac{1}{p(X)}$, where X is drawn according to the probability mass function $p(x)$. Thus

$$S_H = E_p \log \frac{1}{p(X)} \qquad (4.6)$$

Certain properties must be satisfied. This is done axiomatically or by answering some natural questions such as "What is the average length of the shortest description of the random variable?" As a result we find

$$\begin{aligned} &S_H \geq 0 \\ &0 \leq p(x) \leq 1 \text{ implies } \log(1/p(x)) \geq 0 \end{aligned} \qquad (4.7)$$

Here is an example. Let

$$X = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases} \qquad (4.8)$$

then

$$S_H(X) = -p\log p - (1-p)\log(1-p) \equiv S_H(p) \qquad (4.9)$$

We see that S_H = 1 bit when $p=1/2$. **Figure 4-2** shows the basic properties of entropy. It is a concave function of the distribution and equals 0 when $p=0$ or 1. This makes sense because when $p=0$ or 1, the variable is not random and there is no uncertainty. The entropy is maximum when $p=.5$, which corresponds to the maximum value of the entropy.

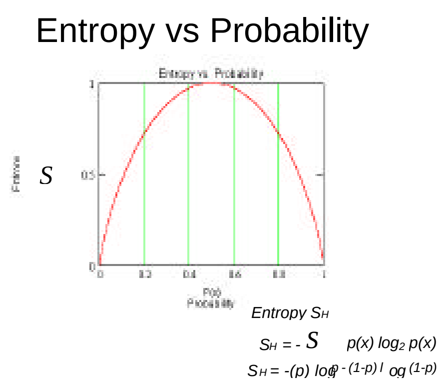


Figure 4-2 Entropy vs Probability

Here is another example. Let

$$X = \begin{cases} a & \text{with probability } 1/2, \\ b & \text{with probability } 1/4, \\ c & \text{with probability } 1/8, \\ d & \text{with probability } 1/8. \end{cases} \qquad (4.10)$$

The entropy of X is

$$S_H(X) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{4}\log\frac{1}{4} - \frac{1}{8}\log\frac{1}{8} - \frac{1}{8}\log\frac{1}{8} = \frac{7}{4} \text{ bits} \qquad (4.11)$$

Suppose we wish to determine the value of X with the minimum number of binary questions. An efficient first question is "Is $X=a$?" This splits the probability in half. If the answer to the question is no, the second question can be, "Is $X=b$?" The third question is "Is $X=c$?" The resulting expected number of binary questions is 1.75. This turns out to be the expected number of binary questions required to determine the value of X. It can be shown that the minimum number of binary questions required to determine X lies between $S_H(X)$ and $S_H(X+1)$.

Let's now introduce the definitions for joint and conditional entropy and mutual information.. These are key facets of the technology transfer models proposed.

Joint entropy $S_H(X,Y)$ of a pair of discrete random variables (X,Y) with a joint distribution (X,Y) can be considered to be a single vector-valued random variable. It is defined as

$$S_H(X,Y) = -\sum_{x\in\Xi}\sum_{y\in\Psi} p(x,y)\log p(x,y) \tag{4.12}$$

which can also be expressed as

$$S_H(X,Y) = -E\log p(X,Y) \tag{4.13}$$

The conditional entropy of a random variable given another is defined as the expected value of the entropies of the conditional distributions, averaged over the conditioning random variable. If $(X,Y)\sim p(x,y)$, then the conditional entropy $S_H(Y|X)$ is

$$S_H(Y\,|\,X) = \sum_{x\in\Xi} p(x) S_H(Y\,|\,X=x) \tag{4.14}$$

$$= -\sum_{x\in\Xi} p(x)\sum_{y\in\Psi} p(y|\,x)\log p(y|\,x) \tag{4.15}$$

$$= -\sum_{x\in\Xi}\sum_{y\in\Psi} p(y,x)\log p(y|\,x) \tag{4.16}$$

$$= -E_{p(x,y)}\log p(Y\,|\,X) \tag{4.17}$$

This is shown in the Venn diagram in **Figure 4-3**.

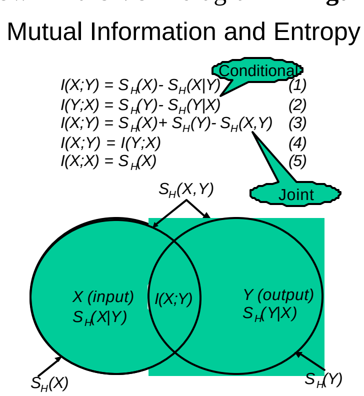


Figure 4-3 Mutual Information, Joint and Conditional Entropy

Relative entropy or the Kullback Leibler distance between two probability masses $p(x)$ and $q(x)$ is defined as

$$D(p\,\|\,q) = \sum_{x\in\Xi} p(x)\log\frac{p(x)}{q(x)} \tag{4.18}$$

$$= E_p \log\frac{p(X)}{q(X)} \tag{4.19}$$

Similar to earlier developments, we use the convention based on continuity of arguments that $0\log\frac{0}{0} = 0$ and $p\log\frac{p}{0} = \infty$. (Cover 1991, p18)

While it is not a true distance between distributions, it is useful to think of relative entropy as a "distance" between distributions. The mutual information which was introduced before is the measure of the amount of information that one random variable contains about another random variable. It is the reduction in the uncertainty of one random variable due to the knowledge of the other. Assume we have two random variables X, and Y with a joint probability mass function $p(x,y)$ and marginal probability mass functions $p(x)$ and $p(y)$. The mutual information $I(X;Y)$ is the relative entropy between the joint distribution and the product distribution $p(x)p(y)$, i.e.,

$$I(X;Y) = \sum_{x\in\Xi}\sum_{y\in\Psi} p(x,y)\log\frac{p(x,y)}{p(x)p(y)} \tag{4.20}$$

$$= D(p(x,y)\,||\,p(x)p(y)) \tag{4.21}$$

$$= E_{p(x,y)}\log\frac{p(X,Y)}{p(X)\,p(Y)} \tag{4.22}$$

It is important to see that the mutual information $I(X;Y)=I(Y;X)$

$$I(X;Y) = S_H(X) - S_H(X\,|\,Y) \tag{4.23}$$

The mutual information $I(X;Y)$ is the reduction in uncertainty of X due to knowledge of Y. By symmetry it follows that

$$I(Y;X) = I(X;Y) = S_H(Y) - S_H(Y\,|\,X) \tag{4.24}$$

That is X says as much about Y as Y says about X. Since $S_H(X,Y) = S_H(X) + S_H(Y\,|\,X)$ we have

$$I(X;Y) = S_H(X) + S_H(Y) - S_H(X,Y) \tag{4.25}$$

In addition, we see that

$$I(X;X) = S_H(X) - S_H(X\,|\,X) = S_H(X) \tag{4.26}$$

The mutual information of a random variable with itself is the entropy of the random variable.

Mutual information and the symmetry we see here is what will enable the conservation principle to be met. As X correlates with Y it is realized in the same amount of mutual information. This is easy, as we saw in **Figure 4-3**.

To bridge the gap between communication theory and capacity of human performance an analogy is made between, a human accepting input and generating output and a communication system. This is seen as the overlap in a Venn diagram, where input variance is represented by the circle to the left and the output variance is the circle to the right, and at the intersection is the amount of transmitted information. Miller (Miller 1956) suggests that an individual is a communication channel. He states for a human, “when we increase the amount of input information, the transmitted information will increase at first and will eventually level off at some asymptotic level.” He indicated that this is the *channel capacity* of the observer, the human.

Now we have enough information theory to understand the models. Consider the entropy as a representation of the terms in a vocabulary, which are available to the researchers in a time step. A researcher reaches into the pool of messages, which are constituted by terms. We can compute entropy contribution of a term in a given time step by computing $p(x)$ for the term. Summing all of the terms we have the entropy at time step k.

The balance of the paper will discuss the findings of the model. The details will follow. Here is a brief, heuristic summary of the data using illustrations.

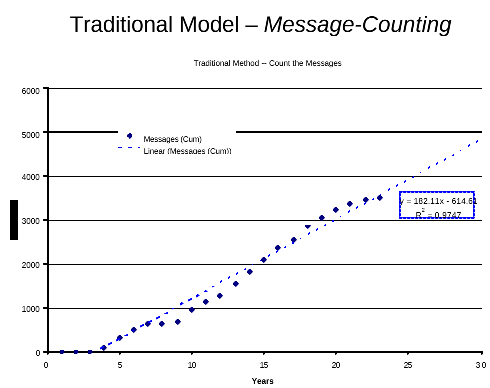


Figure 4-4 Message Counting Linear Model

The message counting model seen in Figure 4-4 typically used provides a very good correlation and is quite linear. Possibly an information theoretic and dynamical systems model can be built that enables richer analysis.

For an information – communication model to work we need to determine the change in entropy over a time step. In Figure 4-5, we see how entropy and messages vary over time. Messages are a conserved extensive quantity, and the information entropy is related to the message terms.

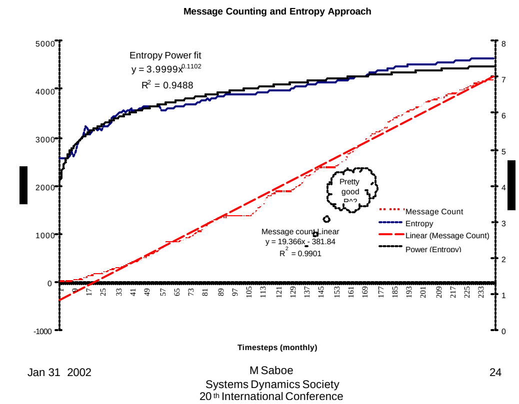


Figure 4-5 Entropy and messages over time

In Figure 4-5, we see that we would like an illustration of the joint entropy related to technology at a given time step. Further, we would like a method to compare to different technologies, Figure 4-6. This is done through the mechanism of relative entropy.

Figure 4-6 illustrates two technologies, we now have a mechanism to determine how close these technologies are in a crude sense. However, there are other factors are work. For example, what is the mind share, the volume of nodes operating on the messages?

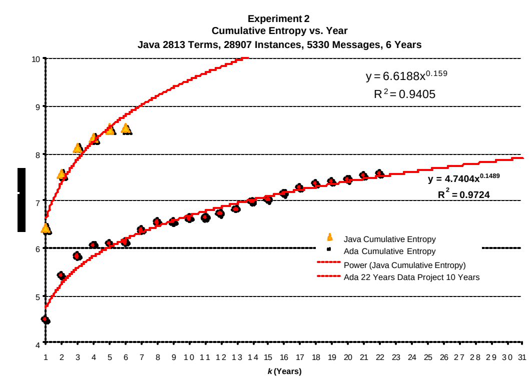


Figure 4-6 Entropy vs time

4. *Interacting Subsystems*

Let’s imagine a super system (the community’s world of knowledge) that consists of two subsystems. These subsystems represent what is known and what is unknown at a given time. The sum of the two subsystem’s extensive variables, messages N, and nodes V is constant. Here the conserved extensive variable properties are N messages, and the sum of all the nodes, v which is the volume. This will define a control volume. Now we will take a virtual partition and have it progress expanding subsystem A to the right. As this partition passes over some nodes, effort is made by the nodes and they "discover" a term. They stimulate and change the internal configuration of the system by converting an undiscovered term (a null) into a communicated discovered term.

This can be seen in Figure 4-7. On the left hand side we see "!!!s" representing terms that have been discovered (answers), on the right hand side of the partition we see "???s" representing terms that are yet undiscovered (questions). The Venn diagrams indicate the subsystems A and B, joint, conditional entropies and mutual information, as illustrated earlier. Now examine what this looks like with a sample alphabet as in Figure 4-8. The nulls {} are terms that have not yet been discovered at the frontier of the research in time. In this simplified example, we are assuming a fixed set of terms in the alphabet, and a fixed number of nodes. This will permit the development of the general relationships between extensive and intensive variables in a state equation. Later, once we have seen these relationships, we can start with an initial condition representing the number of terms and nodes, and add more vocabulary to the system or more author nodes any way we wish. This permits the design of a desired solution in the form of an engine.

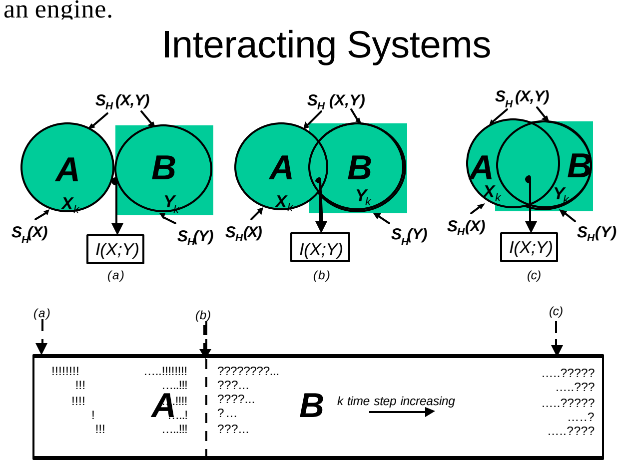


Figure 4-7 Interacting Systems A and B

FirstOfPubYear	CountOfdesc	descriptor_text	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992
1980	3	procedure-oriented-languages	{}	3	3	3	3	3	3	3	3	3	3	3	3	3
1981	140	Ada	{}	{}	14	16	22	30	30	43	51	57	101	101	109	140
1983	4	systems-analysis	{}	{}	{}	{}	1	1	1	1	1	1	1	1	1	4
1981	39	software-engineering	{}	{}	1	2	3	11	11	18	19	21	36	36	36	39
1981	32	program-compilers	{}	{}	5	6	9	12	12	14	18	19	30	30	31	32
1981	8	operating-systems-computers	{}	{}	1	1	1	3	3	3	4	5	7	7	7	8
1984	18	distributed-processing	{}	{}	{}	{}	0	1	1	1	2	4	13	13	15	18
1986	34	software-tools	{}	{}	{}	{}	0	0	{}	4	5	8	22	22	23	34
1984	27	real-time-systems	{}	{}	{}	{}	0	1	1	1	2	4	8	8	12	27
1989	8	object-oriented-programming	{}	{}	{}	{}	0	0	{}	{}	{}	{}	3	3	3	8
1988	11	scheduling	{}	{}	{}	{}	0	0	{}	{}	{}	1	5	5	6	11
1981	23	programming	{}	{}	1	1	2	6	6	7	8	8	17	17	17	23
1991	9	software-metrics	{}	{}	{}	{}	0	0	{}	{}	{}	{}	{}	{}	1	9
1992	4	standards	{}	{}	{}	{}	0	0	{}	{}	{}	{}	{}	{}	{}	4
1986	3	structured-programming	{}	{}	{}	{}	0	0	{}	1	1	1	2	2	2	3
1989	6	fault-tolerant-computing	{}	{}	{}	{}	0	0	{}	{}	{}	{}	4	4	5	6
1989	2	software-reliability	{}	{}	{}	{}	0	0	{}	{}	{}	{}	1	1	1	2
			!!!	!!!	!!!	!!!	!!!	??	??	??	??	??	??	??	??	??

Figure 4-8 Subset of an alphabet in two Interacting systems !!! and ???

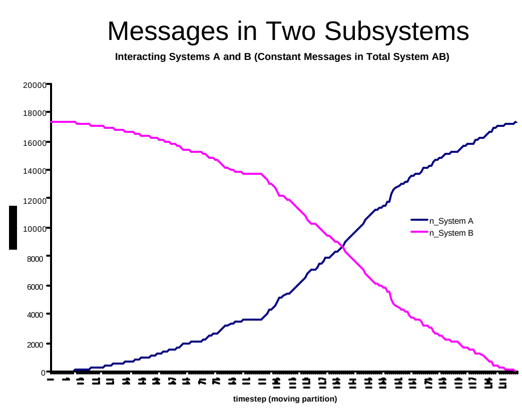


Figure 4-9 Messages in two subsystems

In Figure 4-9, we see that as system *A* expands, the number of terms discovered increases, at the same rate that the number of terms undiscovered decreases. This model satisfies our conservation principle for extensive quantities.

Next, in Figure 4-10, we examine the entropy relationship. The horizontal line at the top of the figure is the joint entropy of the system. Since this is a closed system, this is not changing, however, the internal distribution will change. That entropy related to subsystem A will increase, as there are more and more choices to make in order to get complete information. Subsystem B will decrease from a high entropy (all of the unknown terms) to a lower entropy as there becomes less and less left to be discovered. The lower curve shows the mutual information. When the distance between the center of the two probability masses, or subsystems, decreases, there is a higher correlation.

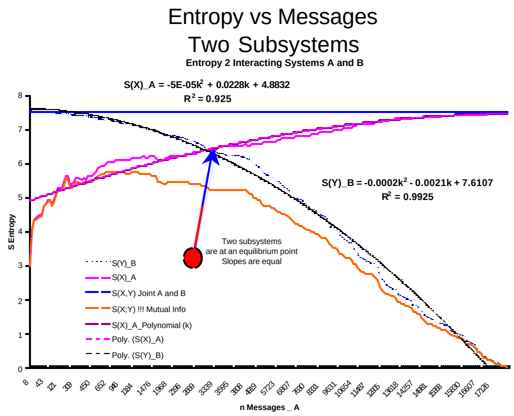


Figure 4-10 Entropy vs Messages Two interacting Systems

Following reasoning similar to that used in statistical, and condensed particle physics (Schroeder 2000) (Fraundorff 2000), we can find some useful relationships. The slope of the curves of the two subsystems give us some important information about thermal equilibrium. Recall the from the canonical ensemble discussion of free energy, that the temperature T is the parameter controlling free energy, or the conserved property. In this case of messages, we can write

$$\frac{1}{T} = \frac{\Delta S_H}{\Delta n} \tag{4.27}$$

So the temperature is related to slope of the change in entropy to change in messages curve. When the curves in the figure cross over, the system is at an equilibrium point. Let's look at a general relationship that shows the increase in one system is related to the negative slope or, the decrease in the other.

$$\frac{\Delta S_A}{\Delta n_A} = -\frac{\Delta S_B}{\Delta n_A} \tag{4.28}$$

The incremental change in $S_{A,}$ divided by the change in n_A messages, is equal to the change in entropy, $S_{B,}$ for system B again compared to the change in the conserved quantity, in this case n_A. Rewriting we get

$$\frac{\Delta S_A}{\Delta n_A} + \frac{\Delta S_B}{\Delta n_A} = 0 \tag{4.29}$$

The second term has a B in the numerator and A in the denominator. Dn_A is the same as $-Dn_B$, since what we discover in messages is the same as what is removed from the undiscovered system. We can rewrite this for a system at equilibrium as

$$\frac{\Delta S_A}{\Delta n_A} = \frac{\Delta S_B}{\Delta n_B} \tag{4.30}$$

The thing that is the same for both systems when they are at thermal equilibrium is the slope of the entropy message graph. This slope must somehow be related to the temperature of the system. The 2nd law of thermodynamics tells us that the conserved property will tend to flow *into* the subsystem with the *steeper* entropy vs. message graph, and *out* of the object with the *shallower* entropy vs. message graph (Schroeder 2000 p87).

According to Schroeder, the former "wants to" gain the free conserved property (messages) in order to increase its entropy. If there is an imbalance between the two subsystems, the latter doesn't so much "mind" losing a few messages (since the entropy will not decrease much). A steep slope must correspond to a *low* temperature, while a shallow slope corresponds to a *high* temperature.

Now we can see in the lower curve of Figure 4-13, the relationship of the temperature (the right had Y axis) of sub-system A as the partition moves over the time steps. More activity increases the temperature. The temperature is measured in degrees as we would in a physical system; however, these degrees are developed from information units. This is "the" fundamental temperature unit developed from the relationship of entropy, and the conserved quantity.

Note that there are temperature fluctuations. This is consistent with Prigogine's observation about evolving systems. A dynamical system will help explain these fluctuations.

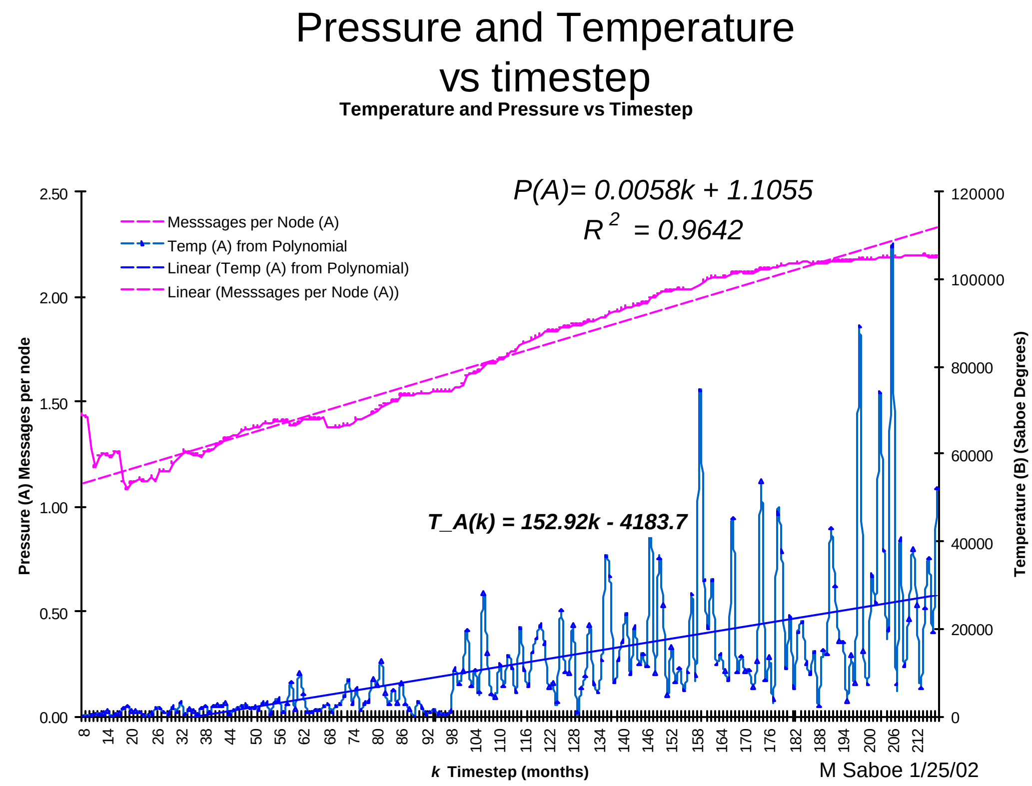


Figure 4-11 Pressure and Temperature °SaboeÓ vs time – two interacting systems

Pressure is defined as the <messages> processed per node, where the <messages> represent the average in the time step per node. The pressure can also be seen to increase as the temperature increases. While messages are not physical molecules as in a thermodynamic system, they seem to behave as a gas might, as the temperature goes down the pressure goes down.

Figure 4-13 shows the relationship directly between pressure and temperature.

Typically, a state diagram viewed by engineers is a temperature – entropy, or *T-S*, diagram. In the lower curve of Figure 4-14, the *T-S* is illustrated. This is the entropy of sub-system *A* with entropy (upper *x* axis) and temperature (secondary *y* axis on the right).

The figure also shows entropy of subsystem *A* (left *Y* axis) and messages *n* on the *x* axis. From this information in a closed system we can see the trends for a given technology over time. In a way we have the ability to define the heat capacity[7] (say C_p, heat capacity at constant pressure) in bits. This allows us to move to an open system, like an engine and add nodes, *volume*, and increase message flow. We can then compute our effort required from a desired "engine" to develop a technology.

$$\Delta U = \dot{n} C_p \Delta T \qquad (4.31)$$

This also implies the equivalent of Carnot's cycle, which can tell us the maximum efficiency we can expect.

Another interesting point is that the set of sets of terms, reduced to primitive message combinations follows a Boltzmann distribution, Figure 4-12. On the *x* axis, is the q-level, representing the number of terms in set. The lower curve on the *y*-axis is the frequency of sets. The upper curve assigns a weight to each set. It is interesting to note, as well, that these curves plotted over the time steps examined (up to 21 years) essentially remain stationary.

This permits conjecture in the deeper meanings of the distribution of terms. Further, state transitions moving from one q-level to another, must somehow be affected by an impulsive stimuli of some sort. That implies both the notion of kinetic and potential "energy". These topics are subject for future research.

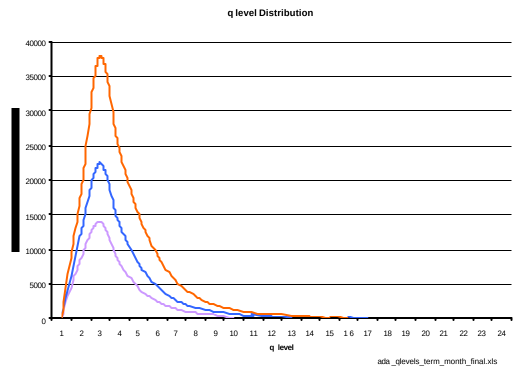


Figure 4-12 Boltzmann Distribution of Sets of Terms (primitive messages)

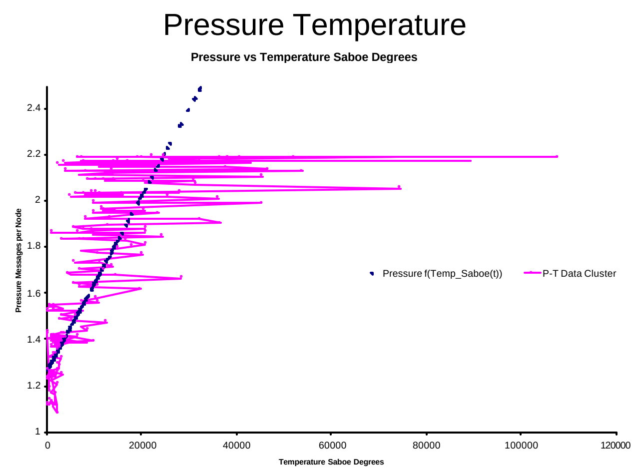


Figure 4-13 Pressure vs Temperature ^{o}Saboe Ó

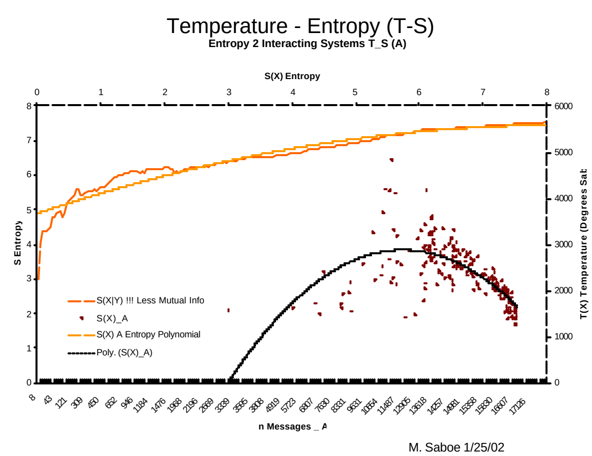


Figure 4-14 Entropy -- Messages, and Temperature – Entropy

[7] Heat capacity for sate equations are property relations and as such are independent of the type of process. C_p is the amount of "stimuli" transferred to a system per unit "message" per unit degree rise during a constant pressure process.

5. *Dynamical Systems Model*

This is the information theoretic view. Now we marry up a dynamical systems model with an information theoretic model. When both stabilize in a rate represented by equations of the same form, we have a match.

Figure 4-15 shows a map of the state space. The legend shows The Java Entropy map marked with a triangle (▲) and a dashed line. The marker represents data, the dashed line is an indicator of the curve that the data would be fit to. Similarly the circle (❍) and dashed line legend for the Ada points. The state space map is shows that the data is oscillating in the early stages. This shows that the vocabulary and threads of research have not settled down at first. Based on observation, see **Figure 4-15**, as the entropy increases, but at declining rate, the data starts to approach the $y=x$ line. The spacing between each data point gets closer together. This indicated that the data is moving toward a stabilizing attractor basin. A method to quantify this stabilization is discussed in the next sub-section.

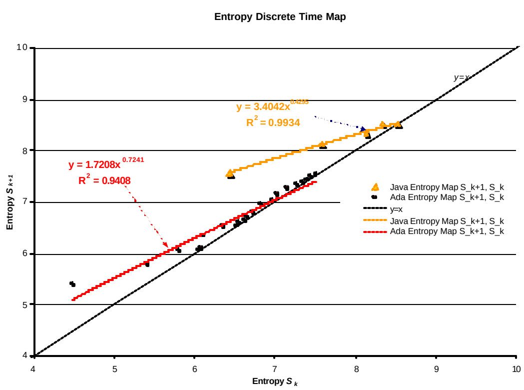


Figure 4-15 Java and Ada State Space Finite Difference Map Entropy (S_{k+1}, S_k)

6. *State Space Representation*

Recall that the bakers' transformation example illustrates the mixing of a spot of sauce on the piece dough, then folding and stretching of dough. In technology maturation, a node is locally taking in a chunk of dough, messages out of the pool of messages persistent in history, and mixing them along with new information, i.e. a new term, which represents yet another spot on the dough. These areas contain remnants from bakers' transformations of other nodes that performed the mixing and adding function throughout time. A performing node may perform a number of iterations. Other nodes also perform the folding, stretching and mixing function before and concurrent with the local node. The nodes successively repeat the iteration action. We can let X be the function that represents the value corresponding to the application of n bakers' transformations.

$$X_{n+1} = F(X_n) \qquad (4.32)$$

The various functions X_n are functions of internal time. The internal time is an *operator* like the one used in quantum mechanics. The age of partition X_n is the number n of iterations i that are to be performed to go from X_o to X_n. The eigenvalue of the characteristic equation has a relationship in natural units to entropy. This relationship is through the Lyapunov exponent, which gives the stretching rate per iteration averaged over the trajectory.

So in **Figure 4-15** we see a plot of a one-dimensional map. Taking the derivative in this case yields l. The goodness of fit is determined through the finite difference method known to determine convergence and stability points in dimensions using the Lyapunov number and the Lyapunov exponent λ. The Lyapunov numbers quantify the stability of an orbit around an attractor. The Lyapunov numbers are the absolute values of the eigenvalues of the Jacobian matrix at a fixed point. The Lyaponuv exponent is the logarithm of the Lyapunov numbers (Farmer 1983).

$$l = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln | f'(x_i) | \qquad (4.33)$$

These dimensions represent an entropy measure for non-linear systems in stable or chaotic regions.

We compute entropy two ways. One is from data, which is acquired experimentally. The other is from a model of the process of transferring (transforming) information. The experimental entropy data is related to the information we know about a topic. We refer to this as Shannon's entropy (S_H). The data S_H is gathered over time steps k. We perform regression on this data and have as a result a function that is of the power law form e.g. $y=bx^m$, where m is the slope and b is the intercept in log-log linear form. We also have a model of a non -linear dynamical system. The Lyaponuv exponent of a map gives the sensitive dependence upon initial conditions that is characteristic of chaotic behavior. Further discussion can be found in Prigogine (Prigogine 1983), Farmer, York Ott, (Farmer 1983), McCauley, (McCauley 1993), and Baker (Baker 1990). The following description follows the development found in Farmer (Farmer 1983) and Baker (Baker 1990).

The eigenvalue of the Jacobian[8] of the finite difference equations representing the dynamical system is also related to entropy. In fact, the relationship is through the Lyaponuv number. The Lyaponuv number and this relationship is defined as $J_n = [J(x_n)\ J(x_{n-1}) ..\ J(x_1)]$ where $J(x)$ is the Jacobian matrix of the map $J(x) = (\partial F/¶x)$ with $j_1(n)$ ³ $j_2(n)$… ³ $j_p(n)$ are the magnitudes of the eigenvalues of J_n. The Lyaponuv numbers are

$$l_i = \lim_{n\to\infty} [j_i(n)]^{1/n}, \ i = 1,2,\ldots,p \qquad . (4.34)$$

The Lyaponuv exponent is the smallest, positive, real *n*th root taken. We follow Farmer's assumption that *almost every* (Farmer's emphasis) initial condition in the basin of the attractor has the same Lyapunov numbers. This followed from his empirical evidence, and the data in this model does not meet the exceptional conditions that he identifies.

6. One Dimensional Finite Difference Representation of S_{H_k}

We determine the one-dimensional model for computation of this entropy for the *TechTx Basic Entropy* model in a form compatible with the two dimensional micro level model. This is

$$S_{H_{k+1}} = f(S_{H_k}) \qquad (4.35)$$

$$l = f'(\bullet) \qquad (4.36)$$

The macro entropy is partitioned and allocated to the performer and affiliated organization nodes. This enables computation of the system entropy at the nodal level. This provides the method of computing the Lyaponuv dimension from λ to measure the non-linear system entropy $S_{B_{micro}}$, at the micro level or for simplicity of notation, S_B. Note that this differs from the entropy S_H in Figure 4-6, which is the information entropy, NOT the entropy measure for the stability or chaos of the system.

The general form for the transformation is $S_{H_{k+1}} = f(S_{H_k})$. We have from our earlier *TechTx Basic Entropy* discussion the macro entropy vs time. While we recognize that we have to partition and allocate the entropy to the performing nodes, we can use the macro function for illustrative purposes here. We develop the relationships using a power law here. However, as experimentation progressed, it became apparent for the technology we were evaluating, the exponent was always almost 1. As the power law may be the right fit for some technologies, we develop this more general relationship here. For the linear fit, the derivative reduces simply to a constant – the slope *m.* At the end of the day for the linear fit proved to be a very good and simple relation that gave most satisfactory results. Having fit the entropy over time, if we have a power function, it is in the general form of $S_{H_k} = bk^m$

To get to the general finite difference form, we have

$$S_{H_k} = bk^m$$

$$k = \left(\frac{S_{H_k}}{b}\right)^{\frac{1}{m}} \qquad (4.37)$$

$$S_{H_{k+1}} = b(k+1)^m$$

Recall the general form of the finite difference transform is

$$S_{H_{k+1}} = f(S_{H_k}) \qquad (4.38)$$

To get to the derivative, we use (4.37) eliminate *k* resulting in

$$S_{H_{k+1}} = b\left[\left(\frac{S_{H_k}}{b}\right)^{\frac{1}{m}} + 1\right]^m \qquad (4.39)$$

To find l we get

$$l = \frac{dS_{H_{k+1}}}{dS_{H_k}} = \left[\left(\frac{S_{H_k}}{b}\right)^{\frac{1}{m}} + 1\right]^{m-1} \left(\frac{S_{H_k}}{b}\right)^{\frac{1}{m}-1} \qquad (4.40)$$

Recall that λ was required to compute the Lyaponuv dimension from λ to measure the non-linear system entropy, S_B to quantify the stability of the system.

7. Two Dimensional Finite Difference Representation of S_{H_k}

Similarly, we develop a two dimensional model using the finite difference method. For *n*

[8] The Jacobian matrix is simply the derivative of a *p*-dimensional map function *F*. We use the form $X_{n+1} = F(X_n)$, where X is a *p*-dimensional vector

dimensional maps there are n Lyapunov exponents l_i, since stretching can occur for each axis.

A two dimensional model[9] is used for the computation of the Lyapunov dimension from l to measure the non-linear system entropy S_B.

$$\begin{cases} S_{H_{k+1}} = F(S_{H_k}, N_{i_k}) \\ N_{i_{k+1}} = G(S_{H_k}, N_{i_k}) \end{cases} \qquad (4.41)$$

Functions F and G are defined as one-to-one functions in R. We assume that the partial derivatives exist. Now using l as defined in (4.41) or **(4.42)** $l = \lim_{n\to\infty} [j_i]^{\frac{1}{n}}$ where j_i are the eigenvalues of $|A - j_l I| = 0$ and A is the Jacobean of transformation is defined as $D\mathbf{T}$

$$DT = \frac{\partial(F,G)}{\partial(S,N)} = \begin{bmatrix} \dfrac{\partial F}{\partial S_{H_{k+1}}} & \dfrac{\partial F}{\partial N_{i_{k+1}}} \\ \dfrac{\partial G}{\partial S_{H_{k+1}}} & \dfrac{\partial G}{\partial N_{i_{k+1}}} \end{bmatrix} \qquad (4.42)$$

Here we are computing F and G to develop the transfer function and to correlate these two dimensions to determine S_B from λ, the Lyapunov exponent. The interesting feature of the bakers' transformation is that it is a dissipative function in state space if the sum of the exponents is negative.

The micro level and macro level computations both should approach or diverge from stability in the same manner, *if* the models are correct. In this case S_H *would* yield a strong correlation to S_B,. In this model there is an adjustable performance index parameter, to reflect efficiency. Elsewhere this related to the learning curve. The performance index parameters are adjusted to tune the micro model and to match the S_B. Using the *TechTx Learning Curve* model (Saboe 2002), we are provided a method to identify the performance bands and half-life of performance improvement, for maturing of the technology.

5. *TECHTX* FEEDBACK ENTROPY MODEL

1. *Communication and Control Model*

9 The two-dimensional model for the bakers transformation using both information and entropy is found in Saboe 2002. Only the one dimensional model is developed in this paper.

Let's give an example of information being exchanged at the micro level. Consider some coupled nodes in a communication system. This example is adapted from Brown (Brown 2000). This system described will be represented in a dynamical system model, which ends up being the bakers' transformation.

We can develop and represent the system in a model of information and the state as it flows from the advocate and receptor as seen in Figure 5-1. Model the following communication nodes, a *sender* (***S***), a *receiver* (***R***), and a *consumer (**C**)*. A simple function with inputs as messages and outputs as messages associated with each node carries the dynamical information about each node.

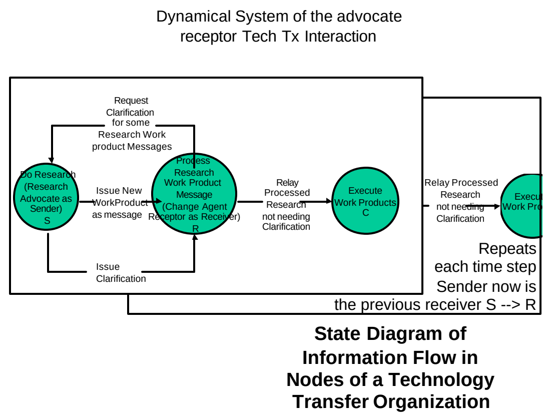


Figure 5-1. Dynamical System Model of Advocate-Receptor Interaction.

The sender is an advocate. This is a researcher, or in the terms of Fowler (Fowler 1994) an advocate and producer. The sender issues new work products as messages. The receiver is a change agent, or the receptor. The sender develops research, advances and publishes a message as a work product, thesis, article, technical report, demo, etc. The message is observable, e.g. measurable and countable. We can generally only measure output. We can measure output in terms of messages and terms from which the messages are made up. Except for one type of input, it is usually difficult to quantify, or measure all of the input.

The receiver receives the message. If the message is understood completely, i.e. no need for clarification, the receiver retransmits the processed message and a local state transition occurs on the node, as the receiver becomes a sender. The consumer node becomes a receiver, and so on further down the technology transition food chain. On the other hand, if some percentage of the messages is not understood, the receiver asks for clarification in terms of feedback from the sender. The sender then sends clarification in response for the request for clarification.

Once the consumer understands the message, the consumer can execute the work products. Since a change agent becomes a sender, and the consumer becomes a receptor, each is capable of issuing requests for clarification and providing clarification.

This elemental system (Figure 5-2a) consists of a send unit and a receive unit. The receiver unit is able to retransmit or execute an action when there is little uncertainty in the terminal action to be taken. At that point, the receiver executes the action and becomes a send unit, since someone else (another potential receive unit) can witness the evidence of a signal. Let's assume for the moment a clear, noiseless signal from the sender. If the receive unit understands the encryption and protocol of the sender, it is able instantaneously to resend the message or to act. No effort is required to handle the encryption and protocol.

If the message received is well understood, the unit ***R*** (at time step t_k) can receive the messages from unit ***S*** (sent at time step t_{k-1}), immediately and resends or performs an action, observable as a message, to another (or the same) receiver at a later time step (t_{k+1}). Figure 5-2a shows this basic state transition model. Note, that there is also a term ***p'*** representing message state transition arcs from outside the local control volume. The message traffic *from* the receiver ***R*** is a sum of the messages from the earlier send unit and multiple streams persistent in history that are available to the receive node and selected (filtered) as input. The sum of the messages is available to be processed by node ***R***.

2. Entropy in the Communication Control Model

We can also have the case where there are messages with entropy (noise, or unknown signal) as input to ***R***. This can be accommodated as seen in Figure 5-2b. Now, we add the concept of a "think" state transition. This is the case where the messages received could not be effectively processed. Some internal processing is required. There is yet another type of "think" state transition. This is represented by feedback in order to clarify the entropy, noise or non-signal received. Figure 5-3 illustrates the elemental notion presented in Figure 5-2b and adds two feedback loop state transition arcs p_4 and p_5. For initial model development and clarity, we assume that the quantity of messages in the think loop p_3 are equivalent to the number of messages sent back to the send unit in p_4. These are subsequently fed to a receive unit as clarification at some later time step as p_5. It is possible that the send unit has to use multiple time steps and its own think loop. Further, it is possible that the receive unit has to do more internal processing (and learning) which could store, for more than one time step, a number of prior messages awaiting action. We want to avoid or minimize a design that has this characteristic. The system would appear to have slow response to transients, and the hysterisis effects resulting from these time step delays can put the node and system in an unstable mode of operation. While some of this effect is able, the model should be able to accommodate these aspects as well. These refinements can be added later.

The nodes can be in two states, x_k, y_k. The state represented by variable y_k is the quantity of messages or tasks orders that have been executed by an organizational unit, or node at time t_k. The state x_k is the quantity of messages / task orders received by the organization at time t_k. x_k consists of two parts. One is the quantity of messages / task orders that arrive from the outside the organizational node. The second part is the set of internal messages / task orders that must be processed/executed by the unit due to the content of the messages / task orders processed in the previous time step (feedback) t_{k-1}.

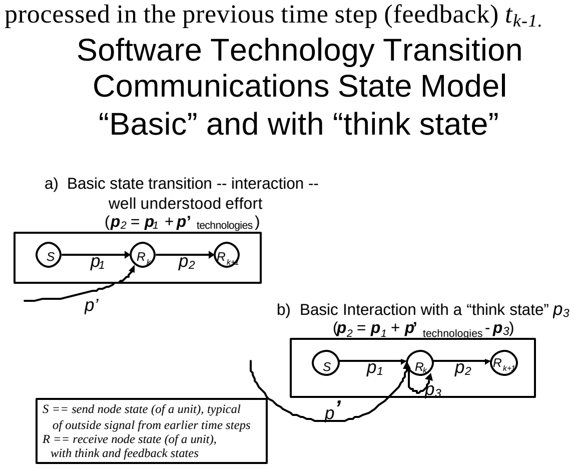


Figure 5-2. Software Technology Transition Basic and "Think" State.

(Source: Saboe 2001)

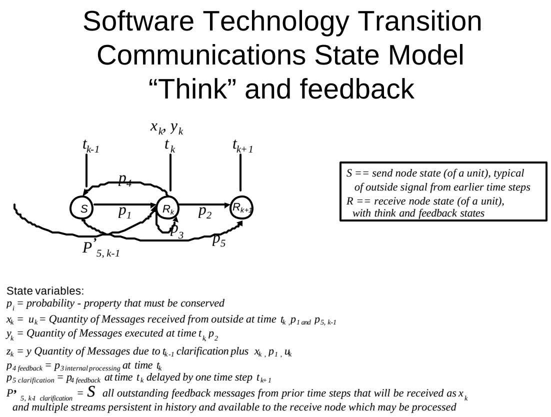


Figure 5-3. Software Technology Transition "think" and Feedback.

(Source: Saboe 2001)

On the other hand, let's assume that the receiver has to process some internal messages in order to unpack the message. Now there is a delay before the message can be resent. Going a little further, if the receiver received noise, an unclear signal, or unknown signal it may have to request clarification, delaying a time step or do some additional correction processing. This uses up node capacity. If the message is simple and concrete, or agrees in abstraction (state level) or is at a higher level meta-statement, the amount of processing and effort that it takes to correct the poor signal is less than one that is more complicated and more densely packed. From this, we might say that abstraction is a form of information hiding. Encapsulation of this form provides leverage and can reduce the "entropy" of the system.

6. DYNAMICAL SYSTEMS MODEL

Assume we have available a macro level model of technology transfer to represent the community level technology maturation. That macro model can identify the stability and convergence of an ensemble of nodes. The macro model can be partitioned into a number of nodes (organizational units and sub units that compose the organizational units). The macro model is represented in terms of entropy dimensions of natural measure (Farmer 1983) i.e. both the information entropy S_H and the bakers' transformation entropy S_B, representing the transfer (transform) function. We now would like to develop a model that represents the interaction between nodes at the micro level. This model will complete a linkage from macro to micro levels and permit implementation models (infusion, learning, etc) to bridge to the macro-micro infrastructure scale models. This section will explore a feedback model at the organizational node and sub-organizational node level we incorporate control theory and use the bakers' transformation.

The model should incorporate a factor for learning, and address requests for clarification and the ability to model the process load in requesting clarification messages and receiving clarification messages. This model will permit tuning an organization to ensure efficient processing of technology messages. We will develop a node response curve and associated system response curve. Determination of the bakers' transformation entropy from the Lyapunov number and exponent will permit an assessment of the node performance in terms of stability and confidence of convergence to a steady stable state, or chaotic state.

1. *Assumptions*

Assume nodes made up of people and machines that can do a task, such as publish a work product as a message. A node is modeled in terms of the messages it receives verses the messages it processes. The work product (message) is the representation of something that can be understood by communicating in terms familiar to the sender and receiver. Just as a map is not the road system but symbols from a vocabulary of terms, which represent a common understanding of the lay of the land of a road system. The terms are measured in information units – bits. As input, the processing node receives work product. These represent messages. Output from a node is also observed and measured in messages. A technology generating or processing node produces the output by acting on input to reduce uncertainty in the cause and effect relationship involved in achieving a desired result. This is reasonable since this is what elements of a node do. This is true for the activities of researchers, producers in general as advocates, or receivers, change-agent and consumer as receptors. This assumption is also consistent with the observation by Rogers (Rogers 1983). Within this context, we examine the meaning of the concepts of stability, equilibrium, attractors, chaos, eigenvalues, and eigenvectors, and the relationship to technology transition, system node dynamics. Convergence of an organizational node on a fixed point depends on the nature of the eigenvalues of the derivative of the dynamical system at the fixed point. The direction of convergence depends on the direction of the eigenvectors. We discuss seven cases elsewhere (Saboe 2002). A useful term that will frequently arise in the following discussion is an orbit. The orbit of a dynamical system is that sequence of points in the state-space phase plane that corresponds to successive time steps in the system.

2. Context

We assume that all of the nodes have functions of equivalent form. Rather than model the efficiency in a node, the efficiency is built into the tuning parameter b. This performance index can be attributed to nodes in a performance band (See Saboe 2002 *TechTx Learning Curve Model)*. The nodes, in different performance bands, inherit the performance parameters of their band. The node is modeled in terms of the messages it receives versus those it carries out or processes. The individual nodes are assumed heterogeneous, varying in size and composition, or a mix of people with varying skills and tools to perform the function. For ease in validation computations, we assume the organizational nodes that have say a performance index in the range of +/- 1σ of the mean, all would have the same efficiency (learning curve function parameters). Should we wish to calibrate an individual node or all of the nodes in the band the model will still be applicable. The capacity of a node in the band can be calculated. The amount and complexity of the message flow acted on, and generated, applies pressure to an organizational node. Demands on the organizational node as sender or receiver components are among the pressures that require modeling and analysis. But we do know this intensive property from the macroscopic analysis. Other pressures are internal to an organizational node to ensure smooth functioning. These internal pressures come in the form of messages as well, procedures, interfaces, meetings, collaborations and other interactions that consume resources. These are important facets to model since they provide feedback pressures on the components. External pressures are also among the features that determines organizational node dynamics and should be modeled.

All of the pressures mentioned so far can be thought of as messages passing between organizational nodes and between the organizational nodes and the environment. This concept facilitates modeling organizational node states that can be organized as messages received by the component and processed by a component. In this respect, the organizational nodes are analogous to a communications network. The analog is simple and useful. There are however at least two important differences. One is that an organization will adapt to and absorb pressures that would cause a network to breakdown. This is because the network is not hardwired. It is also difficult to predict the breakdown capacity in advance. We could address ranges of capacity by banding the organization into performance index bands. This does not however mean that a node is at capacity. The potential for the technology transfer system to breakdown is important to model. A simple source of collapse is when the demands on the system exceed its ability to adapt and the node reaches a state of demoralization. We have mechanisms to model this, however for purposes of illustrating the model, are at or below capacity. We can ignore for now this breakdown at over capacity issue.

The model for organizational dynamics is drawn from (Brown 2000). This model can be represented in state space using messages (*N*). This is closely related to the entropy as we saw earlier. The state space is mapped onto the x-axis (input) and y-axis (output) as follows: *x*, the input N_k, in messages, and the output in *y*, N_{k+1}, where k represents the time step (the internal time as an *operator*, and not a number). Where k represents internal time iterations as discussed earlier. We would not have synchronous discrete time steps in a network that includes nodes comprised of organizations and people.

$$N_{k+1} = f(N_k) \qquad (6.1)$$

This function represents the bakers' transformation. For the ensemble of nodes performing the function $N_{k+1} = f(N_k)$, we have the vector (capital *F)* representation $N_{k+1} = F(N_k)$

We narrow our discussion from the ensemble of nodes that can appear on the network or disappear to a typical group of nodes, the sender, receiver and consumer.

The model uses two state variables for the system node representing the messages received and messages processed. We shall apply the message information in terms of the number of messages of the incoming and processed messages. The significance of the system of equations is that the eigenfunction characteristic equation represents the bakers' transformation of folding, stretching and rotating. The eigenvalue of this dissipative function is also entropy and represents mixing. Elsewhere (Saboe 2002), the potential significance of the values of the eigenfunction are discussed.

Now we will develop the equations for this model. The relationship between the state transition diagram and a dynamical system is shown in **Figure 6-1**.

The sender publishes messages u_k (a number of messages) at time step k. Input messages at time step k to the receiver are indicated by x_k (a number of messages). The output messages from the receiver is at time step k is given by y_k. (a number of messages) Some percentage of the messages output from a prior time step y_{k-1}, are indicated by b, a rational number.

This process is repeated for the next time step x_{k+1} and y_{k+1}. The crossed circle immediately to the left of the receiver node represents the collection point where the different parts of the input message stream is combined for the input message count x_k. In **Figure 6-1**, $f(x_k)$ represents the function to transform the input messages into output messages. It takes a time step to complete the processing. A way to view the nodes processing is that for a message to move through a node, it takes a time step.

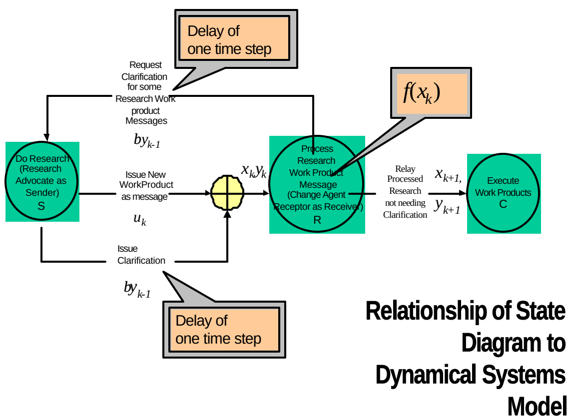


Figure 6-1. Dynamical Systems Model.

The x_k state variable consists of two parts. One part of the state is the messages that come from outside the receiver node u_k. The second part of the state variable is clarification of messages that was requested from the previous time step y_{k-1}. Initially we assume that the quantity of messages processed (y_k) is a function of x_k. This function has the following simple properties:

(1) if $x_k = 0$ then $y_{k+1} = 0$
and
(2) as $x_k \to \infty$ then $y_k \to 0$

Condition (1) is obvious – no input, then no output. Condition (2) says that the system grinds to a halt if the message demand is too great. We can assume that as the number of messages received becomes infinite, the messages processed have to approach some limiting value, which is the capacity of the system. The system can be represented by the following equations. However, for the systems we are seeing, we are not at capacity, and this condition can be finessed out of the picture in low pressure, low temperature situations. We can determine when this happens by partitioning the macroscopic community into smaller and smaller partitions. Then we can observe the performance of nodes with a technology and in the environment of the day.

$$\begin{aligned} x_{k+1} &= b\,y_{k-1} + u_k \\ y_{k+1} &= f(x_k) \end{aligned} \qquad (6.2)$$

$f(x_k)$ is called the node response curve. Elsewhere (Saboe 2002) the important relations between the node response curve and the system response curve are developed. We need only concern ourselves, for this exposition, on the node response curve and its ultimate relationship to the macroscopic information theoretic model.

The above is a second order system of finite difference equations with the response curve $y_{k+1} = f(x_k)$ represented by the following three-dimensional dynamical system. Where z_k, clarification from the prior time step, is substituted for y_{k-1} and using the mapping (Kreyszig 1993 p419) we get

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \\ z_{k+1} \end{pmatrix} = \mathbf{T} \begin{pmatrix} x_k \\ y_k \\ z_k \end{pmatrix} = \begin{pmatrix} b\,z_k + u_k \\ f(x_k) \\ y_k \end{pmatrix} \qquad \mathbf{(6.3)}$$

Let's consider the time step $k+1$. The periodic points determine the dynamics of the system. In particular, the fixed points are of interest. These are the equilibriumpoints. The coordinates of the fixed points X, Y, Z are given by

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} b\,z_k + u_k \\ f(x_k) \\ y_k \end{pmatrix} \qquad (6.4)$$

The fixed point condition becomes $x=u+bf(x)$.

The derivative of the transformation T is given by the Jacobian

$$J = \frac{\partial(X,Y,Z)}{\partial(x,y,z)} \quad (6.5)$$

$$= \begin{pmatrix} \frac{\partial X}{\partial x} & \frac{\partial X}{\partial y} & \frac{\partial X}{\partial z} \\ \frac{\partial Y}{\partial x} & \frac{\partial Y}{\partial y} & \frac{\partial Y}{\partial z} \\ \frac{\partial Z}{\partial x} & \frac{\partial Z}{\partial y} & \frac{\partial Z}{\partial z} \end{pmatrix} \quad (6.6)$$

$$DT = \begin{pmatrix} 0 & 0 & b \\ f'(x) & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad (6.7)$$

Where ***D*T** is the Jacobean of transformation **T.** Find the eigenvalues j_i which are the roots of the characteristic equation:

$$|A - jI| = 0 \quad (6.8)$$

Where A is the Jacobian of the DT transformation, and I is the identity matrix. More specifically the determinant

$$|DT - jI| \quad (6.9)$$

is the characteristic equation when set equal to zero.

$$-j^3 + b\, f'(x) = 0 \quad (6.10)$$

There are three eigenvalues for the solutions of the equation $j^3 = b\, f'(x)$. There are two complex conjugate eigenvalues and one real eigenvalue. The three eigenvalues may be represented as

$$j_1, j_2 e^{(2p\,i/3)}, j_3 e^{(-2p\,i/3)}$$

where (6.11)

$$j_i = (b\, f'(x))^{1/3}$$

is the real root of the equation. From the model, we conclude that $\left|j_i^3\right| < b\left|f'(x)\right|$. The system is stable when $\left|j_i\right| < 1$, in equilibrium when the norm is $\left|j_i\right| = 1$, and unstable when $\left|j_i\right| > 1$ (Farmer 1983, Baker 1990, Brown 2000).

This gives some insight into the structural stability aspect. The control theory element of the current research model addresses mixing, and structural changes due to feedback from external nodes. The value of the norm (<1, =1, >1, real imaginary, etc) of the eigenfunction characteristic equation assimilation of reality based on experiences from prior time steps.

From this, we see that for small enough b or large enough u_0 we can achieve stability. For the technology transition system, we desire stability and convergence. With a stable model at the organizational level, we have organization nodes, which are not thrashing or wasting effort. With stable nodes, we can build a stable infrastructure composed of those nodes. This will also yield convergence of the technology.

The data that we can measure is the number of messages published at some time k. We can also measure output $y_{k+2,\, k+1, k,\, k-1,\, k-2}$. The output message data is simply the offset published by a time step e.g. $u_{(t-c)}$. The difficulty we have is, that the macro data to empirically support $f'(x) = \frac{d\mathbf{Y}}{d\mathbf{X}}$ cannot be arrived at directly.

The derivative $f'(x)$ can be obtained using parametric differentiation. Our system curve from empirical data is the output y, which represents u offset by an interval c from a prior time step. Initially, for the data examined, this interval was one year. In effect, this provides an immediate memory for chunking of three registers because it take three time steps to clear all of a message when there is a request for clarification.

As this immediate memory, represented in time steps is expanded, the error from the modeled to predicted should start to diminish. Therefore:

$$\mathbf{Y} = y_t \equiv u_{(t-c)} \quad (6.12)$$

$$\mathbf{X} = b\, f(x) + u_t = b u_{t-c} + u_t \quad (6.13)$$

Now deriving from (6.12) and (6.13) we get

$$f'(x) = \frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{d\mathbf{Y}/dt}{d\mathbf{X}/dt} \quad (6.14)$$

The following result was obtained using parametric differentiation of (6.14) and substituting (6.12) and (6.13)

$$f'(x) = \frac{u'_{(t-c)}}{b\, u'_{(t-c)} + u'_{(t)}} \quad (6.15)$$

We can substitute $f'(x)$ into (6.11) which defines the real eigenvalue:

$$j = \left(b \frac{u'_{(t-c)}}{b u'_{(t-c)} + u'_{(t)}} \right)^{1/3} \qquad (6.16)$$

or explicitly to enable programming from the data sets

$$j_i = \left(b \frac{\frac{du_{(t-c)}}{dt}}{b \frac{du_{(t-c)}}{dt} + \frac{du_{(t)}}{dt}} \right)^{1/3} \qquad (6.17)$$

The point where the graph intersects the line $y=x$ is the equilibrium point. The slope of $y=u+bf(x)$ at the fixed point is the real eigenvalue of the matrix DT(X_0). By changing the parameter b we change the shape of the graph and thus we change the slope where the fixed point is found. Also by changing u_0, we change the location of the fixed point along the horizontal axis and thus the eigenvalue. By starting u_0 at 0, we first have a fixed point whose real eigenvalue is positive and less than 1. This is ideal in that it indicates that the system will converge to a point where it remains stable and made sense as the review of the various characteristics of the eigenvalue were developed. See these graphs and the various interpretations of their meaning in the appendix.

For the moment, let's go back to the model consisting of sender, receiver and consumer, Figure 5-2. Now let's focus in on the receiver, look at the inputs, and outputs of this node. It turns out that any of these nodes looks like a receiver in the general sense. The sender can also be picking up new messages from others, in which case the sender acts like a receiver. The sender can also be requesting clarification and be receiving clarification in the same manner as the receiver. Likewise the consumer gets input and outputs. So our model can be seen in **Figure 6-2** to have all of the features but represented only in a single node, the receiver. We also have a useful sign convention. All of the inputs to the node are positive and outputs are negative.

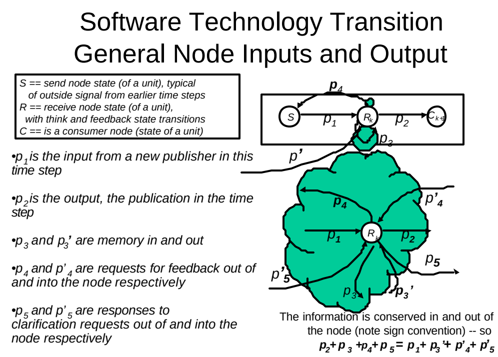


Figure 6-2. General Node Inputs and Outputs.

We are now in a position to think of an ensemble of nodes. Essentially a distribution of these nodes performing the bakers' transformation. Just like a physical system or communication system, we now can speak of a macro stochastic process in terms of entropy and information.

Let's go to the basic equation (4.27). Recall our conserved property is messages N – information in our case. Using Shannon's entropy S_H and N for the number of messages we get

$$\frac{1}{T_{S_H}} = \frac{\partial S_H}{\partial N} \qquad (6.18)$$

From the section dealing with the information theoretic aspects of Interacting Subsystems, we saw how T varied with a time step. Now we can observe the control model, S_B as a function of the same time steps. We have the opportunity to relate the two entropy measures, S_H and S_B since they are related to the same information system of messages N. We are dealing with the same information flows, hence the same system, so this seems reasonable. Recall S_B is related to Lyapunov's exponent λ, which comes from the eigenvalue j.

We found the relationship of messages verse time step in Figure 4-5 was very satisfactorily modeled as a linear equation for this technology set. (It could be different for other technologies, this is why we have dealt with the relationships in terms of functions, eigenvalues and derivatives.) In this case, the derivative of the linear model reduced to a constant in equation (4.40).

Initially, to determine the *form* of the functions, the average value *b»10%* was used. This was done by iterative guesses of a fixed *b*. This approximation of *b* was used to satisfy the macroscopic rate of change of entropy. The curves can be seen in Figure 6-3. By observation, we see both the entropy measures as a function of time step. S_H,, the information theory entropy measure is on the left y axis, and the S_B which comes from the eigenvalue of the micro control model (hence in the range of 0 to 1), is on the right hand y axis. Both curves are of the same form. They follow the power law, in this case, with a good R^2. This suggests that the model does approximate the observed conditions.

Entropy S_H and Dynamical System j

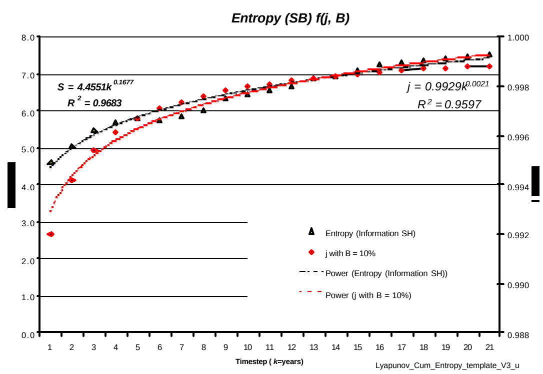


Figure 6-3 Macro Equilibrium S_H and Eigenvalue l Stabilization

At this point we are considering the "community" a large node. In the real world, the community is partitioned into a volume of performing nodes, and different performance rates. Elsewhere (Saboe 2002) methods of relating these partitions to learning curves are discussed. The learning curve implies **b** varies with tasks performed, and this in turn can be shown in terms of time steps.

We can also comment a bit further on the relationship of the two entropies. In order to do this, we have to develop S_B a bit more since it is based on the real eigenvalue λ.

$$\begin{cases} t \equiv \{term\} \\ 2^t = \{msg\} \end{cases} \qquad (6.19)$$

Where 2^τ is a family of all the subsets of set τ, often called the power set. Here is an example.

$$t = \{A, B, C, D\} \qquad (6.20)$$

$$2^t \equiv \left\{ \begin{array}{l} \{\}, \\ \{A\},\{B\},\{C\},\{D\}, \\ \{A,B\},\{A,C\},\{A,D\},\{B,C\},\{B,D\},\{C,D\}, \\ \{A,B,C\},\{A,B,D\},\{B,C,D\},\{A,C,D\}, \\ \{A,B,C,D\} \end{array} \right\}$$

(1.21)

Now when the number of elements in $|\tau| = 4$, we get $2^{|t|} = 2^4 = 16$. Note also the distribution of sets. This makes available the partition function (the most useful formula in statistical mechanics.) A number of useful relationships can be developed from the set distributions.

Recall that $S_H \equiv -\sum p \log_2 p$ from (4.4)

Let's define the operators in the transformation of the control model.

$$\mathrm{B} \equiv \{ b, f(\bullet), g(\bullet) \} \qquad (6.22)$$

where $b, f(\bullet)_0, g(\bullet)_0$ define a transformation S_B that works on **B.**

$\mathrm{B} \equiv \{ (b, f) \}$ which is a set of operator combinations. Simplified we have

$$S_B \equiv \sum j(b, f) \qquad (6.23)$$

$$j_0 \equiv \text{real eigenvalue of} \begin{pmatrix} 0 & 0 & b \\ f' & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \qquad (6.24)$$

where

$$j_0 = \sqrt[3]{b\, f'} \qquad (6.25)$$

We sum the λs over the number of send units in the sample that publish.

We are now in the position to relate the two entropies S_H and S_B. It also makes intuitive sense. The idea that Kolmogorov had is there are objects and there are descriptions (encodings) of objects, and the complexity of an object is the minimal size of this description. If we have one publisher, and the publisher encodes a message, we can sum all of the publishers and messages (a countable number) and say some real things about the ensemble of messages (objects) and publishers (elemental control model nodes). This can be represented by a program (this is a process) in a finite length for the nodes and messages generated. That is there is a "decidable" partial ordering defined on the set. The term "decidable" means there is an algorithm for y' and y'', whether $y' \leq y''$ or not.

On an intuitive level, (per Uspensky) the elements of a "space can be taken as informations, and $y' \leq y''$ means that the information y'' is a refinement of the information y' (and hence y'' is closer to some limit value to which both y' and y''

serve as approximations.” This even sounds like technology maturation.

Now let’s take these two entropy approaches (plotted on the left hand y-axis and now coincident) and determine *b*. *b*, as we recall represents the amount of feedback required and is plotted on the right hand y-axis. In Figure 6-4, we see that initially the feedback required is very high, which then bottoms out, and then increases. As shown in Figure 6-4, we are asking for the amount of feedback of the system, this technology’s world, required for the entire ensemble of publishers. If we adjust for the number of authors and as shown in Figure 6-5, we can see the effect of “mindshare”. In this figure, *b* per author is the asymptotic curve plotted against the primary y-axis. Publishers, i.e. nodes, which represent the volume is the increasing curve plotted against the secondary y-axis. We are now representing the system on a per node basis. This is the now in terms of messages and feedback per unit volume (nodes) an extensive variable.

We have more authors absorbing a smaller and smaller fraction of the information persistent in the world. In other words, more information means more decisions (the number of decisions are directly related to entropy), and there is a limit to how many decisions a node can make per unit time.

These two figures’ curves also seem to provide some hint as to where innovators, and early majority (the advancement phases of the technology) shift to early and late majority (the diffusion phase of the technology’s life).

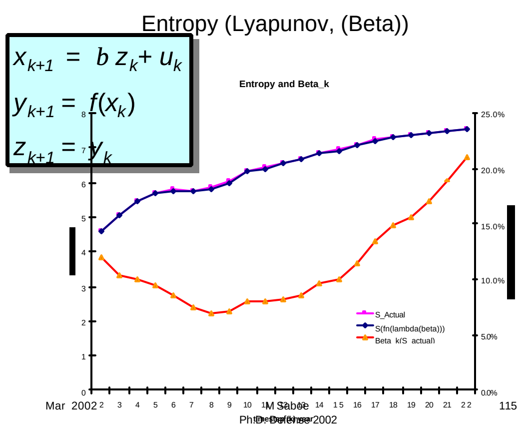


Figure 6-4 *b*, **feedback required varies with time**

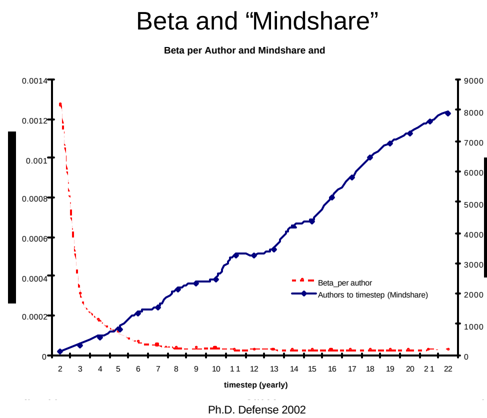


Figure 6-5 *b* **adjusted for “mindshare”**

7. A “TECHNOLOGY TRANSFER ENGINE”

Now that we have established the basic relationships of the *TechTx Entropy* models, lets put it in the framework of a system. We can put it all together as an evolutionary, technology transfer system that has probabilistic effects at the macro level and deterministic, dynamical effects at the microstate level. We have to the tools to analyze a program and represent it as an engine. These technology transition dynamics tools permit us to engineer a solution to get maximum efficiency out of our resources.

It is useful to define a control volume that is typical of the system. In a traditional continuous system in a physical world, a control volume identifies boundaries of the system. In such a continuous system, say an engine, a mass flows a distance and contributes to the work performed. It is not unusual to partition up a continuous control volume into stages, e.g. a compressor, a combustor, a diffuser. As the mass (in our case a message *n)* flows from stage to stage we can consider its macro state property transition at the system and locally of the nodes (compressor, combustor, diffuser). There are *n* messages flowing each one unique, but we can characterize the state of an ensemble, so the system and the nodes take on different state properties at different stages of the process for the complete elaboration of the message-node state combinations. For now let’s look at all three stages in terms of the technology transfer dynamics message model with the message moving through the control volume. This causes both local and system level state transitions.

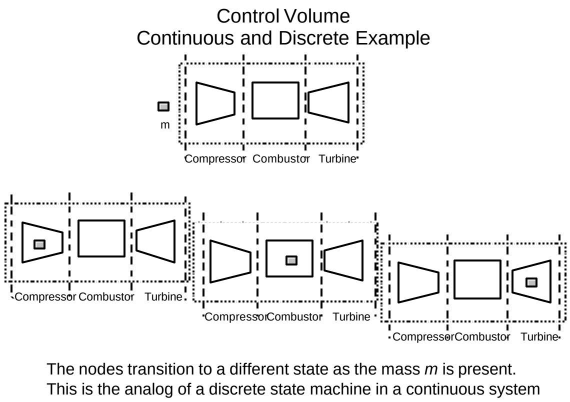


Figure 7-1 Illustration of a Control Volume -- a Continuous System or as a Discrete State Machine

Similarly, in this discrete state machine, we have drawn the boundaries around the three subsystems. Full elaboration of all of the messages *(n)* states within the control volume would represent all the possible states of the bounded system. With this we can represent an individual interaction, an organizational interaction or even a macro technology transfer system such as the economy and analyze and prescribe using cycle diagrams like **Figure 7-2**. This type of diagram is familiar to most engineers.

Let's examine some of the state diagrams and system quantities in **Figure 7-2**.

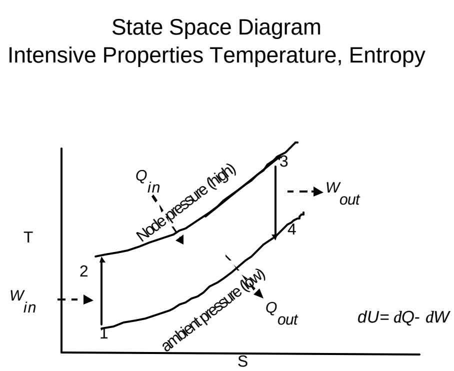


Figure 7-2 Cycle Diagram

One could image that the compressor, which moves a technology from the null or ambient state (with focused research as in academic and basic research facilities.) This is the sate transition process path from 1-2.

Here we have a process depicted in macro state space that originates at point 1 with T_{lo},P_{lo},S_{lo} which are the ambient temperature and pressure of the surroundings, a reservoir. In the normal sense of use, we see this as work, energy or heat. In technology transfer dynamics, we can think of this as effort, which is added to the system, yielding "energetic" messages. We see an *isentropic*, (constant entropy) compression as the system moves along the path 12 to T_2,P_{hi},S_{lo}. This says the temperature is increasing because some effort is being done to reduce the volume, or increase the messages in the same volume of performing nodes in which the interaction between entities occurs.

More occurrences of terms consistently show up in messages. Terms are combined to get to concepts that are more powerful. While there are more messages for the same volume, i.e. nodes, the message term content has higher density. In the model proposed, there would be fewer nodes, but doing very intense research, i.e. producing much high quality messages. They closely interact and publish messages generally within the confines of the system.

During the progression form state point *2* to *3*, energy in the form of effort is added at a constant pressure. Entropy, S_{lo} increases to S_{hi}. Think of this as a demonstration. No new basic research is being performed, the science is being scaled up and loaded with a lot of energy that will make it attractive to consumers. This occurs when the technology is *diffused* from state *3* to *4.* A high pressure, concentrated set of messages escapes into a larger volume to get a drop back to ambient. In order for this to happen the message entities must some how move to a bigger volume, must some how escape. This is where work is taken out, as products are delivered to a market (ambient). This is shown as a constant entropy line, which a rapid drop from T_{hi},P_{hi},S_{hi}, at state point *3*, to state point *4*, T_4,P_{lo},S_{hi}. Work, in thermodynamic terms, is represented by extensive property rate changes. For example,

$$W = \dot{n}C_p(T_3 - T_4) \qquad (7.1)$$

Where *W* is work yield, is the result of some stimuli as we saw in (4.31),and $\dot{n}$ is the message flow rate in terms of messages per time step. This is a very ideal cycle description that would be a typical start point to determine the minimum nodes (authors) and messages to be managed to achieve a technology transition objective. Recall, we are developing the basis for an engineering model. Reality will demand that we can't achieve a constant entropy line from P1-P2. It will likely follow a polytropic path. Further, efficiencies of the components must be addressed. Finally, we recognize that people, and organizations are not machines. We would have to calibrate and experiment with known components to determine component (node) efficiency *h*, and failure rates (burnout or collapse of morale from over work.

8. VALUE AND USAGE

The research tied together fundamental elements underlying technology transition. Currently, systematic techniques for assessing macro mechanisms for transferring software engineering technologies has been thoroughly reviewed and systematized. This paper developed the fundamental elements of an *industrial model* of a *software technology transition engine*. The mechanisms were developed utilizing information theory, communication theory, chaos control theory, and learning curve principles. The combination of those scientifically sound mechanisms provides a basis for assessing, and / or prescribing a portfolio of technologies and the implementing macro infrastructure. A program manager armed with a simple browser interface, augmented and driven by these algorithms, should be able establish the answers to questions similar to those posed in Figure 8-1. Linkages to lower level models and implementation methods are provided. This research provides the *theoretical framework, and an entropy based engineering model, for a practical method for a program manager to establish a high capacity transition channel, which accelerates technology maturation and insertion*.

This model can be used in a prescriptive manner. From a large research investment portfolio manager's perspective, e.g. the Department of Defense Science and Technology base, this enables decisions to be made as to the number of nodes (researchers) which should be funded. It also provides a basis for policy development and decisions using quantitative approaches in concert with the current qualitative assessments.

The model was validated (Saboe 2002) on data samples assess the following technologies: software engineering, software technology transfer, Ada, Java, abstract data types, rate monotonic analysis, cost models, software work breakdown structures. Also included in Saboe 2002 is an extensive annotated bibliography on software technology transfer and related references, and a bibliography including related material from philosophy, psychology, math, physics, thermodynamics, management, economics, game theory, technology transfer, software engineering, and systems engineering. The appendix of this paper illustrates a number of signatures of systems that may be found in various technology's studies.

A broad area of future research is now open for examination. The application of the fundamental tool set, and the definition of *information temperature*, provides a robust capability to describe, develop, or analyze a software technology transition engine. Further, the model was developed to accommodate any evolutionary process. One could conceivably apply this to the software development process. Finally, it seems within the realm of imagination that software itself could be modeled using the relationships presented herein.

Now, it is left to the community to determine whether this is satisfactory to open the discussion that supports the following logic:

- since we should be able to accept that a process is just a program (Osterweil 1987) and
- software can represent the program, and
- the engine is the representation of a process that was based on axiomatic and logical transfers from established science and engineering (physics and thermodynamics)
- The basic elements of the ***physics of software*** have been developed.

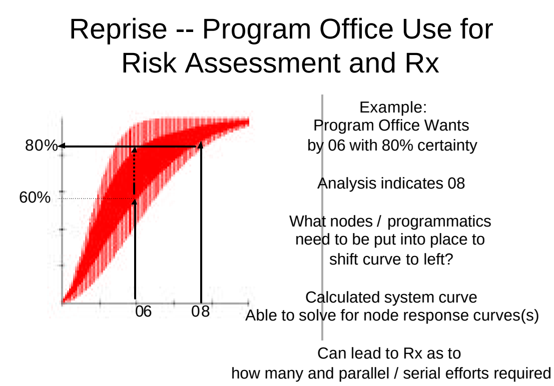


Figure 8-1 Usage of the Model when implemented in a tool

Michael S. Saboe is an Associate Director at the US Army Tank Automotive Research, Development and Engineering Center in Warren, Michigan. He runs the US Army Next Generation Software Engineering Technology Area where he has responsibility for the US Army Main Battle Tank Software sustainment activities. He has nearly 30 years of software experience that began with real time software efforts in 1973 on gas turbine engine control systems. He was the chair of the DoD Joint Logistics Commanders Computer Resource Management group, and has been a DARPA PI since the 1980's. Saboe received his Ph.D. in Software Engineering at the Naval Postgraduate School in Monterey CA. He is a professor at the Naval Postgraduate School as well.

References:

(AAAS 1972) Liebig and After Liebig: A Century of Progress in Agricultural Chemistry. American Association for the Advancement of Science, Bulletin 16, 1972.

(Abbott 1989) Abbott, Michael M. and Nan Ness, Hendrick, *Theory and Problems of Thermodynamics*, McGraw-Hill, New York, 1989.

(Allen 1977) Allen, Thomas J., *Managing the Flow of Technology*, MIT Press, Cambridge, MA, 1977.

(Allen 1983) Allen, Thomas J., Diane B. Hyman, and David L. Pinckney, "Transferring Technology to the Small Manufacturing Firm: A Study of Technology Transfer in Three Countries," *Research Policy*, 12, pp. 199-211, 1983.

(Anderson 1981) Anderson, John R., ed., *Cognitive Skills and Their Acquisition*, Lawrence Erlbaum Associates, Hillsdale, N.J., 1981.

(Baker 1990) Baker, Gregory L. and Gollub, Jerry B., *Chaotic Dynamics and Introduction*, Cambridge University Press, 1990.

(Bayes 1763) Bayes, Thomas, Rev., "Studies in the History of Probability and Statistics: IX, Thomas Bayes's Essay Towards Solving a Problem in the Doctrine of Chances", Barnard, G. A, *Biometrika*, Vol 45, Issue 3 / 4 Dec. 1958, pp.293-315.

(Boehm 1989) Boehm, Barry, and Ross, R., "Theory W Software Project Management: Principles and Examples," *IEEE Transactions of Software Engineering*, July 1989, pp. 902-916.

(Boehm 1998) Boehm, Barry, Egyed, A., Kwan, J., Port, D., and Shah, A., "Using the WinWin Spiral Model: A Case Study," *IEEE Computer*, July 1998, pp. 33-44.

(Boehm 1999) Boehm, Barry B. TBD.

(Boehm 2000) Boehm, Barry and Victor Basili, "Gaining Intellectual Control Of Software Development", *IEEE Computer*, May 2000, pp. 27-33.

(Brown 2000) Brown, Ray, "Private Communication," *Toward a Control Theory for C^3* Systems, Research Fellow, Raytheon Systems Company, Falls Church, VA 22042, September 2000, through May 2001.

(Buxton 1991) Buxton, J.N. and Malcolm, R., *Software Technology Transfer*, 17-23.

(Cover 1991) Cover, Thomas M. and Thomas, Joy A., *Elements of Information Theory*, John Wiley and Sons, Inc., N.Y., Chichester, Brisbane, Toronto, Singapore, 1991.

(Eveland 1990) Eveland, J.E. and Tornatzky, L.G., "The Deployment of Technology," in *The Processes of Technological Innovation*, L.G. Tornatzky and M. Fleischer , editors, pp. 117-148, Lexington Books, Lexington, Massachusetts:, 1990.

(Farmer 1983) Farmer, J. Doyne, "The Dimension of Chaotic Attractors," *Physica* 7D, pp. 153-180, North-Holland Publishing Co., 1983.

(Fichman 1993) Fichman, Robert G. and Kemerer, Christ F. "Adoption of Software Engineering Process Innovations: The Case of Object Orientation," *Sloan Management Review*, winter, pp. 7-22, 1993.

(Fichman 1994) Fichman-RG; Kemerer,-CF, "Toward a theory of the adoption and diffusion of software process innovations," in *Diffusion, Transfer and Implementation of Information Technology*, in Levine, Linda, ed., proceedings of the IFIP TC8 Working Conference on Diffusion, Transfer and Implementation of Information Technology, Software Engineering Institute, Carnegie Mellon Institute, Pittsburgh, PA, North Holland, Amsterdam,, London, New York, Tokyo, 1994.

(Forrester 2000) DoD Evolutionary Acquisition Workshop Sep 2000.

(Fowler 1990) Fowler, P., "Technology Transfer As Collaboration: The Receptor Group," *Proceedings of the 12th International Conference on Software Engineering* pp. 332-333, IEEE Computer Society Press, Nice, France, U.S., 1990.

(Fowler 1992) Fowler, P. & Levine, L.,"Toward A Problem Solving Approach To Software Technology Transition," in J. Van Lemwen (Ed.), *Proceedings of the IFIP 12th World Computer Congress*, vol. ,pp. 57-64,. Madrid, Spain, The Netherlands: North Holland, Elsevier Science Publishers, 1992.

(Fowler 1992A)Fowler, P. & Maher, J., "Foundations For Systematic Software Technology Transition," *Software Engineering Institute Technical Review* '92, pp. 1-32, 1992.

(Fowler 1994)Fowler, Pricilla and Levine, L., "From Theory to Practice: Technology Transition at the SEI," *IEEE Proceedings of the Twenty-Seventh Annual Hawaii International Conference on System Sciences*, pp. 483-497, 1994.

(Fraundorf 2000) Fraundorf, P., "Heat Capacity in Bits," xxx.LANL.gov/cond-mat/9711074 v2, October 1, 1999, rev. October 30, 2000.

(Gulliksen 1934) Gulliksen, H., "A Rational Equation of the Learning Curve Based on Thorndike's Law of Effect," *Journal of General Psychology*, 11 pp. 395-434, 1934.

(Hanakawa 1998) Hanakawa, Noriko; Morisaki, Syuji; Matsumoto, Ken-ichi, "A Learning Curve Based Simulation Model for Software Development," pp. 350-359, Nara Institute of Science and Technology, IEEE, 1998.

(Hargadon 1997) Hargadon, Andrew and Sutton, Robert I., *Technology Brokering and Innovation in a Product Development Firm*, Cornell University, 1997.

(Jaakkola 1995) Jaakkola, Hannu, :Comparison and Analysis of Diffusion Models", p 65-82, 1995.

(Jaynes 1957) Jaynes, E.T., “Information Theory and Statistical Mechanics,” *Physical Review*, Vol. 106, Number 4: pp.620-630, May 15, 1957.

(Knecht 1974) Knecht, G.R., “Costing, Technological Growth and Generalized Learning Curves,” *Operations Research Quarterly*, vol. 25, no.3, pp. 487-491, 1974.

(Kwon 1987) Kwon, T. H. and Zmud, R.W., "Unifying the Fragmented Models of Information Systems Implementation" in *Critical Issues in Information Systems Research,* eds. J.R. Boland and R. Nirshheim, John Wiley & Sons, New York , 1987.

(Langley 1981) Langley, Pat and Simon, Herbert A., “The Central Role of Learning in Cognition,” in Anderson, John R, ed., *Cognitive Skills and Their Acquisition*, Lawrence Erlbaum Associates, Hillsdale, N.J., 1981.

(Leagans 1979) Leagans, J. Paul, *Adoption for Modern Agricultural Technology by Small Farm Operators: An Interdisciplinary Model for Researchers and Strategy Builders,* CIAM 69, NTIS PB82-234527, Cornell University Department of Education, Ithaca, New York, 1979.

(Leonard-Barton 1988) Leonard-Barton, D., "Implementation Characteristics of Organizational Innovations," *Communication Research* 15, pp. 603-631, 1988.

(Lewis 1981) Lewis, Clayton , “Skill in Algebra,” in Anderson, John R, ed., *Cognitive Skills and Their Acquisition*, Lawrence Erlbaum Associates, Hillsdale, N.J., 1981.

(Li 1993) Li, Ming and Vitanyi, Paul, *An Introduction to Kolmogorov Complexity and its Applications*, ISBN 0387-94053-7, Springer-Verlog, New York, 1993.

(Mandlebrot 1953) Mandlebrot, Benoit, “An Informational Theory of the Statistical Structure of Language,” in *Communication Theory,* (ed. By Willis Jackson), Butterworths, 1953.

(Mazur 1978) Mazur, J. & Hastie, R., “Leaarning As Accumulation: A Reexamination of the Learning Curve,” *Psychological Bulletin*, 85 (6), pp. 1256-1274, 1978.

(McCauley 1993) McCauley, Joseph L., *Chaos, dynamics, and fractals, An algorithmic approach to deterministic chaos*, Cambridge University Press, 1993.

(Miller 1956) Miller G.A. “The Magic Number Seven Plus Or Minus, Two: Some Limits On Our Capacity For Processing Information”, *Psychological Review*, 63, pp. 81-97, 1956.

(Moore 1987) Moore, G. C., "End-User Computing and Office Automation: A Diffusion of Innovations Perspective," *Infor* 25 pp. 214-235 , 1987.

(Moore 1991) Moore, G., *Crossing the Chasm: Marketing And Selling Technology Products To Mainstream Customers,* Harper Business, 1991.

(Munster 1969) Munster, Arnold, *Statistical Thermodynamics*, Vol. I, Springer-Verlag, Berlin, 1969.

(Nembhard 2000) Nembhard, David A., “An Individual-Based Description of Learning within an Organization,” in *IEEE Transactions on Engineering Management*, Vol 47, No. 3, August, 2000.

(Newell 1981) Newell, Allen and Rosenbloom, Paul S., “Mechanisms of Skill Acquisition and the Law of Practice,” in Anderson, John R., ed., *Cognitive Skills and Their Acquisition*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1981.

(Newton 1726), Newton, Isaac, The Principia, Mathematical Principles of Natural Philosophy, translated by I. Bernard Cohen and Anne Whitman, University of California Press, Berkeley, 1999.

(Nishiyama 2000) Nishyama, Tetsuto; Ikeda, Kunihhiko; Niwa, Toru, “Techhnology Transfer Macro-Process, A Practical Guide for the Effective Introduction of Technology,” *International Conference of Software Engineering*, Limerick, Ireland, ACM, pp. 577-586, 2000.

(O’Marchu 1997) O’Marchu, Diarmuid, *Quantum Theology*, Crossroads Publishing, New York, 1997.

(Osterweil 1987) Osterweil, Leon, “Software Processes are Software Too,” Proceedings, 9th *International Conference on Software Engineering*, pp. 2-13, 1987.

(Pennings 1987) Pennings, J.M., "Technological Innovations in Manufacturing," in *New Technology as Organizational Change*, eds. J.M. Pennings and A. Buitendam,, pp. 197-216, Ballinger., Cambridge, Massachusetts, 1987.

(Pfleeger 1999) Pfleeger, S.L., “Understanding and Improving Technology Transfer in Software Engineering,” *The Journal of Systems and Software* 47, pp. 111-124, 1999.

(Piaget 1963) Piaget, J., *The Origins of Intelligence in Children*, WW Norton, New York, 1963.

(Piaget 1967) Piaget, J., *Six Psychological Studies*, Random House, New York, 1967.

(Piaget 1977) Piaget, J., *The Development of Thought: Equilibration of Cognitive Structures*, Vicking Press, New York, 1977.

(Planes 2002) Planes, Antoni and Vives, Eduard, “Entropic Formulation of Statistical Methods,” in Journal of Statistical Physics, Vol. 106, Nos. 314, February 2002.

(Prigogine 1980) Prigogine, Illya, *From Being to Becoming*, San Francisco, W. H. Freeman, 1980.

(Prigogine 1983) Prigogine, Illya, “The Rediscovery of Time,” a Discourse Prepared for the Isthmus

Institute, presented to the American Academy of Religion, December, 1983.
(Prigogine 1984) Prigogine, Illya, *Order Out of Chaos*, Bantam Books, New York, 1984.
(Prigogine 1997) Prigogine, Illya, *The End of Certainty*, Free Press Division of Simon and Schuster, New York, 1997.
(Raghavan 1988) Raghavan, Sridhar, "Diffusion Software Engineering Innovation,", THO218-8/88 *IEEE*, pp. 116-118, 1988.
(Raghavan 1989) Raghavan, S.A.; and Chand, D.R. "Diffusing Software-Engineering Methods*," IEEE Software Vol 6, no. 4*, July, 1989.
(Redwine 1984) Redwine Jr., Samuel T., et al, *DoD Related Software Technology Requirements, Practices, and Prospects for the Future*, IDA Paper P-1788, 1984.
(Rogers 1981) Rogers, Everett M., and D. Lawrence Kincaid, *Communication Networks: Toward a New Paradigm for Research*, Free Press, New York, 1981.
(Rogers 1983) Rogers, E. M., *The Diffusion of Innovation*, 3rd ed., Free Press, New York, 1983.
(Rogers 1975) Rogers, Everett M., "Where We Are in Understanding Innovation", paper for the *East-West Communication Conference on Communication and Change: Ten Years Later*, 1975.
(Rogers 1995) Rogers, E. M., *The Diffusion of Innovation*, 14th ed., Free Press, New York, 1983, 1995.
(Saboe 2001a) Saboe, Michael S., Grumbacher, P., Kloska, P., "Experience with the WinWin Process for Planning Software Infrastructure and Technology", to be published Int'l Conference on Group Systems Hawaii, 2002.
(Saboe 2001b) Saboe, M., "Technology Transfer: An Annotated Bibliography," Software Engineering Institute Web Page (to be posted), 2001.
(Saboe 2002) Saboe, Michael S., *A Software Technology Transition Engine – Industrial Model, A High Capacity, Accelerated, High Efficiency, Entropy Based Model*, Naval Postgraduate School Ph.D. Software Engineering Dissertation (Pending March 2002).
(Schroeder 2000) Schroeder, Daniel V., *Thermal Physics*, Addison Wesley Longman, San Francisco, CA, 2000.
(Schum 1994) Schum, David A., *Evidential Foundations of Probabilistic Reasoning, Wiley Series in Systems Engineering*, John Wiley, New York, 1994.
(Shannon 1948) Shannon, Claude, *A Mathematical Theory of Communication*, Bell Labs, 1948.
(Shaw 2001) Shaw, Mary, "Coming-of-Age of Software Architecture Research," *Keynote Address at the 23rd International Conference on Software Engineering*, www.cs.cmu.edu/~shaw/, May 2001.
(Snoddy 1926) Snoddy, G.S., "Learning and Stability," *Journal of Applied Psychology*, 10, pp. 1-36, 1926.
(Speigel 1998) Speigel, Murray, *Advanced Mathematics for Scientists and Engineers, Schaum's Outline Series*, McGraw Hill, 1998.
(Van de Ven 1991) Van de Ven, A.H., "Managing the process of Organizational Innovation" in *Changing and Redesigning Organizations*, ed. G.P. Huber, Oxford University Press, New York, 1991.
(Vigil 1994) Vigil, and Sarper, H. "Estimating the Effects of Parameter Variability on Learning Curve Model Predictions," *Int. J. Prod. Econ*., vol. 34, pp. 187-200, 1994.
(Wiio 1980) Wiio, Osmo A., *Information and Communication: A Conceptual Analysis*, Helsinki, Finland, University of Helsinki, Department of Communications Report, 1980.
(Yakovenko 2000) Yakovenko, V.M., A. Dragulescu, "Statistical Mechanics of Money", *Eur. Phys. J. B* 17 723-739 (2000)
(Yelle 1979) Yelle, L.E. "The Learning Curve: Historical Review and Comprehensive Survey," *Decision Sci*., vol. 10., no 2, p. 302-328, 1970.
(Zelkowitz 1995) (Zelkowitz, Marvin V., "Assessing Software Engineering Technology Transfer Within NASA," NASA technical report NASA-RPT-003095, National Aeronautics and Space Administration, Washington, DC, January 1995.
(Zelkowitz 1998) Zelkowitz, Marvin V., Dolores R. Wallace and David Binkley, *Understanding The Culture Clash In Software Engineering Technology Transfer*, University of Maryland Technical Report, 2 June 1998.
(Zipf 1949) Zipf, George Kingsley, *Human Behavior and the Principle of Least Effort*, Addison-Wesley Press, 1949.
(Zipf 1965) Zipf, George Kingsley, *Human Behavior and the Principle of Least Effort and Introduction to Human Ecology*, Hafner Publishing, New York, 1965, (facsimile of the 1949 edition).

APPENDIX

This appendix gives some insight into the control theory element of the current research model when maps exhibit eigenvalues in the following seven cases. The value of the norm (<1, =1, >1, real imaginary, etc) of the eigenfunction characteristic equation assimilation of reality based on experiences from prior time steps. A detail discussion can be found in Saboe 2002.

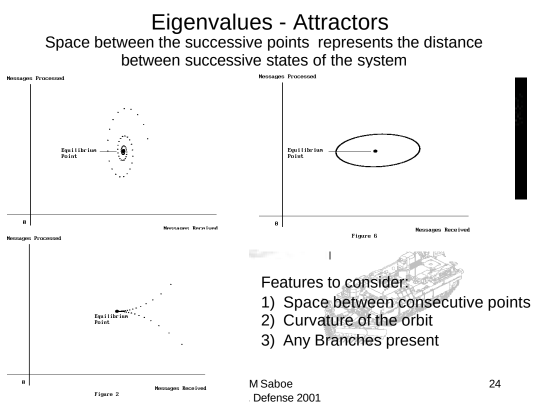
The eigenvalue of the system is significant. The eigenvalue represents an entropy metric, a measure. A more robust representation of the bakers transformation using a two dimensional model for the bakers transformation is found in Saboe 2002. This represents bakers transformation which consists of both information and entropy as suggested in equation (4.1)

$$\begin{cases} S_{H_{k+1}} = F(S_{H_k}, N_{i_k}) \\ N_{i_{k+1}} = G(S_{H_k}, N_{i_k}) \end{cases}$$

(4.1)

This system represents the dynamical aspect of folding and mixing, as well as the information theoretic approach representative of the alphabet, vocabulary and grammar of a technology.

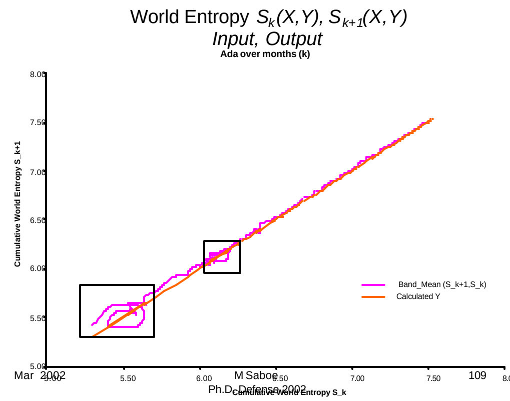
Using straight forward methods from control theory, we develop the characteristic equation of the system and find the eigenvalue. The figure below illustrates some representative attractors. The space between the successive points represents the distance between the successive states of the system. Curvature and branches are also features to be considered and explained.

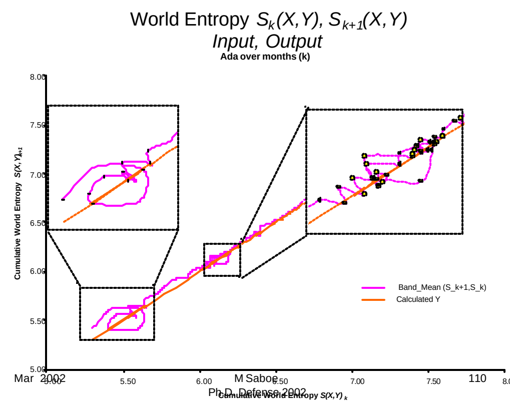


Eigenvalues in these state space maps represent attractors (sinks) and repellors (sources). Eigenvalues near 1.0 are generally indicative of smooth, orderly transitions between states. Eigenvalues near zero may represent rapid convergence to a single position or state. If the nodes represent organizations rather than individual publishers, the size of the organizational nodes should be reflected in the size of the eigenvalues.

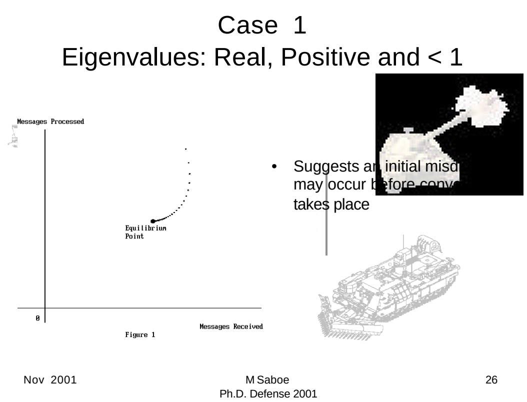
The following two figures illustrate actual examples of observed instabilities and convergence for the Ada programming language. This is representative of map based on a theoretical case 5, below, where the eigenvalue has a complex conjugate with the real part less than one. It appears that the system stabilizes and then

The empirical data for Ada has a confidence interval of $\pm 0.3\%$. This is developed from over 118,000 data points. Other data sets yield similar confidence intervals.

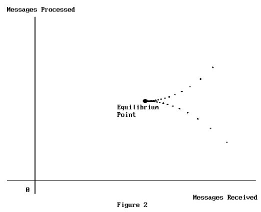


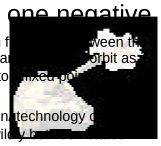


Seven cases that may be seen are illustrated below. Each figure briefly describes a case that may be observed in actual data.



Case 2
Eigenvalues: one positive and one negative

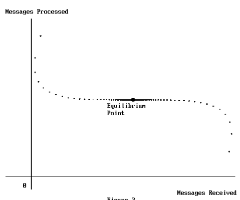




- The system f[...]ween the separate bra[...]bit as converges to [...]
- Organization technology c[...] oscillates wil[...]
- Depends on scale -- could be a small node capable of extremes
- Messages processed greater than received
  - could be explained by node size
  - capability, attrition/retention/insertion
  - new actors without organization/ technology experience

Nov 2001 M Saboe Ph.D. Defense 2001 27

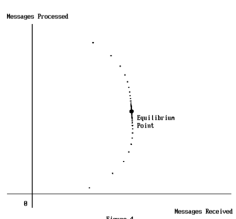
Case 3
Eigenvalues: one positive and one negative -- but size of Eigenvalues make a di[...]

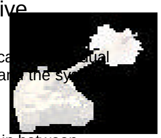


- From a dynamica[...] systems cases 2 and 3 lo[...]k similar - tech transfer or o[...]ganizational perspective intro[...]
- Messages processed exceed messages received!
  - Significant discovery opens up a bottleneck
  - Critical mass achieved - phase change
  - Accounts for memory and could happen if multiple time steps are required to process or learn

Nov 2001 M Saboe Ph.D. Defense 2001 28

Case 4
Eigenvalues: both negative

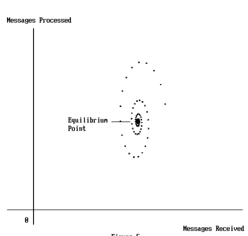


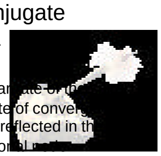


- Similar to ca[...] branches a[...] the sy[...] oscillating
- Relationship between messages received to processed is changing with each time step
  - Top branch, messages received increasing, while processed is decreasing
  - Bottom branch, opposite is true

Nov 2001 M Saboe Ph.D. Defense 2001 29

Case 5
Eigenvalues: Complex conjugate with the real part <1

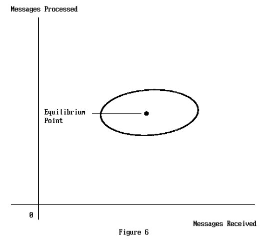


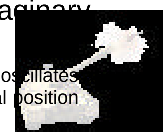


- The angular [...] and the rate of conver[...] should be reflected in th[...] organization[...]
- Eccentricity should make sense
- *Could there be a real basis for a software evolution spiral model?!*

Nov 2001 M Saboe Ph.D. Defense 2001 30

Case 6
Eigenvalues: purely imaginary

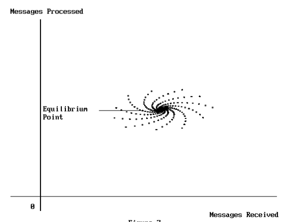




- System oscillates [...] a central position
- Intuitively reasonable as long as sensible
  - angular velocity
  - eccenticity
  - average diameter of orbit (semi major and minor axes of the ellipse)

Nov 2001 M Saboe Ph.D. Defense 2001 31

Case 7
Eigenvalues: Complex Conjugate with negative real part having absolute valu[...]

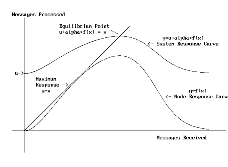


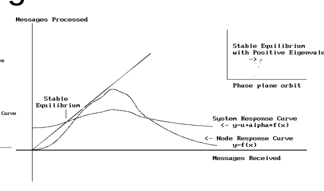
- Similar to c[...] different as [...]llows
- System leapfrogs over previous states as it converges to equilibrium
- This must be explained in terms of the nodes modeled

EigenFunction_examples.vbw

Nov 2001 M Saboe Ph.D. Defense 2001 32

Forcing Bifurcation





- Fixed α 0.3 -- increasing amount of clarification required - - increases overhead
- λ_0 shifts from positive to negative with increasing *u*

- Arrow points to stable equilibrium to which the node converges
- *Branches explained because system is three dimensional and has two complex conjugate eigenvalues that are causing a delay rotation*

Nov 2001 M Saboe Ph.D. Defense 2001 34

Additional discussion on these cases and bifurcation will be found in a paper to be published in October 2002, at the Monterrey 2002 workshop.

Back to the Top