Entropy Metrics for System Identification and Analysis

Article *in* Systems Engineering · June 2014

DOI: 10.1002/sys.21259

2 authors:

Joseph James Simpson

University of Washington Seattle

Mary Simpson

Entropy Metrics for System Identification and Analysis

Joseph J. Simpson, Mary J. Simpson
System Concepts, LLC

Abstract

Whole system metrics are valuable tools for use in systems science and engineering. Entropy metrics are defined, developed and demonstrated in this paper. Based on classical systems engineering methods and practices, these entropy metrics indicate the degree of order/disorder in any given system. A physical-entropy based metric and an information-based entropy metric are aligned with the two primary components of a system: objects and relationships. The physical-entropy based metric is called a relational score, and the information-based metric is called an object score. A subsystem score, based on the relational score and object score, is also developed and presented in this paper. A well-defined process, using these metrics, is used to evaluate the reduction of entropy and complexity associated with any specific system. The metrics and processes developed in this work have a prose component, a graphic component, and a mathematical component. These three components are aligned with the systems science techniques developed by John N. Warfield.

Introduction

The primary purpose of this work is to develop basic structural models that support the interpretive structural models commonly called N-Squared Charts and Automated N-Squared Charts. John N. Warfield augmented system structural modeling to support human reasoning about complex systems as well as to provide computer-based assistance in the communication of systematic mathematical reasoning[1]. Warfield divided system structural models into two, distinct, generic types: 1) basic structural models and 2) interpretive structural models. Basic structural models are system models that are rooted in mathematics, and evolved out of matrix and graph theory. Basic system structural models carry no domain specific, empirical or substantive information. Basic structural models are application domain independent. Interpretive structural models are those focused on representing empirical, substantive knowledge about a specific system domain, type or application. Considering these definitions, both N-Squared Charts and Automated N-Squared Charts are interpretive structural models.

In 1977, Andrew P. Sage published *Methodology for Large-Scale Systems* that proposed a "restrictive use of the words interpretive structural model[2]" that blurred the distinct line earlier drawn between the concepts of basic structural models and interpretive structural models. Sage's text bypassed discussion of basic structural models in his description of interpretive structural models, assigning the mathematical components of basic structural modeling to the practice of interpretive structural modeling. In 1981, Steward presented his interpretation of interpretive structural modeling that included both domain-specific information and domain-independent mathematical models, based predominantly on his work with design structure matrices[3]. In this work, the line between basic structural models that contain no empirical and/or substantive information, and interpretive structural models, was completely erased. Mixing these two types of system structural models in a manner that recognizes only one type of structural model application increases the complexity of using these structural model types in system analysis methods.

The authors reestablish the line and distinction between basic structural models and interpretive structural models as a mechanism to reduce the complexity associated with system structural analysis. Examples of basic structural model analysis methods are presented later in the paper as they apply to N-Squared Charts and Automated N-Squared Charts.

In this paper, a system is defined in two, complementary ways: a construction-rule definition and a function-rule definition. The construction-rule definition of a system[4] is "A system is a relationship mapped over a set of objects." The function-rule definition for a system[5] is "A system is a constraint on variation." The construction-rule system definition is the foundation of a number of classical system engineering graphical analysis and representation techniques including N-Squared Charts, Automated N-Squared Charts and Design Structure Matrices. The function-rule definition is also used to explore specific system configurations and determine whether these configurations are feasible, optimal and/or applicable to a specific system and deployment context. The construction-rule definition is applied to both the system objects and the system organizing relationship in a concurrent fashion. The constraints associated with the function-rule definition may be applied to the objects only, the relationship only, or a combination of both the objects and the relationship. The function-rule definition is mainly applied to the analysis of the system object sequence represented on the matrix diagonal in the examples presented in this paper.

Based on these two system definitions, two system entropy metrics are developed and demonstrated: the relational score and the object score. A third measure, called a subsystem score, is also developed and demonstrated. The two system entropy metrics are domain independent, and provide a consistent set of system metrics across all system application domains. The subsystem score has a domain-dependent value set that must be developed and applied by the system experts using the analysis technique. The subsystem score also has a domain-independent analysis component. These entropy metrics, demonstrated in three core examples, provide a direct conceptual and computational connection between systems science and the practice of systems engineering. Systems science, as developed by Warfield, uses three equivalent communication channels in the analysis and evaluation of systems: prose, graphics and mathematics[6]. In this work, a common graphical system representation – a square matrix found in N-Squared Charts, Automated N-Squared Charts, Design Structure Matrices, Reachability Matrices, Inclusion Matrices, and Implication Matrices - is combined with prose and mathematics to develop and refine a common set of system evaluation methods. The entropy examples in this paper only address entropy metrics for the N-Squared Chart and Automated N-Squared Chart techniques.

The basic system properties of structure, order and configuration are aligned with two basic entropy forms and types: physical entropy and information entropy. Physical entropy and its metric, the relational score, are defined and presented, along with information entropy and its metric, the object score. Finally, a subsystem metric and score are defined, applied in the examples and discussed. Two examples demonstrating these entropy forms, scores and metrics, are cited from systems engineering literature. The N-Squared Chart example is referenced from the work of Lano where the technique was being used in the system design mode. In the system design mode, the relationship is known and the determination of an acceptable set of objects and object configurations, is the objective of the activity[4,7]. The Automated N-Squared Chart example is referenced from the work of Hitchins, where the technique is used in the system discovery mode. In system discovery mode, the objects are known and the controlling relationship and the object configurations are the topics of study. The final example presents an application of the subsystem compression and expansion technique that integrates the relational score, the object score and the subsystem score.

Part One: System Structure, Order, and Configuration

Systems have some common characteristics that include structure and behavior. Warfield's *Science of Generic Design*[6]*,* further details the well-defined process named Interpretive Structural Modeling (ISM)[6].that was initiated in the 1970's. ISM focuses on describing the structure of a system. The goal of ISM is to develop a structured representation of a problem set or system previously viewed as unstructured and/or disordered. Steward extended and expanded the structural modeling component of ISM to include Design Structure Matrices (DSM) methods and techniques[8,9]. DSM techniques provide a well-established process for system structural analysis. DSM uses domain-dependent, structural configuration values to determine if a given system structural configuration is more valuable than another structural configuration of the same system. N-Squared Charts (N2C), developed by R.J. Lano, are also used to evaluate the structure of a set of system nodes and their interfaces[10]. The main system functions are subordinated to system interfaces in the N-Squared Chart approach. The N-Squared Chart approach provides a well-defined set of process steps that address system interface configurations and interface values. The automated N-Squared Chart (AN2C) method, developed by Derek K. Hitchins, adds the use of evolutionary computation to the evaluation and analysis of system structure[11]. Only the N2C and AN2C analyses and approaches are addressed in detail in this paper. There is a direct connection from these methods to DSM and ISM that the authors have addressed in other papers[12,13,14].

Each of these previously described system analysis techniques has well-defined processes that address the reduction of disorder and complexity. The N2C approach uses manual methods to analyze and evaluate the system structure. The AN2C approach uses a combination of human and automated analysis procedures. The DSM approach also uses a combination of human and automated analysis techniques. The Abstract Relation Type (ART) AN2C and ART DSM approaches use similar automated analysis techniques. This paper expands the authors' previous work in complexity reduction, and adds a well-defined process for the evaluation, measurement and reduction of system disorder. The primary contribution of these techniques is to establish well-defined processes that address and integrate both human and automated aspects of system evaluation. One of the design goals of these new metrics and techniques is to apply computer resources in areas where computers have proven beneficial, and to develop a human accessible interface that supports and enhances human performance in manual system evaluation tasks. Figure 1 depicts an overview of methods discussed and/or introduced in this paper.

Systems Analysis Methods	Aspects A	B	C	D	E	F
N-Squared Charts (N2C)	X		X			
Design Structure Matrices (DSM)	X		X			
Automated N-Squared (AN2C)		X	X	X		X
ART DSM Grouping		X	X	X	X	X
ART AN2C		X	X	X	X	X

Where

A = Manual Systems Analysis
B = Application of Evolutionary Computation
C = System Complexity Reduction
D = System Physical Entropy Concept
E = System Information Entropy Concept
F = Automated System Complexity Reduction

Figure 1. System Concept Allocation

These system analysis techniques all address the reduction of disorder in systems. The reduction of system disorder increases system analysis effectiveness by reducing both cognitive and computational complexity. Physical entropy is associated with the increase of disorder in any given system. The aforementioned system analysis techniques are related to physical entropy by the introduction of order into a system structure. The automated N-Squared Charts (AN2C), developed by Derek K. Hitchins, applies an 'N-Squared' Score system structure designed to measure the maximum configuration entropy associated with a given system configuration represented by the AN2C. The AN2C physical entropy concept is based on an appeal to energy efficiency and interface disorder; the more ordered the system, the lower the N-Squared Score.

Similar to the N2 Score developed by Hitchins, Simpson and Simpson developed a method to calculate the reduction of interface complexity associated with N-Squared Charts[13,14]. This complexity reduction technique is represented as an Abstract Relation Type (ART). ART methods reduce cognitive complexity using graphics, prose and mathematics to represent the system evaluation information. ART methods can reduce computational complexity by focusing computational resources only on the current system of interest by discarding consideration of non-system variables. Abstract Relation Types have been discussed and published for both the N2C and DSM system analysis methods.

The determination of the basic structural model components of the Automated N-Squared Chart is presented next to demonstrate the difference between basic structural models and interpretive structural models. The Automated N-Squared Chart is presented as a directed graph and a matrix that encode observed, empirical information about the system of interest. There is no specific mathematical model information associated with either the graph or the matrix. Using the fundamental equivalence among a directed graph model, a matrix model and a prose description of a binary relation, the system organizing structural relation can be determined and documented. Figure 2 shows equivalent representations for the "is-connected-to" organizing relation, which is irreflexive, asymmetrical and transitive. The features of the basic structural model are independent of the system content and/or domain of application. The relational attributes of the system structuring relation were used by the authors to create their evolutionary computational approach for identifying subsystems in this specific example.

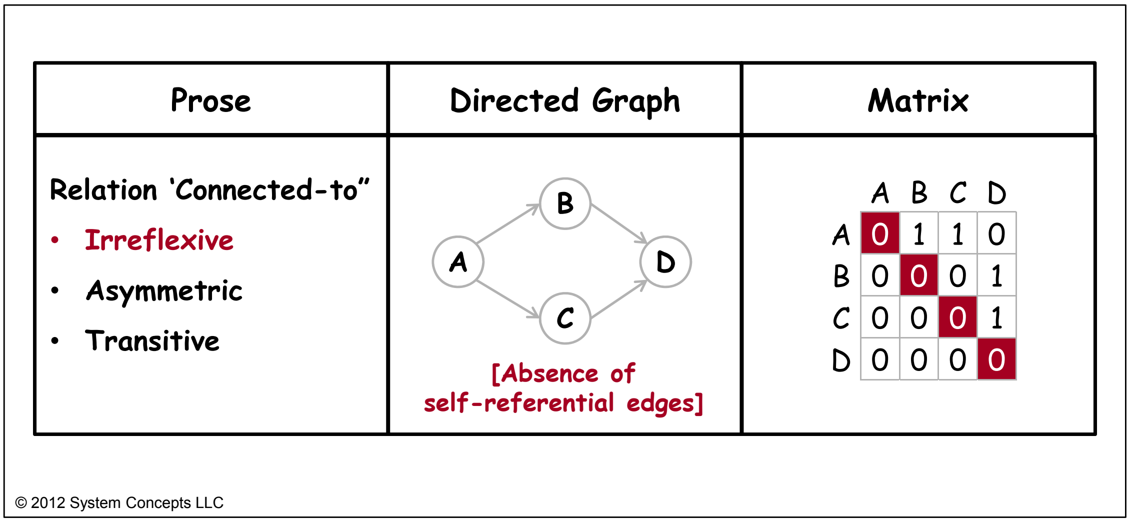

Figure 2. Fundamental Equivalence for a Binary Relation

The N2C ART calculation method uses a set of matrices designed specifically for this purpose. The N2C ART technique has a marking space, a value space, and an outcome space (see Figure 3). Each of these spaces must have at least one matrix, but can contain more than one matrix under some conditions. The N2C ART marking space matrix is used to record the *system structure*. The N2C ART value space matrix is used to record the *value associated with* the system structure. The N2C ART outcome space is used to record the output from the system value function that is applied to the marking space and the value space[13].

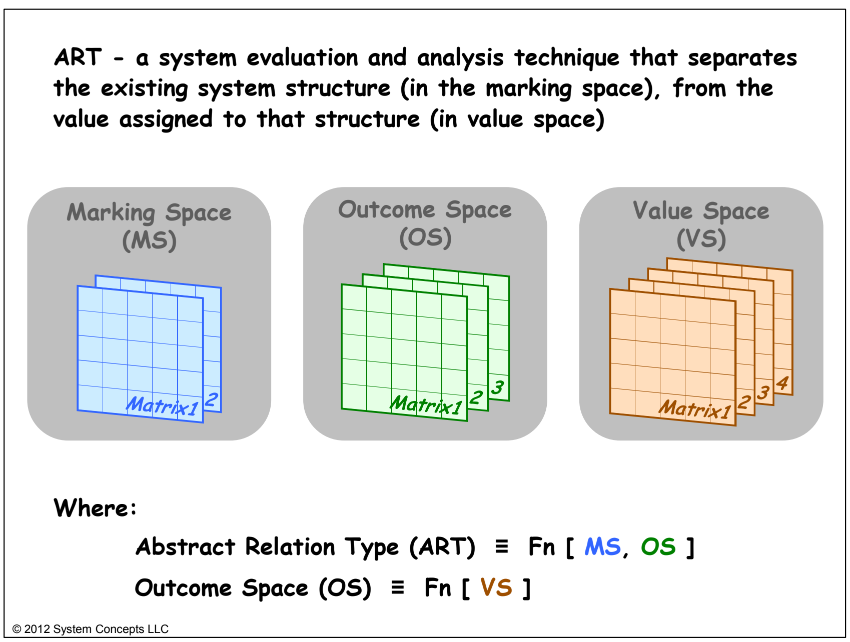

Figure 3. Notional depiction of Abstract Relation Type (ART)

The DSM ART approach is similar to the N2C ART approach; however, at this time the DSM ART approach has two basic grouping processes while the N2C ART approach only has one. In this work, the authors expand and refine the existing complexity reduction metrics to include specific physical entropy metrics and information entropy metrics. More N2C ART and DSM ART methods may be developed to expand the application of this technique. The concept of *system configuration entropy* (similar to physical entropy) is introduced as a total system metric. The concept of *system information entropy* (similar to information entropy from information theory) is also introduced as a total system metric. These two concepts and metrics are used to define a subsystem group metric. These concepts are discussed in more detail in the next section, and specific metric application examples are provided later in the paper.

Entropy Forms and Types

Entropy is a 'whole-system' metric. Two forms or types of entropy are considered in this work: entropy in thermodynamics, and entropy in information theory. Thermodynamic entropy will be called physical entropy, and information theory entropy will be called information entropy. The N2C ART and the DSM ART approaches use a matrix representation of a system structure. The diagonal of the matrix represents the system objects (or nodes) and the off-diagonal notations represent the relationship (is-connected-to) among the system objects (or nodes).

The ART system representation, in the marking space, focuses on the relationship that is mapped over the objects to create a system (see Figure 4). The "is connected to" relationship is used to create the N2C ART and the DSM ART.

> The *information entropy metric is associated with the system object sequence on the matrix diagonal,* and is called the **object score**.
>
> The *physical entropy metric is associated with the off-diagonal system relational interfaces*, and is called the **relational score**.

This approach to measuring system entropy provides a specific metric for each of the primary components of a system: the objects (object score) and the relationships (relational score).

Object A	Potential Connect	Potential Connect	Potential Connect	Potential Connect
Potential Connect	**Object B**	Potential Connect	Potential Connect	Potential Connect
Potential Connect	Potential Connect	**Object C**	Potential Connect	Potential Connect
Potential Connect	Potential Connect	Potential Connect	**Object D**	Potential Connect
Potential Connect	Potential Connect	Potential Connect	Potential Connect	**Object E**

Figure 4. ART Marking Space Graphical Map of System Components

Physical Entropy Relational Score

The primary property or characteristic associated with physical entropy is disorder. System states of high disorder have a high entropy rating, system states of low disorder have a low entropy rating. The energy used to maintain the system in a state of high disorder is not available to support the system functions. In these cases of high disorder, the system can be greatly degraded. This form of entropy motivated Hitchins' work and his development of the N2 Score for AN2C analysis[11].

The N2C ART relational score is based on the N2C feed-forward interface, and a feed-backward interface. Figure 4 provides a graphical map that depicts the system definition: "a system is a relationship mapped over a set of objects." Each square located on the diagonal of an N2C matrix has four faces: top, right, bottom and left. Each of these four faces represents a specific type of system interface that, when taken together, create a flow through the matrix from the upper left corner to the bottom right corner. Figure 5 shows the four faces, and delineates the type of potential interfaces starting at the top and moving clockwise: top, Forward Receive (FR); right, Forward Send (FS); bottom, Backward Receive (BR); and left, Backward Send (BS). Each of the primary system description prose elements introduced here – the feed forward interface, the feed backward interface, forward send, forward receive, backward send, and backward receive – are mapped directly to a graphic and a mathematical representation. These system flow and interface concepts support more refined and detailed system analysis techniques that are developed later in the paper. Further, these flow and interface concepts define and establish another set of basic structural models that are independent of the area of application or the empirical information content contained in the mathematical matrix model. Similar to the "is-connected-to" system structural relation, these flow model descriptions are used to evaluate any specific relational connection configuration using computer support and mathematical functions and equations.

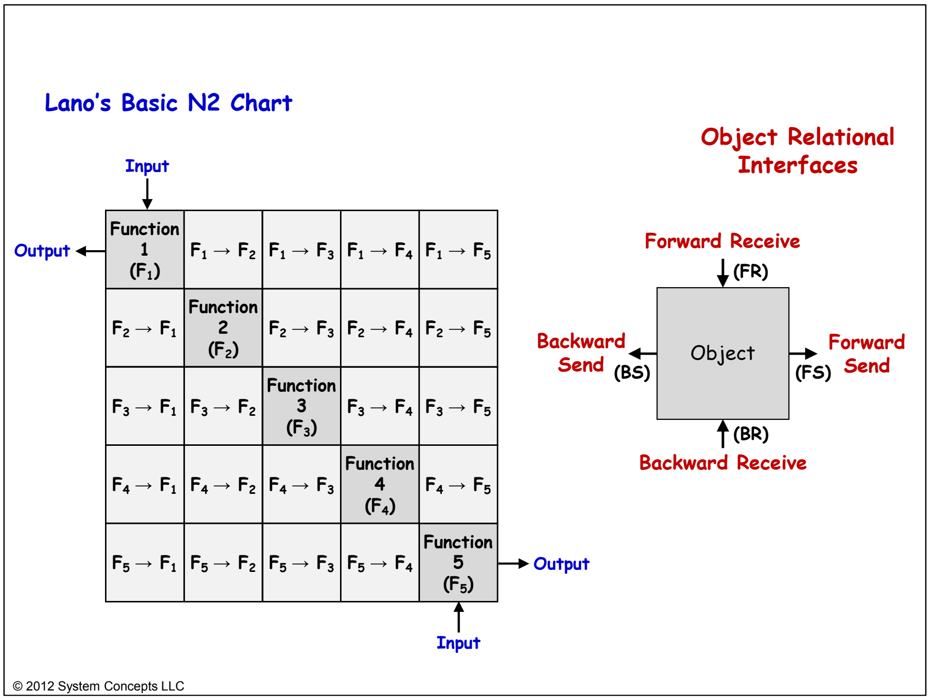

Figure 5. Object Relational Interfaces and Relation Flow

Figure 6 provides the detailed combinatorial relational interface calculations associated with a **9 by 9** N-Squared Chart. Since there is always one object in each row of a 9 by 9 N2C, the number of potential relational interface cells in any given row is (N-1), or eight (8). Figure 6, then, shows the number of different ways that each of the relational interfaces may be combined. The Sum column represents the **total** number of combinations, by row, which can be achieved by the combination of the number of cells and the selected number of relational interfaces. This calculation procedure was adapted and modified from Hitchins' Appendix A, "Configuration Entropy as a Useful Measure of Systems."

# of Cells	Number of Relational Interfaces								Sum of Row
	1	2	3	4	5	6	7	8	
1	1								1
2	2	1							3
3	3	3	1						7
4	4	6	4	1					15
5	5	10	10	5	1				31
6	6	15	20	15	6	1			63
7	7	21	35	35	21	7	1		127
8	8	28	56	70	56	28	8	1	255

Figure 6. Combinations of (N-1) Relational Interfaces for an N-Squared Chart

Row one depicts the selection of one cell, and one relational interface.

- There is only one (1) way to combine one cell and one interface.
- That combination is reflected in row one, column one, as one (1).
- The sum for row one is one (1), reflected in the “Sum of Row” column for row one.

Row two depicts the selection of two cells, and one, or two relational interfaces.

- For row two, column one, one relational interface may be placed in either one of the two selected cells – so there are two ways to combine two cells and one interface.
- That combination is reflected in row two, column one, as two (2).
- For row two, column two, there is only one (1) way to combine two interfaces and two cells.
- That combination is reflected in row two, column two, as one (1).
- The sum for row two is three (2 + 1), reflected in the “Sum of Row” column for row two as three (3).

Row three depicts the selection of three cells, and one, two, or three relational interfaces.

- For row three, column one, one relational interface may be placed in any one of the three selected cells – so there are three ways to combine three cells and one interface.
- That combination is reflected in row three, column one as three (3).
- For row three, column two, there are three (3) ways to combine two interfaces and three cells.
- That combination is reflected in row three, column two, as three (3).
- For row three, column three, there is only one (1) way to combine three interfaces and three cells.
- That combination is reflected in row three, column three as one (1).
- The sum for row three is seven (3 + 3 + 1), reflected in the “Sum of Row” column for row three as seven (7).

The remainder of the table is calculated in a similar manner.

This relational interface combinatorial calculation can now be applied to a matrix of a given size to create a maximum relational score for that matrix. A key complexity reduction technique developed in this paper is the ability to reduce the size of a matrix that represents a specific system in a manner that is well-defined and reversible. Using these techniques, a specific system may be viewed at multiple levels of reduction and/or detail. Two classical systems engineering examples from the literature are evaluated in this paper. In both cases the original system matrix is a 9 by 9 matrix. After the system evaluation and analysis techniques are applied to the 9 by 9 matrices, they are both reduced to a 5 by 5 matrix system model. In the next section of the paper, the maximum relational score for a 9 by 9 matrix and a 5 by 5 matrix will be developed as a basis to support further analysis and discussion of the system analysis examples presented in later sections.

Each interface cell in a matrix can be filled with a one (1) to create a ‘maximum’ relational score for a fully populated matrix. The directional (asymmetric one-way links) nature of the N2C connections in a fully populated matrix, completely defines the sequence order of the objects on the matrix diagonal. As shown in Figure 7, the relational score for each row is the sum of the forward-send (FS) score, the forward-receive (FR) score, the backward-send (BS) score and the backward-receive (BR) score for each object in the matrix. Based on the calculations shown in Figure 6, the first diagonal square in a nine (9) by nine (9) matrix of an N-Squared Chart (that is, the upper left corner square) has a 'forward-send' value of 255, a 'forward-receive' value of zero (0), a 'backward-send' value of zero (0), and a 'backward-receive' value of 255. The Figure 7 values are determined using the following process.

Each object on the matrix diagonal has four relational interfaces with flow direction associated with each interface: the forward send, the forward receive, the backward send and the backward receive. Starting with Object A, each of these interfaces is evaluated.

Forward Send

- Starting at the upper left hand corner of the matrix in Figure 7, the number of 1's in row A is determined to be eight (8).
- The right hand interface of Object A is the Forward Send interface.
- The score associated with the Forward Send interface of Object A is recorded in the FS column (first column) of the first row of the matrix relational score board.
- The value, based on the number of combinations of relational interfaces, is given by the sum column of eight (8) cells in Figure 6 - a value of 255.

Forward Receive

- The upper left hand corner of the matrix in Figure 7 shows no input cells associated with the top interface of Object A.
- The top interface of Object A is the Forward Receive interface.
- The score associated with the Forward Receive interface of Object A is recorded in the FR column (second column) of the first row of the matrix relational score board.
- The value, based on the number of combinations of relational interfaces, is zero (0).

Backward Send

- The upper left hand corner of the matrix in Figure 7 shows no output cells associated with the left hand interface of Object A.
- The left hand interface of Object A is the Backward Send interface.
- The score associated with the Backward Send interface of Object A is recorded in the BS column (third column) of the first row of the matrix relational score board.
- The value, based on the number of combinations of relational interfaces, is zero (0).

Backward Receive

- Starting at the upper left hand corner of the matrix in Figure 7, the number of 1's in column A is determined to be eight (8).
- The bottom interface of Object A is the Backward Receive interface.
- The score associated with the Backward Receive interface of Object A is recorded in the BR column (fourth column) of the first row of the matrix relational score board.
- The value, based on the number of combinations of relational interfaces, is given by the sum column of eight (8) cells in Figure 6 - a value of 255.

The second, and all ensuing rows, are calculated in a similar manner. (Starting with Object B on the matrix diagonal of row two, the number of 1's populating row two on the right hand side of Object B - the Forward Send interface - is determined to be seven (7). The sum row for seven cells in Figure 6 is consulted, and a value of 127 is determined to be the correct entry for column one, row two of the matrix relational score board representing the Forward Send interface of Object B. The Forward Receive top interface of Object B is determined to have 1 populated cell associated with this interface. The sum row for one (1) cell in Figure 6 is consulted to determine a value of one (1) for column two, row two of the matrix relational score board. The Backward Send interface of Object B is associated with one (1) populated cell. The sum row for one (1) cell in Figure 6 is consulted to determine a value of one (1) for column three, row two of the matrix relational score board in Figure 7. The fourth column of row two of the matrix relational score board is associated with the Backward Receive interface of Object B. There are seven (7) populated cells associated with the Backward Receive interface of Object B. The sum row for seven (7) cells in Figure 6 is consulted to determine a value of 127 for row two, column four of the matrix relational score board.)

Figure 7 displays the fully populated 9 by 9 matrix, and its' associated maximum relational score of 2008. The symmetry of the upper triangular (forward) and lower triangular (backward) portions of the matrix fully determine the order of the objects on the diagonal. The forward and backward interfaces determine the flow through the matrix from the upper left to the lower right.

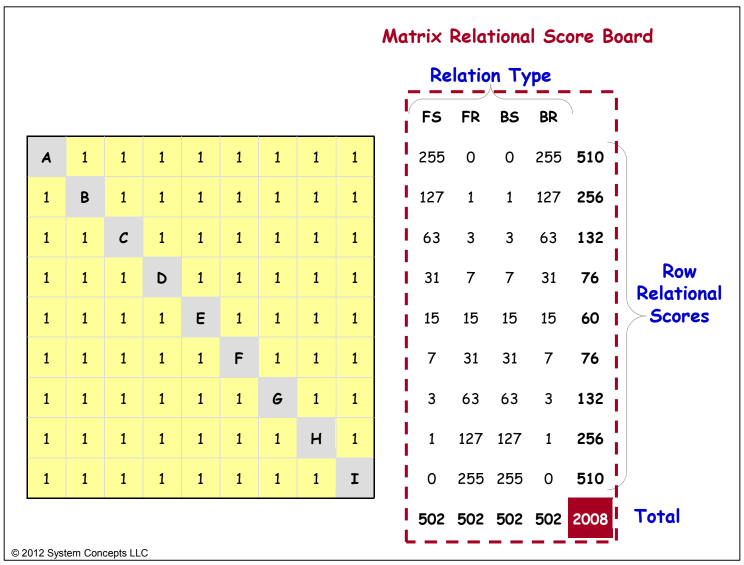

Figure 7. 'Fully Populated' 9 x 9 Matrix (N-Squared Chart)

The fully populated matrix relational score board displays a number of symmetric components. These symmetric components provide the numeric foundation upon which system evaluation and analysis computational techniques are based. These analysis techniques may be applied to system matrix representations that are not fully populated. In the partially populated system matrix models, the configuration *objects* on the diagonal combined with the *relational configuration,* determine the information value assigned to any sequence of objects on the matrix diagonal. For example, consider the fully-populated matrix shown in Figure 7. The fully-populated configuration fixes the order of the objects on the diagonal; they cannot change. Since the order of the objects on the diagonal is fixed, the matrix object score is zero and the object sequence contains no information.

Using the same basis (using the number of combinations shown in Figure 6, incorporating the same forward and backward interfaces, and applying the object interface definition), similar values are assigned to a fully populated 5 by 5 matrix in Figure 8. The associated maximum relational score for the 5 by 5 matrix is 104. The ability to reduce the maximum level physical entropy relational score associated with a given system model from 2008 to 104 is an excellent tool for reducing computational and cognitive complexity. The information entropy object score will be discussed and developed next.

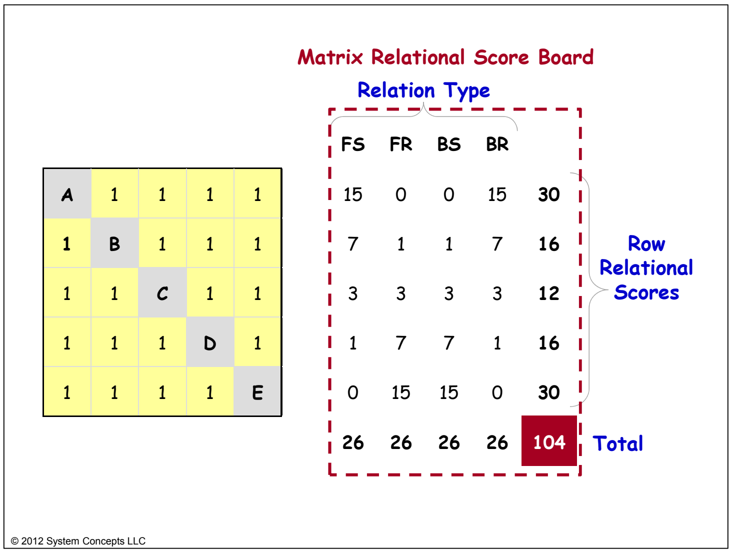

Figure 8. 'Fully Populated' 5 x 5 Matrix (N-Squared Chart)

Information Entropy Object Score

The primary property associated with information entropy is message uncertainty. The primary purpose of N-Squared Charts, Automated N-Squared Charts, DSM, ART N2C or ART DSM is the encoding and communication of empirical information associated with a system that is comprised of the objects displayed on the matrix diagonal. An object score has been developed to address the exchange of information between and among individuals using basic system engineering techniques (N2C, DSM, AN2C, ART DSM-Grouping, and ART AN2C). The object score enables and enhances the communication of large sets of data and information. This process of communication among individuals is similar to the types of message communications activities addressed by Shannon's Information Theory. The two primary components of information theory, message channel capacity and quantitative information metrics, are used as the basis of the object score that is developed in this work. The object score is associated with the specific sequence, or linear arrangement, of the objects on the diagonal of the marking space.

An information entropy metric has been developed to address the object sequence permutations on the matrix diagonal. In the fully-populated examples shown in Figures 7 and 8, all relational interfaces for the 9 by 9 and the 5 by 5 matrices are known. As a consequence, the sequence for these objects is fixed, and cannot change. There can only be an object sequence of A, B, C, D, E, F, G, H, and I for the 9 by 9 matrix. Similarly, there can only be an object sequence of A, B, C, D, and E for the 5 by 5 matrix. There is no possibility of changing the object arrangements. Shannon created an information entropy metric associated with the message source that measures the average amount of information conveyed by a message[15]. If the probability of a message occurring is 100 percent (well known in advance), the message contains zero (0) information. The message also has an information entropy rating of zero (0). The sequences A through I (9 x 9), and A through E (5 x 5), contain zero information, and can therefore be condensed into a single object node without losing any associated message information potential. These fully populated matrices have sequences with an object score of zero (0). When a matrix is *not* fully populated, the information entropy object score creates a substantially different picture.

The maximum object score is calculated using a system relation that does not constrain the permutations and ordering of the system objects on the matrix diagonal. Using a 5 by 5 matrix that has 5 objects on the diagonal, there are 5 factorial or 120 possible arrangements of these 5 objects. The object names, A, B, C, D and E are primitive symbols that are used to make up the non-primitive (compound) system symbols that are used in the object score analysis. Figure 9 shows the 120 system symbols associated with a 5 by 5 matrix. When calculating the maximum object score, all system symbols are equally probable with each system symbol having a one in 120 (1/120) chance of acceptance.

a b c d e	b a c d e	c a b d e	d a b c e	e a b c d
a b c e d	b a c e d	c a b e d	d a b e c	e a b d c
a b d c e	b a d c e	c a d b e	d a c b e	e a c b d
a b d e c	b a d e c	c a d e b	d a c e b	e a c d b
a b e c d	b a e c d	c a e b d	d a e b c	e a d b c
a b e d c	b a e d c	c a e d b	d a e c b	e a d c b
a c b d e	b c a d e	c b a d e	d b a c e	e b a c d
a c b e d	b c a e d	c b a e d	d b a e c	e b a d c
a c d b e	b c d a e	c b d a e	d b c a e	e b c a d
a c d e b	b c d e a	c b d e a	d b c e a	e b c d a
a c e b d	b c e a d	c b e a d	d b e a c	e b d a c
a c e d b	b c e d a	c b e d a	d be c a	e b d c a
a d b c e	b d a c e	c d a b e	d c a b e	e c a b d
a d b e c	b d a e c	c d a e b	d c a e b	e c a d b
a d c b e	b d c a e	c d b a e	d c b a e	e c b a d
a d c e b	b d c e a	c d b e a	d c b e a	e c b d a
a d e b c	b d e a c	c d e a b	d c e a b	e c d a b
a d e c b	b d e c a	c d e b a	d c e b a	e d d b a
a e b c d	b e a c d	c e a b d	d e a b c	e d a b c
a e b d c	b e a d c	c e a d b	d e a c b	e d a c b
a e c b d	b e c a d	c e b a d	d e b a c	e d b a c
a e c d b	b e c d a	c e b d a	d e b c a	e d b c a
a e d b c	b e d a c	c e d a b	d e c a b	e d c a b
a e d c b	b e d c a	c e d b a	d e c b a	e d c b a

Figure 9. System Symbols for a 5 x 5 Matrix

However, if the system is constrained to prohibit Object E from occupying the first element in the sequence, then the set of feasible system symbols is reduced from 120 to 96. This reduction in the number of acceptable system symbols increases the probability that any system symbol from the constrained group is acceptable.

In cases where the matrix is not fully populated, and the existing relational interfaces allow the reordering of objects on the diagonal, then the object score is greater than zero and the diagonal object sequence contains information. In cases where there are subsets of diagonal objects that have a fixed order, this subset of objects may be compressed into one diagonal cell. This matrix size reduction, using subsystem compression, is the heart of matrix system model complexity reduction. The object score applies only to the objects on the system matrix diagonal. The relational score applies only to the relational interfaces of the system matrix.

When a sequential subset of objects is identified, these objects and their related interfaces may be integrated into a grouping called a subsystem that has an associated subsystem score. The subsystem score may take different forms depending on the type of system modeled using the ART N2C techniques.

Subsystem Score

The authors' systems approach[4,7] combined with basic systems analysis methods and concepts are used in one of two basic modes: system discovery mode or system design mode.

- "In discovery mode, the things or objects are known and the primary activity is the determination of the relationships between the objects."
- "In the design mode, the relationship (function) is known and the objects and the configuration required to provide that relationship are the subject of study."

The subsystem score is based on a combination of the relational score, the object score, and system analysis mode. Unlike the relational and object scores, which are domain independent, the subsystem score can be domain dependent and contain domain-dependent semantics. Subsystem grouping may be performed based on characteristics of individual objects, and/or on the identification of objects that populate a subspace of the system space.

Part Two: Examples of System Relational, Object and Subsystem Scores

In the following section, three examples (two from SE literature and an original integration example) are evaluated to provide an overview of the processes that develop and apply the physical and information entropy metrics. The first example is referenced from a 1979 publication, wherein manual analysis methods and techniques evaluated the system design problem. The second example is referenced from a book published in 2003 that demonstrated automated systems analysis that supports the manual identification of subsystems embedded in a disordered system matrix model. . The third example integrates the relational score, object score and subsystem score to demonstrate subsystem compression and expansion techniques.

As mentioned earlier in the paper, the entropy metrics and their associated mathematical processes are basic structural models. Basic structural models are domain independent, and are used to help the human analyst achieve the desired outcome from the system evaluation activity. These basic structural models *augment and assist* the analytical procedure and do not override human judgment, or in any way invalidate the classical methods that have been used effectively for decades. The authors selected these specific literature examples because they believe that these are examples of interpretive structural models that do not have a clear connection to the mathematical techniques associated with basic structural models. In these examples, the authors augment these classical examples with the basic structural modeling capability to enable the use of these techniques in new and different applications.

Example One: N-Squared Chart

The N-Squared Chart (N2C) analysis example from Lano's *A Technique for Software and Systems Design* was selected to support the examination of an N2C example that was completed using manual techniques[10].

N2C Relational Score

This N-Squared Chart's relational score calculation technique, and its' respective values, are shown in Figure 10.

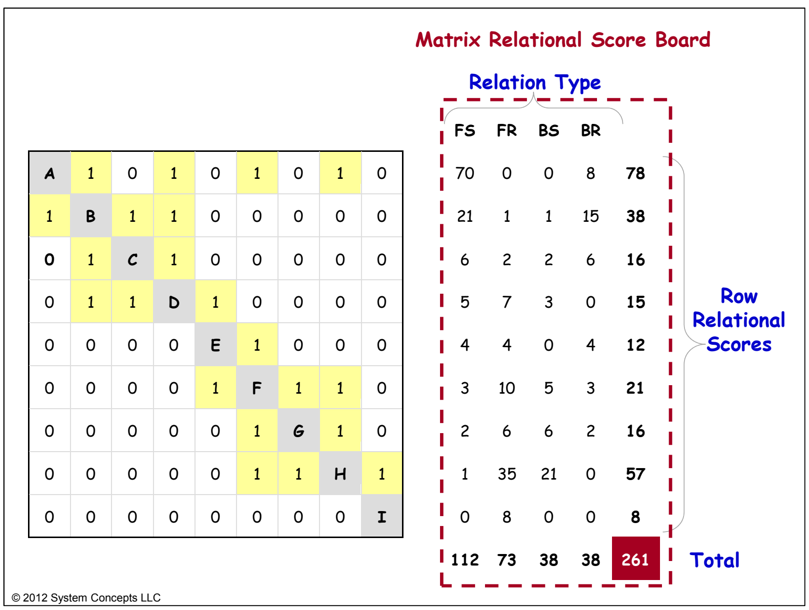

Figure 10. Manual N-Squared Chart Example

In this example, the relational score is 261. This example presents the output from the manual analysis discussed in Lano's work, and provides the associated relational score for the physical entropy analysis developed by the authors. Lano's matrix example also introduces a type of system and subsystem analysis process that is similar to the information entropy process developed by the authors. The subsystem grouping is depicted in Figure 11, with the subsystems highlighted in green. This systems analysis process identifies areas of 'highly-coupled' relational groups, and compresses these groups into subsystems that fit in a single cell. In this case, the system and design information was collected from human experts by the system engineering team. The manual N-Squared Chart methods were used to graphically encode and analyze the proposed system components, interfaces and structure. The graphical representation of the system structural information increased the communication ability of the engineering and product team members.

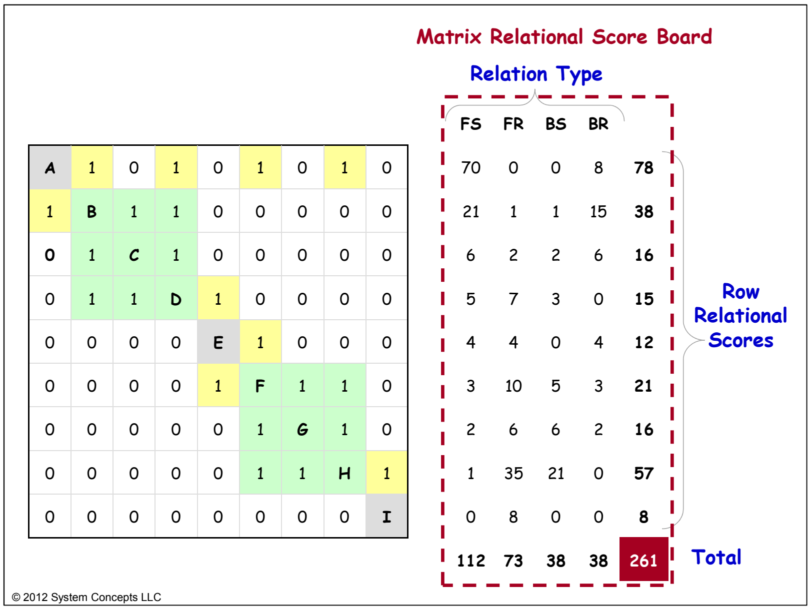

Figure 11. Well-Defined Object Sequence

The sequence of Objects B, C and D is well-defined by the highly-coupled relational marks highlighted in green in the upper left hand corner of the matrix. The B-C-D Object sequence was determined to be a subsystem group. In a similar fashion the sequence of Objects F, G, and H is well-defined by their respective highly-coupled relational marks, and is identified as a subsystem group. In both cases, the fixed nature of the sequence coupled with the fixed relational marks, enables the compression of each of these subsystem groups into a single cell. Because each subsystem sequence is defined and well known, each subsystems series of objects contains no information, and the information entropy is zero (0) for both subsystem sequences.

In this manual example, systems analysts worked with the engineering team members, and produced a simpler system representation that is shown in Figure 12. This simpler representation of the system was a more effective conceptual design tool, due to the previously reported cognitive complexity reduction[14]. This simpler matrix representation is identical to the system produced if the two green highlighted areas in Figure 11 are compressed into two one-cell object representations, cell D and cell F. The relational score for the reduced (5 x 5) manual example is 34. The maximum relational score for a 5 by 5 matrix was found to be 104; the maximum relational score for a 9 x 9 matrix was 2008. This simplified matrix example reduces the maximum relational score from 2008 to 104 or 19.3 times. This manual example also shows the original larger 9 x 9 matrix relational score of 261 and the simplified matrix (5 x 5) with a relational score of 34, which produces a relational score reduction of 261/34 or 7.7 times.

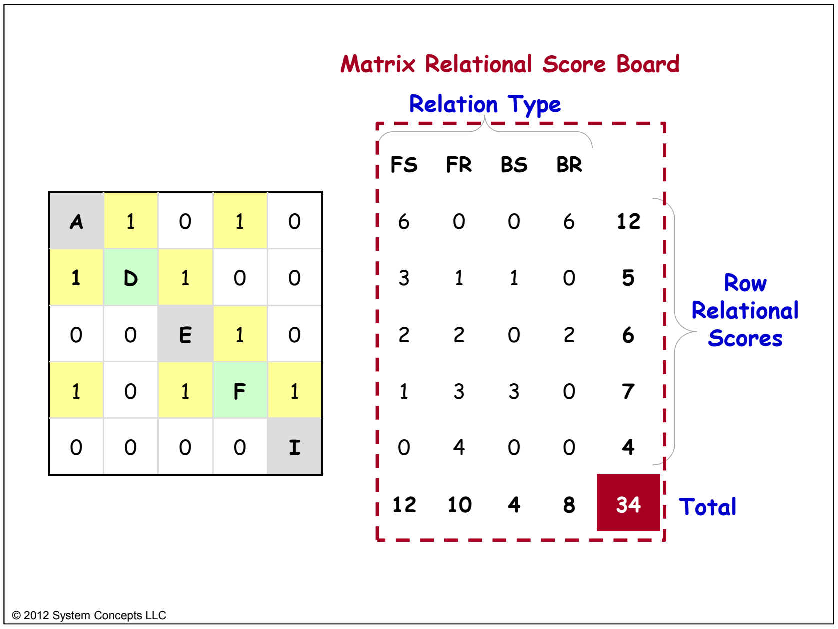

Figure 12. Relational Score for Compressed Matrix Example

N2C Object Score

As shown in the Figures 11 and 12 N-Squared Charts, columns B, C and D were identified as a subsystem and compressed into column D, and columns F, G and H were identified as a subsystem and compressed into column F. This resulted in a smaller number of system objects, and a different object sequence. The information entropy metric has been developed to address the object sequence permutations on the matrix diagonal. In the example shown in Figure 11, all of the relational marks for these two groups of objects are known. As a consequence, the sequence for these objects is fixed. The fixed object sequences of B,C,D and F,G,H are then identified as subsystem groups. There is no possibility of changing the object arrangement in the subsystem groups. Each subsystem group is then compressed into a single cell on the matrix diagonal, as shown in Figure 12. Shannon created an information entropy metric associated with the message source that measures the average amount of information conveyed by a message[15]. If the probability of the message content is 100 percent (well known in advance), then the message contains zero (0) information, and also has an information entropy rating of zero (0). The sequences B, C, D and F, G, H contain zero information, and can therefore be compressed into two object nodes (D and F) on the matrix diagonal without losing any associated message information potential. Each of these two sequences [(B,C,D) and (F,G,H)] has an object score of zero (0).

N2C Subsystem Score

As previously defined, the subsystem score is based on a combination of the relational score, the object score, and the system mode. Discovery mode and design mode may be interleaved when necessary. The example presented here – an N2C technique – was applied in the system design mode to allocate functions to implementable components. *In this specific example*, the **subsystem score** is the **ratio of the final number of objects on the diagonal to the original number of objects** on the diagonal. This ratio is 5/9 or a subsystem score of 55.56 percent. A subsystem score of 55.56 percent indicates a 44.44 percent reduction in the number of system objects. The reduction in the number of objects reduces cognitive complexity by compressing the subsystems with no information. Computational complexity is also reduced by eliminating a significant number of matrix components. Unlike the relational and object scores, which are domain independent, the subsystem score can be domain dependent and contain domain-dependent semantics. As a consequence, we address only the domain-independent aspects of the subsystem score in this paper.

N2C System Entropy Metric Calculation Process

The system evaluation process associated with manual N-Squared Chart analysis approach is different than the Automated N-Squared Chart process that includes an automated or evolutionary computational component. In this manual N2C case, only one larger system representation was developed – an ordered representation. When automated analysis techniques are used, there are at least two basic matrix forms: the original disordered system matrix model and the final ordered system matrix model. The system entropy calculation steps for the manual N2C process accommodate the absence of the original disordered matrix.

The system entropy metric calculation process has thirteen basic steps, grouped into four general phases. Phase one (1) is analysis initialization and base metric calculation. Phase two (2) is matrix analysis and evaluation. Phase three (3) is production of the compressed system matrix. Phase four (4) is final metric calculation and matrix size compression evaluation.

Phase One – Analysis Initialization and Base Metric Calculation

- M-1. Identify the system of interest and collect data.
- M-2. Develop original unstructured system marking and value spaces.
- M-3. Calculate the maximum relational score for the system.
- M-4. Calculate the maximum object score for the system.

Phase Two - Matrix Analysis and Evaluation

- M-5. Use manual techniques to find a minimum relational score (structure system).
- M-6. Evaluate the system structure with the minimum relational score, and find areas that have a zero (0) object score.

Phase Three - Production of Compressed System Matrix

- M-7. Compress each 'object-score zero area' into one diagonal object node.
- M-8. Calculate the maximum relational score for the new compressed system structure.

Phase Four - Final Metric Calculation and Matrix Size Compression Evaluation

- M-9. Calculate the reduction in the system maximum relational score between the two different sized matrices.
- M-10. Calculate the minimum relational score for the new compressed system structure.
- M-11. Calculate the reduction in the minimum relational score.
- M-12. Calculate the maximum object score associated with the new compressed system structure.
- M-13. Calculate the reduction in the system maximum object score.

Table 1 shows the calculation steps, and scores, for the N-Squared Chart example shown in Figures 10, 11 and 12.

Table 1. Evaluation of Lano's N-Squared Chart Example

Process Step	Matrix Size	Number of Relations	Maximum Relational Score	Maximum Object Score	Minimum Relational Score	Connection Score Reduction	Maximum Object Score Reduction
Step 1	9 x 9	21					
Step 2	9 x 9	21					
Step 3	9 x 9	21	2008				
Step 4	9 x 9	21	2008	362880			
Step 5	9 x 9	21	2008	362880	261		
Step 6	9 x 9	21	2008	362880	261		
Step 7	5 x 5	8					
Step 8	5 x 5	8	104		34		
Step 9	5 x 5	8	104		34	2008 to 104	
Step 10	5 x 5	8	104	120	34		

Process Step	Matrix Size	Number of Relations	Maximum Relational Score	Maximum Object Score	Minimum Relational Score	Connection Score Reduction	Maximum Object Score Reduction
Step 11	5 x 5	8	104	120	34		
Step 12	5 x 5	8	104	120	34	261 to 34	
Step 13	5 x 5	8	104	120	34		362880 to 120

Lano's N-Squared Chart example (referenced in Figures 10, 11 and 12) was developed using manual analysis techniques. The following section will explore the system entropy metric calculation process in detail using an automated N-Squared Chart example from Hitchins' work.

Example Two: Automated N-Squared Chart

System analysis tasks may be performed on existing systems that are poorly understood, defined and documented. The first step in the system entropy metric calculation process is the identification of the system of interest (set system boundaries), and the collection of relevant data associated with the system. In many cases, individuals associated closely with one component (object) of the system may have limited knowledge about the total system, but maintain a high degree of useful information about the specific subsystem and/or component. In cases of poorly defined and documented systems, the amount of manual effort directed toward the total system analysis can be reduced using evolutionary computation techniques.

The example considered next is taken from the work of Derek J. Hitchins[11], and is discussed in some detail in the following sections. The basic steps given in Hitchins work are expanded here to add further detailed analytical steps. As detailed earlier in the paper, the organizing system relation was identified along with its relational attributes by the authors using the mathematical concepts and processes associated with basic structural modeling. Unlike Lano's work, the Automated N-Squared Chart example starts with a disordered system representation, and uses evolutionary computation techniques to reduce the disorder and to identify sub-systems within the total system. The identification of subsystem groupings is one of the main features of the AN2C approach presented by Hitchins. The N2 Score proposed by Hitchins has proven effective in identifying and grouping subsystems. In Hitchins' example, the systems analysis approach is being used in the discovery mode. The system of interest already exists and is operating, but the specific subsystem structure is not known. The AN2C approach is used to discover, analyze and evaluate the existing disordered system of interest to identify ordered subsystem configurations.

AN2C Relational Score

Within Hitchins' Automated N-Squared Chart technique, the N2 Score developed by Hitchins is effective in subsystem identification. The authors created a software-based, system evaluation technique (ART AN2C) that generated the same final subsystem configurations as those presented by Hitchins[11]. The cognitive complexity associated with the system of interest was greatly reduced, and the subsystem grouping pattern became apparent in the final system configuration. However, when a more detailed relational score and object score analysis are performed on the example, it became clear that there appeared to be a missing relational connection in the analysis published in Figure 8.13 presented in Hitchins' book. Figure 13 shows two AN2C matrix configurations. On the upper left, the original AN2C matrix configuration – a disordered 9 by 9 matrix – and its associated relational score using the system entropy metric calculation process are detailed. The upper left figure also shows the N-Squared Score for that matrix found by applying Hitchins N-Squared Score procedure found in Appendix A[11]. The lower right figure addresses the same disordered 9 by 9 matrix, but now contains the "missing" relational interface – along with the applicable changes in both the system entropy metric calculation and Hitchins N-Squared Score. The increased resolution and analytical power of the entropy metrics create a useful cognitive complexity reduction mechanism as well as an automated cross-check of the presented system structure.

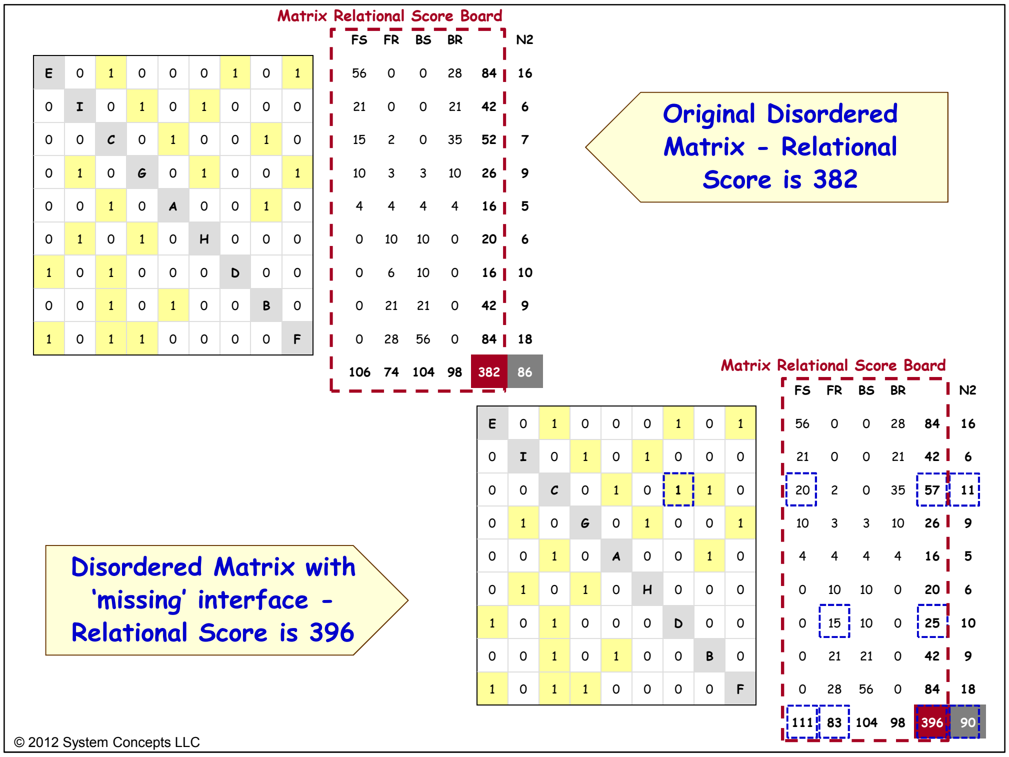

Figure 13. Disordered 9 x 9 Matrices

The more detailed relational scores associated with the disordered matrices presented in Figure 13 highlight the fact that Objects H, D, B, and F have no forward-send or backward-receive relational interfaces, with a zero (0) associated with each of these objects in the applicable scoring column. The numerical values in the scoring columns provide the basis upon which the mathematical functions associated with the prose descriptions of the matrix configurations can be constructed, evaluated and deployed.

AN2C Object Score

For this Automated N-Squared Chart example, once the disordered matrix shown in Figure 13 has been ordered using evolutionary computation, the missing feed-forward relational interface between Object C and Object D is readily identified by the zero (0) in the forward-send column for Object C and the zero(0) in the forward-receive column for Object D (see Figure 14). While the objects are still grouped in the same pattern, the more detailed relational interface configuration information provides further pattern and flow data that can be automatically evaluated to verify and validate the specific interface flow pattern. For example, in Figure 14 it is clear that the system forward flow path from Object C to Object D does not exist. The absence of this connection prevents the forward flow from Objects A, B, and C, to the rest of the system, and creates a distinct boundary in the internal system flow. Once the missing system relational interface is identified and repaired, as shown in Figure 14, there is a clear pathway for internal system flow between and among each of the objects.

An analysis of the detailed relational scores presented in the matrix relational score board sections of Figure 13 shows the presence of multiple zero (0) valued relational interfaces in both of the disordered system relational score boards. In particular, there are four zeros in both the FS and the BR positions on the relational score board in the 6th through 9th row area. In addition, there are five zeros in the FR and the BS positions on the relational score board in the 1st through 3rd row area. In contrast, Figure 14 (representing ordered matrices), shows considerably less zeros. The upper left relational score board (with one relational interface still missing) has zeros (0) in 6 positions; the 1st row area has one 'FR' zero and one 'BS' zero, the 9th row area has one 'FS' zero and one 'BR' zero, and the other two zeros are located where the 'missing' relational interface is found – in the 3rd row under FS and the 4th row under FR. The lower right matrix relational score board – where the relational interface has been inserted - has zeros in 4 positions: the 1st row area has one 'FR' zero and one 'BS' zero, the

9th row area has one 'FS' zero and one 'BR' zero. Given the flow in a basic N-Squared chart (see Figure 5 – Object Relational Interfaces and Relation Flow), these 4 zeros will always be present when the matrix has been ordered – showing the output, input at the upper left hand corner and the input, output at the lower right hand corner. These numerical patterns provide an indicator of the relative order or disorder in a given system structure. Automated evaluation of system order, or disorder, can be based – in part - on this numerical indicator associated with the system relational scores that are organized and displayed in the system relational score boards.

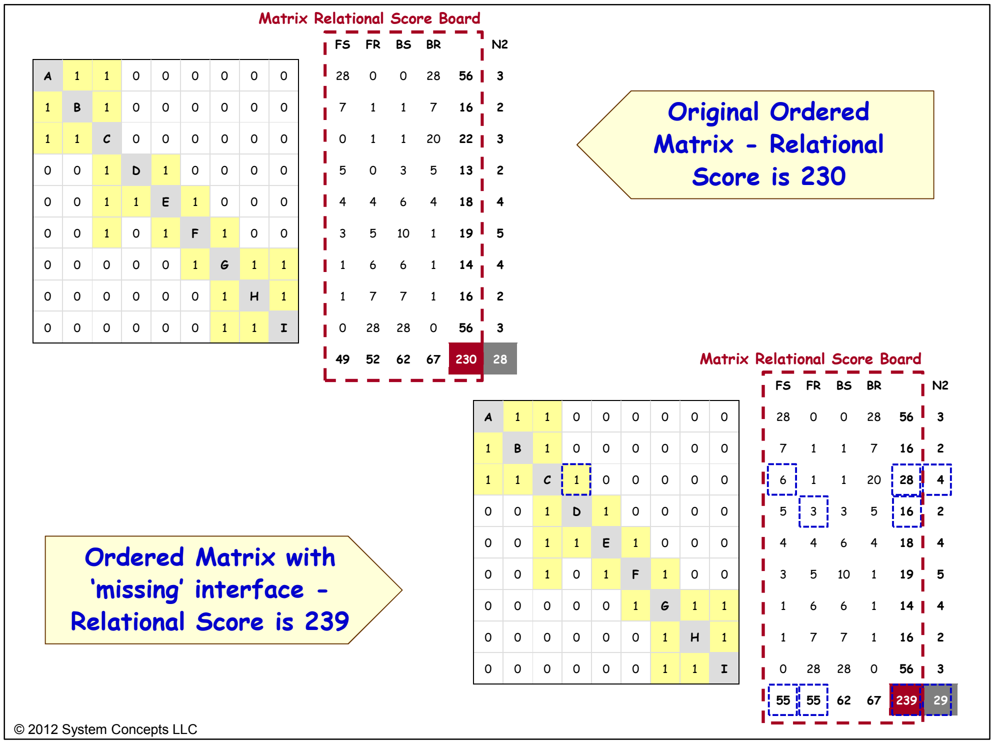

Figure 14. Ordered 9 x 9 Matrices

The ordered system shown in Figure 14 has two sequences with an object score of zero (0). These sequences are the object sequence A, B and C, and the object sequence G, H and I. The highly-coupled relational interface markings for the objects in Figure 14 completely determine the object sequence. Because the object sequence is now well-known, these object sequences contain no information, and have an information entropy rating of zero (0). When a sequence of objects receives a rating of zero (0) for the object score, that sequence of objects can be compressed into a single object cell on the diagonal without losing any system information. Using the object score technique, the resultant compressed system configuration is shown in Figure 15. Once the original system representation has been compressed from a 9 x 9 to a 5 x 5 matrix representation, it is very clear that the missing relational interface between Object C and Object D may create internal system interaction restrictions. However, once the missing relational interface is added to the system interfaces, the system interaction restrictions are lifted.

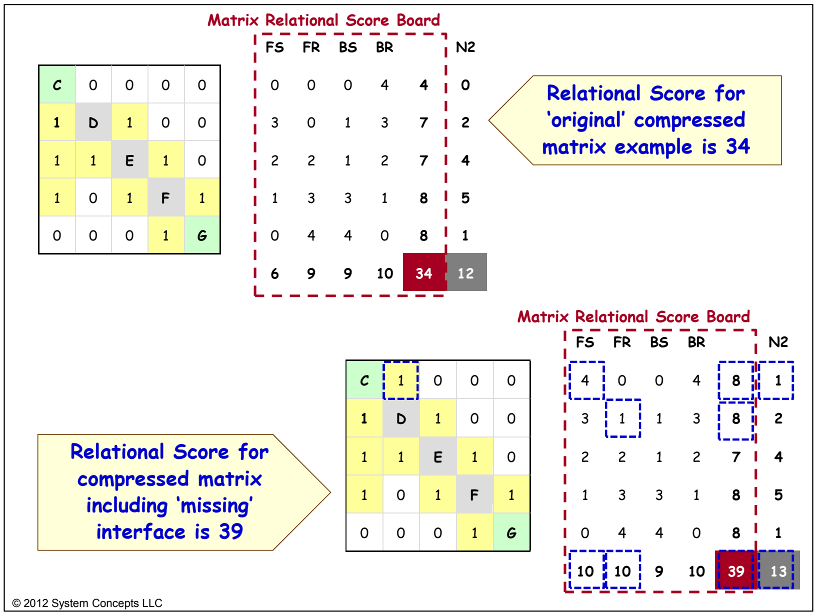

Figure 15. Ordered 5 x 5 Matrices

AN2C Subsystem Score

As previously noted, the identification of subsystem groupings is the main feature of the AN2C approach presented by Hitchins. As a result of the difference in Hitchins and the authors' techniques, the subsystems identified by each technique are different. Hitchins identifies three subsystems; the authors' subsystem score identifies only two. In this specific example, the authors' subsystem score is the ratio of the final number of objects on the diagonal to the original number of objects on the diagonal. This ratio is 5/9 or a subsystem score of 55.56 percent. A subsystem score of 55.56 percent indicates a 44.44 percent reduction in the number of system objects. The reduction in the number of objects reduces cognitive complexity by compressing subsystems with no information. Computational complexity is also reduced by eliminating a significant number of matrix components. Unlike the relational and object scores, which are domain independent, the subsystem score can be domain dependent and contain domain-dependent semantics. As a consequence, we address only the domain-independent aspects of the subsystem score in this paper.

AN2C System Entropy Metric Calculation Process

The system analysis and evaluation process associated with the Automated N-Squared Charts includes an automated or evolutionary computational component. In the automated case the first step is the identification of the system of interest and the development of the disordered matrix. Once the disordered system has been identified, then evolutionary computation can be used to search for other, more ordered or optimal, system configuration alternatives. The manual identification of these alternative system configurations is a major road block in the effective use of these types of techniques. Evolutionary computation techniques are viewed as a technology that can remove this road block. After the original disordered and/or unstructured system configuration has been identified and the marking space matrix developed, an automated process can be used to generate a range of alternative configurations. Domain experts and system experts can then use this range of alternative system configuration as the foundation for system analysis and selection procedures. The general automated process steps are listed below.

The system entropy metric calculation process has thirteen basic steps grouped into four general phases. Phase one (1) is analysis initialization and base metric calculation. Phase two (2) is matrix analysis and evaluation. Phase three (3) is production of the compressed system matrix. Phase four (4) is final metric calculation and matrix size compression evaluation. The processes in Phase Two are automated and provide the ability to evaluate large system matrix models with less cost than manual analysis methods.

Phase One - Analysis Initialization and Base Metric Calculation

A 1. Identify the system of interest and collect data.
A 2. Develop original unstructured system marking and value spaces.
A 3. Calculate the maximum relational score for the system.
A 4. Calculate the maximum object score for the system.

Phase Two - Matrix Analysis and Evaluation

A 5. Use automated techniques to find a minimum relational score (structure system).
A 6. Evaluate the system structure with the minimum relational score, and find areas that have a zero (0) object score.

Phase Three - Production of the Compressed System Matrix

A 7. Compress each 'object-score zero area' into one diagonal object node.
A 8. Calculate the maximum relational score for the new compressed system structure.

Phase Four - Final Metric Calculation and Matrix Size Compression Evaluation

A 9. Calculate the reduction in the system maximum relational score entropy between the two different sized matrices.
A 10. Calculate the minimum relational score for the new compressed system structure.
A 11. Calculate the reduction in the system minimum relational score entropy.
A 12. Calculate the maximum object score for the new compressed system structure.
A 13. Calculate the reduction in the system maximum object score.

Table 2 shows the calculation steps, and scores, for the Automated N-Squared Chart (AN2C) example shown in Figures 13, 14 and 15.

Table 2. Evaluation of Hitchins' Automated N-Squared Chart Example

Process Step	Matrix Size	Number of Relations	Maximum Relational Score	Maximum Object Score	Minimum Relational Score	Connection Score Reduction	Maximum Object Score Reduction
Step 1	9 x 9	22					
Step 2	9 x 9	22					
Step 3	9 x 9	22	2008				
Step 4	9 x 9	22	2008	362880			
Step 5	9 x 9	22	2008	362880	396		
Step 6	9 x 9	22	2008	362880	396		
Step 7	5 x 5	10					
Step 8	5 x 5	10	104				
Step 9	5 x 5	10	104			2008 to 104	
Step 10	5 x 5	10	104	120	39		
Step 11	5 x 5	10	104	120	39		
Step 12	5 x 5	10	104	120	39	239 to 39	
Step 13	5 x 5	10	104	120			362880 to 120

Hitchins' Automated N-Squared Chart (AN2C) example (referenced in Figures 13, 14 and 15) was developed using automated analysis techniques based on evolutionary computation. The system entropy metrics used for the evaluation of both the manual and automated system analysis techniques support the automated analysis of detailed system configurations. The automation of these processes allows the evaluation of much larger systems, and enables the reduction of the human effort required to effectively process system structures of a very large size.

Final Example: A System Integration Analysis using ART AN2C

The two previous examples demonstrated the use of ART N-Squared Chart and ART Automated N-Squared Chart techniques on systems of a similar size, but with different system analysis modes (design and discovery). The first example, from Lano, was an example of system analysis for system design communication and evaluation. The second example, from Hitchins, was an example of system discovery – the discovery of subsystems in an existing, operational system.

In this third example, the authors present another kind of system integration modeling and evaluation task that is supported by ART AN2C processes. This example provides a method to incorporate a new object into an existing system. The system discovery example (Example Two) from Hitchins cited earlier in the paper will be used for illustrative purposes. The initial starting point of this example is the 'corrected' disordered matrix referred to in Figure 13, on the lower right side of that graphic. That matrix – originally containing Objects A through I, is now expanded to include the insertion of a new subsystem containing Objects J through O. The final systems analysis will present not only the original existing objects, but also the configuration of the system after a new subsystem is inserted to support a new system function. This type of system discovery and evaluation is useful when the system context has a number of incomplete, or poorly-defined components that must be modified to support a new operational objective.

Figure 16 shows two matrices. For clarity, on the left is the 9x9 matrix (initial starting point) taken from Figure 13. The relational interface markings remain the same. The objects along the diagonal have been renamed, as the problem being analyzed and solved is a new one. On the right is the resultant 15 by 15 system matrix model, reflecting the inserted (new) subsystem. The new subsystem is highly coupled and is placed in the lower right hand corner, along with two new relational interface marks (see column I, row J and column J, row I) which connect the new subsystem to the original matrix. As can be seen, the original, distributed, disordered set of system components remain in the upper right-hand corner of the 15 by 15 matrix.

Initial 9 x 9 matrix

A	0	1	0	0	0	1	0	1
0	B	0	1	0	1	0	0	0
0	0	C	0	1	0	1	1	0
0	1	0	D	0	1	0	0	1
0	0	1	0	E	0	0	1	0
0	1	0	1	0	F	0	0	0
1	0	1	0	0	0	G	0	0
0	0	1	0	1	0	0	H	0
1	0	1	1	0	0	0	0	I

New matrix (now 15 x 15) with inserted subsystem

A	0	1	0	0	0	1	0	1	0	0	0	0	0	0
0	B	0	1	0	1	0	0	0	0	0	0	0	0	0
0	0	C	0	1	0	1	1	0	0	0	0	0	0	0
0	1	0	D	0	1	0	0	1	0	0	0	0	0	0
0	0	1	0	E	0	0	1	0	0	0	0	0	0	0
0	1	0	1	0	F	0	0	0	0	0	0	0	0	0
1	0	1	0	0	0	G	0	0	0	0	0	0	0	0
0	0	1	0	1	0	0	H	0	0	0	0	0	0	0
1	0	1	1	0	0	0	0	I	1	0	0	0	0	0
0	0	0	0	0	0	0	0	1	J	1	1	1	1	1
0	0	0	0	0	0	0	0	0	1	K	1	1	1	1
0	0	0	0	0	0	0	0	0	1	1	L	1	1	1
0	0	0	0	0	0	0	0	0	1	1	1	M	1	1
0	0	0	0	0	0	0	0	0	1	1	1	1	N	1
0	0	0	0	0	0	0	0	0	1	1	1	1	1	O

Figure 16. Starting Point to Incorporate a New Subsystem into an Existing System

The new system matrix shown on the right hand side of Figure 16 is reduced in size by compressing the subsystem represented by Objects J through O into one (1) cell, Object J. This subsystem compression activity reduces the system matrix model from a 15 by 15 to a 10 by 10 matrix model. This new disordered system model is shown on the left hand side of Figure 17. This disordered system matrix, when processed using the ART evolutionary computational technique, results in the ordered, compressed system model shown on the right hand side of Figure 17.

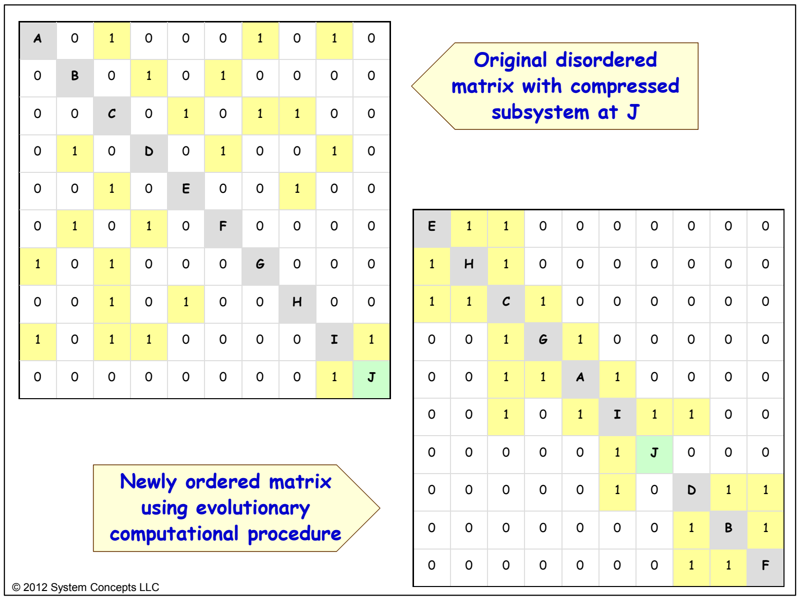

Figure 17. Reduced System Matrix Model in Both Disordered and Ordered Forms

Reducing the matrix size by compressing a subsystem directly reduces the computational complexity associated with the matrix calculations. Once the compressed system model has been ordered, it is then possible to apply a further series of analytical processes. The type of analytical process would depend on the domain of application, and the objectives of the system analysis.

One type of analytical process that could be applied is to expand the subsystem contained in the compressed matrix model. This expansion of the 10 by 10 matrix model back to a 15 by 15 model returns the matrix to the original level of representation, but in a more ordered configuration. This type of expansion and compression of the ordered matrix is shown in Figure 18.

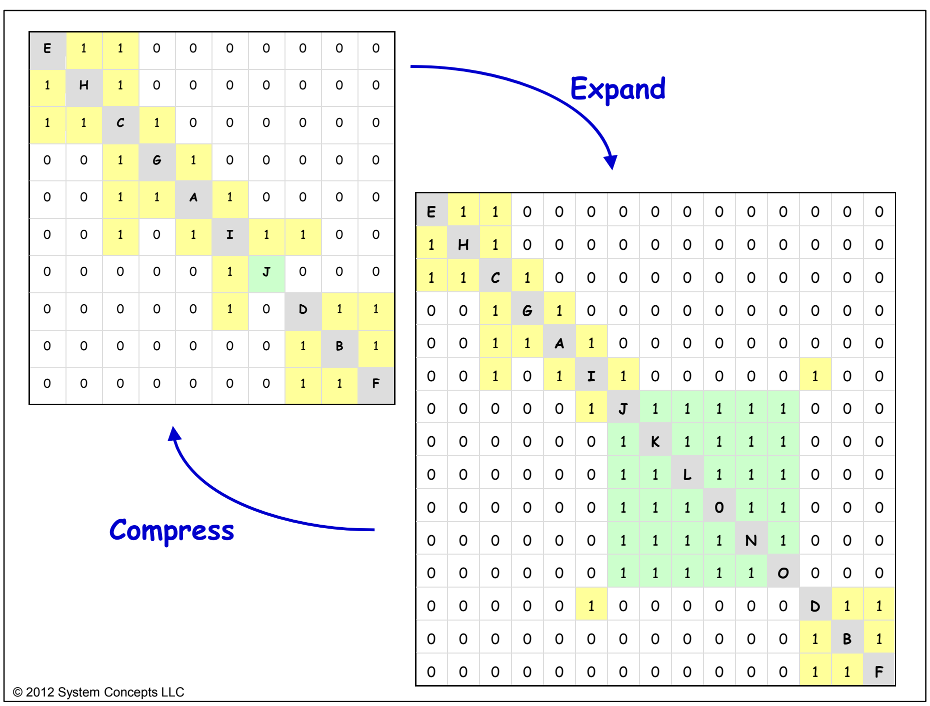

Figure 18. Expansion and Compression of the Subsystem Model

A second type of analytical process is the compression of the other subsystems contained in the reduced matrix model. This compression of Objects E, H, C and D, B, F into Object C and Object D further compresses the 10 by 10 system matrix model to a 6 by 6 system matrix model. The expansion and compression of the subsystems in the compressed matrix is shown in Figure 19. With two levels of subsystem compression, the matrix size has been reduced from 15 by 15 to 6 by 6. This process of subsystem compression and expansion is well defined, reversible, and substantially reduces both cognitive and computational complexity.

These types of analytical procedures may be interleaved, and applied in an adaptive manner to support the specific semantics of the application domain and objectives of the system analysis activity.

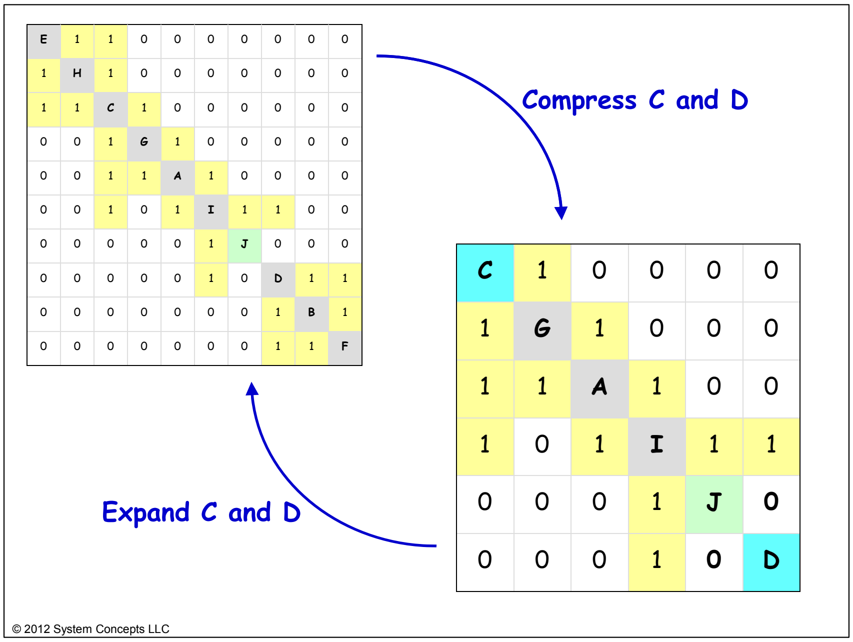

Figure 19. Compression and Expansion of the Reduced Matrix Model Subsystems

Summary and Conclusions

The authors have reestablished the line between basic structural models and interpretive structural models as a fundamental distinction needed to reduce the complexity associated with the analysis of system structures. Those interpretive structural models known as N Squared Charts and Automated N Squared Charts, are used as examples of interpretive structural models that have no clear, direct connection to basic structural mathematical models. In this and other work, the authors augment these interpretive structural models, with basic mathematical structural models.

The whole system metrics identified, developed and presented in this work support both the construction-rule and function-rule definitions of a system. These metrics demonstrate the integration of prose, graphics and mathematics in the practice of systems science and systems engineering. The relational score, associated with physical entropy, is applied to the relational interfaces between and among the system objects. The object score, associated with information entropy, is applied to the ordering of system objects on the matrix diagonal. The subsystem score provides a mechanism to evaluate domain specific system configurations as well as system constraint conditions. Two domain-independent numerical scores, the relational score and the object score, have been defined and their use demonstrated in the examples. The clear connections between and among the two system definitions, and the two numerical system entropy scores, provide a solid computational and graphic foundation upon which other analytical techniques can be based. The direct connection between the two parts of the system matrix representation (objects and interfaces), and the two system scores, facilitates their application and use.

The subsystem score is developed based on the combination, and/or integration, of the relational and object scores. The subsystem group score can be domain and application dependent, and is designed to be tailored and adjusted by domain subject matter experts to support the specific type of analysis needed for the current system evaluation task. The application and semantics of the subsystem score must be documented in a manner that clearly communicates the objective and outcome of the analysis. The prose component is designed to capture this written information, and tie it directly to the semantics of the graphical model, and the behavior of the executable mathematical model. The subsystem metric is a fundamental classification technique that transforms numerical-based data into prose-based information using rules that are domain dependent.

The three examples presented here provide a basic introduction to the metrics and their application. These methods further strengthen the relationship between prose system descriptions, graphical system descriptions and mathematical system descriptions. These foundational systems concepts provide the necessary basis for the production of more detailed automated system evaluation and analysis methods. The authors believe that the disciplined development and application of 'Abstract Relation Type' Automated N-Squared Chart techniques have the potential to reduce the computational and cognitive complexity associated with large-scale systems development projects. More focused research is needed to fully explore this area of system science and engineering.

References

[1] John N. Warfield, *Structuring Complex Systems. Monograph No. 4.*, Battelle Memorial Institute, Columbus, Ohio, 1974.
[2] Andrew P. Sage, *Methodology for Large Scale Systems*, McGraw-Hill Book Company, New York, New York, p. 92, 1977.
[3] Donald V. Steward, *Systems Analysis and Management: Structure, Strategy and Design*, PBI, A Petrocelli Book, New York, 1981.
[4] J.J. Simpson and M.J. Simpson, "Systems and Objects," In *Proceedings, Thirteenth Annual International Symposium of INCOSE, "Engineering Tomorrow's World Today,"* Crystal City, Virginia, July, 2003.
[5] Francis Heylighen, "(Meta)Systems as Constraints on Variation – A classification and natural history of metasystem transformations," In *Free University of Brussels Research Paper, sponsored by Belgian National Fund for Scientific Research (NFWO)*, Brussels, Belgium, 1994.
[6] John N. Warfield, *A Science of Generic Design*, Iowa State University Press, Ames, Iowa, 1990.
[7] J.J. Simpson and M.J. Simpson, "Formal Systems Concepts," *Proceedings, Fourth Annual Conference on Systems Engineering Research,"* Los Angeles, California, April, 2006.
[8] Donald V. Steward, "On an Approach to Techniques for the Analysis of the Structure of Large Systems of Equations," Society for Industrial and Applied Mathematics Review, Vol. 4, No 4, pp. 321-342, 1962.
[9] Donald V. Steward, "Partitioning and Tearing Systems of Equations," Journal of the Society for Industrial and Applied Mathematics: Series B, Numerical Analysis, Vol. 2, No. 2 (1965), pp. 345-365, 1965.
[10] R.J. Lano, *A Technique for Software and Systems Design*, TRW and North-Holland Publishing Company, Amsterdam, 1979.
[11] Derek K. Hitchins, Advanced Systems Thinking, Engineering and Management, Artech House, Inc., London, England, 2003.
[12] J.J. Simpson and M.J. Simpson, "System of Systems Complexity Identification and Control," *Proceedings, Proceedings, IEEE 4th System of Systems Engineering Conference,* Albuquerque, New Mexico, June, 2009.
[13] J.J. Simpson and M.J. Simpson, "A Pragmatic Complexity Framework," *Proceedings, INCOSE Spring 2009 Conference, "Systems Engineering - Affordable Success,"* Suffolk, Virginia, April, 2009.
[14] J.J. Simpson and M.J. Simpson, "Complexity Reduction: A Pragmatic Approach," *Systems Engineering Journal*, Volume 14, Issue 2, June 2011, John Wiley and Sons Ltd., Chichester, UK, June 2011.
[15] C.E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, Vol. 27, pp. 379–423, 623–656, July, October, 1948.