A Framework for Quantifying Complexity and Understanding its Sources: Application to two Large-Scale Systems

by

Pierre-Alain J. Y. Martin

Ingénieur diplômé de l'Ecole Polytechnique, Palaiseau, France, 2004

Submitted to the Engineering Systems Division
in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Technology and Policy

at the

Massachusetts Institute of Technology

September 2004

Signature of Author..
Pierre-Alain J. Y. Martin
Technology and Policy Program, Engineering Systems Division
August 6, 2004

Certified by..
Christopher L. Magee
Professor of the Practice of Mechanical Engineering and Engineering Systems
Director, CIPD
Thesis Supervisor

Accepted by..
Dava J. Newman
Professor of Aeronautics and Astronautics and Engineering Systems
Director, Technology and Policy Program

**A Framework for Quantifying Complexity and Understanding its Sources:
Application to two Large-Scale Systems**

by

Pierre-Alain J. Y. Martin

Submitted to the Engineering Systems Division on August 6, 2004
in Partial Fulfillment of the Requirements for
the Degree of Master of Science in Technology and Policy

ABSTRACT

The motivation for this work is to quantify the complexity of complex systems and to understand its sources. To study complexity, we develop a theoretical framework where the complex system of interest is embedded in a broader system: a complex large-scale system.
In order to understand and show how the complexity of the system is impacted by the complexity of its environment, three layers of complexity are defined: the internal complexity which is the complexity of the complex system itself, the external complexity which is the complexity of the environment of the system (i.e., the complexity of the large-scale system in which the system is embedded) and the interface complexity which is defined at the interface of the system and its environment.
For each complexity we suggest metrics and apply them to two examples. The examples of complex systems used are two surveillance radars: the first one is an Air Traffic Control radar, the second one is a maritime surveillance radar. The two large-scale systems in which the radars are embedded are therefore the air and the maritime transportation system.
The internal complexity metrics takes into account the number of links, the number of elements, the function and hierarchy of the elements. The interface complexity metric is based upon the information content of the probability of failure of the system as it is used in its environment. The External complexity metric deals with the risk configuration of large-scale systems emphasizing the reliability and the tendency to catastrophe of the system.
The complexity metrics calculated based on specific analysis of the ATC radar are higher than those calculated for the maritime radar for all the three levels of complexity indicating that the external complexity is the source for the internal complexity. Thus, not surprisingly it appears that the technical complexity of a system mainly stems from the socio-political complexity of the large-scale system in which it is embedded. More interestingly, the more rigorous and quantitative complexity metrics (Internal and Interface) are approximately linearly related for these two systems. This result is potentially important enough to be tested over a wider variety of complex systems.

Thesis Supervisor: Christopher L. Magee
Title: Professor of the Practice of Mechanical Engineering and Engineering Systems

ACKNOWLEDGMENTS

This research opportunity and the course of my studies at the Massachusetts Institute of Technology were supported by the MIT Center for Innovation in Product Development (CIPD) and THALES Research and Technology.

First of all, I would like to show my gratitude to my advisor Professor Christopher L. Magee who supported me throughout this research and taught me much of what follows in this thesis. He offered me both the freedom and the guidance I needed. I would like to thank him more specifically for his availability, his advice and his mentorship.

I also wish to thank very much M. Philippe Souchay of THALES Communications for the confidence he granted me and for the support he offered me. Without him, this thesis and my experience at MIT would not have been possible.

Many thanks to the numerous persons at THALES who, beside their busy schedule, have helped me to access the data I needed to validate the theoretical work. Thank you to: Josette Diesner, Jan Dijkstra, François Grégoire, Emmanuel Imbert, Alain Johann, Francis Lacoste, Jean-Luc Lambla, Theirry Leroy, Michel Lesage, Bernard Navarro, François Pache, Pierre Saulais, Adrian Smits, Albert Stegeman and Leo Vervoort.

Many thanks also to the persons who helped me to gather data: Jean Chapelon (ONISR), Colette Decamme (ONISR), Olivier Ferrante (BEA), Babak Khorrami (NCSA), Brice Martin-Castex (IMO), Sharron Morin (ICAO), Guilhem Nicolas (BEA), Joshua Summers (CREDO).

Finally, I wish to thank the staff and the students of the Center for Innovation in Product Development for the excellent working atmosphere and motivating dynamics they create at CIPD.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

This thesis which aims at quantifying complexity and understanding its sources is based on a framework which brings out three different complexities: internal complexity which is the complexity of the complex system studied, external complexity which is the complexity of the environment of the system and interface complexity which is defined at the interface of the system and its environment. Each of these three complexities are studied and quantified separately. Metrics are proposed and applied to two test bed systems: an Air Traffic Control (ATC) radar and a maritime surveillance radar. Finally, after analyzing the results, challenging the concepts and the methods, conclusions are drawn on the interplay of these complexities.

A. Framework to study Complexity

The first step to study complexity is the identification of boundaries. It is a major step because it is a prerequisite for computing complexity. It is also a hard step because boundaries are blurred, and then hard to define, due to the recursivity of systems. Indeed, every system can be regarded as a sub-system of a bigger system. The boundaries are also difficult to draw because engineering systems are permeable: they always have links that go from the outside to the inside. These trans-boundary links may justify the extension of the frontier of the system already defined to its connected elements.

To study how the complexity of a device is influenced by the complexity of its environment we propose a three-layered framework. The three sets applied to our study and identified in Figure 1 are: the complex system, the transportation system and the "Political" system. Applied to one of the test bed complex systems, the three sets are: the Air Traffic Control (ATC) radar, the air transportation system and the political system. The first set is a product: it is very focused and concrete. The second one is much broader but it still has material dominance: it may be seen as a socio-technical system. Conversely, the third set: the "Political" system is immaterial and may be seen in comparison as a socio-political system.

Conceptually, three kinds of complexity are identified: internal complexity, interface complexity and external complexity. In the example of the ATC radar, internal complexity is the complexity of the radar itself; external complexity is the complexity of the air transportation system and interface complexity is the complexity at the boundaries of the radar and its environment.

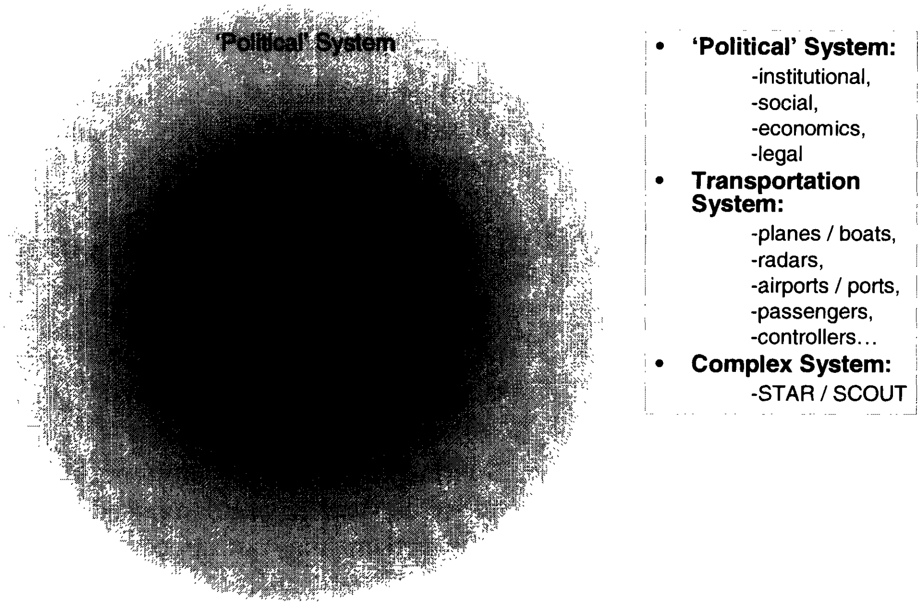


Figure 1 – Framework to study complexity (applied to our study)

This framework has been developed because it is quite natural and because it fits the purpose of this thesis. To validate the hypothesis that the external complexity impacts the internal complexity, the complexity of the radar must be computed. Thus, it is quite natural to isolate the radar as a set. It is all the more natural to do so because the conceptual boundary drawn corresponds to the physical boundary of the radar. The complexity of the environment of the system also must be computed. The environment is, by definition, what is outside of the system. The environment has well-defined inner boundaries, which are the boundaries with the radar, but it does not have well-defined exterior boundaries. The environment is open to the outside and it is divided into two sets: one where complexity is actually computed and one where it is not. So, to validate the hypothesis, the complexity of the air transportation system is computed and a correlation will be sought with the internal complexity described above. Nonetheless, the complexity of the "Political" system which cannot be computed is not ignored because we believe that it shapes the complexity of the air transport system. In this study, interface complexity is mainly a tool to analyze the relation between external and internal complexity. It is also natural to study it because it corresponds to the well-defined boundary between the radar and its environment.

In this thesis, the name of the three complexity metrics will be capitalized while the name of the concepts will not: Internal Complexity means the metric and internal complexity means the concept.

Further discussions on the framework are presented in E.I.2 (p. 62).

B. Internal Complexity

This section on internal complexity defines the terminology used before describing a methodology to represent the internal complexity of complex systems. 48 possibly useful Internal Complexity metrics are defined and then applied to a detailed example in order to determine which one of these is most robust and potentially useful. Finally, the internal complexity of two test bed systems is computed with the metric identified.

I. Definitions and initial approach to internal complexity

1. Definitions

- System

The working definition of system, which is consistent with the Engineering Systems Division definition [1], is the one used by C. Magee and O. de Weck [2]. A system is: "a set of interacting components having well-defined (although possibly poorly understood) behaviors or purposes; the concept is recursive as systems are composed of other lower-level systems. Thus what is a system to one person may not appear to be a system to another."

In this thesis, a system is seen as a set of layers where each layer is a level of decomposition representing a further decomposition of the layer just above it. Each subsystem of a given layer is represented as further decomposed into "smaller" subsystems (in the layer just below).

In this thesis, the term "high" is used to describe a layer at the top of the decomposition of the system (i.e., with a small level number) and the terms "low" or "deep" to describe a layer at the bottom of the decomposition of the system (i.e., with a higher level number).

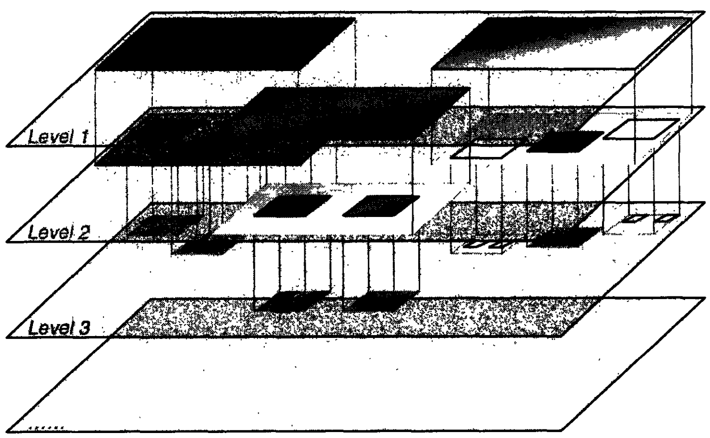


Figure 2 – System model

- Element

An element is a conceptual entity. It can be material or immaterial but it must achieve a purpose. In this paper, by the term "purpose" we mean: to represent a function. In fact, elements are operators and we distinguish five basic operations [2]:

  - Transform or Process,
  - Transport or Distribute,
  - Store or House,
  - Exchange or Trade and
  - Control or Regulate.

These five operations (also called basic functions) are a generator set of every function: every function performed by a given element is a combination of some of these basic functions.

Due to the recursive definition of systems, a system can be regarded as a subsystem or even as an element. However, generally speaking, the term "system" is used for the bulk device while "subsystem" is used at the first levels of decomposition and "element" at the deeper levels.

- Link

In contrast with elements which are entities, links are flows. Links can be flows of [2]:

  - Matter,
  - Energy,
  - Information or
  - Value.

Thus, links are connectors between elements in the sense that they are the operands on which the elements (operators) operate. Since links represent flows, they can be either unidirectional or bidirectional.

- Complex system

A complex system is "a system with numerous components and interconnections, interactions or interdependencies which are difficult to describe, understand, predict, manage, design, and/or change." [2]

2. Initial approach

a. Characterization of internal complexity

Complex can be defined as: "composed of interconnected or interwoven parts" [3] and *complexity* as "a quality of an object with many interwoven elements, aspects, details, or attributes that makes the whole object difficult to understand in a collective sense" [4].
Even if complexity is a macroscopic propriety of a system, its origins are microscopic: complexity arises both from quantity and quality.

Basic observations of complex systems tend to lead us to the immediate conclusion that: the more, the more complex. This idea is withstood by the definition of complexity which states that complexity is the "quality of an object with <u>many</u> interwoven elements..." Some people tend to reject this feature: they claim that quantity is inaccurate and even wrong because it is too simplistic. We agree that quantity cannot alone represent complexity. Nevertheless, quantity plays a major role in complexity even if it needs to be qualified with other features.

Besides, many authors emphasize the role of quantity in complexity. For instance, Sussman [5] regards complex systems as" Complex, Large-scale, Integrated, Open Systems (CLIOS)" and engineering systems as technical CLIOS. "Large-scale" is highly correlated with our idea of quantity. "Integration" is also correlated with quantity because such feature denotes high density: density coupled with large-scale refers to quantity.

Many attempts to quantify complexity are based on an analogy with entropy. The best example is the Shannon entropy [6]. Now, entropy is a holistic measure of the microscopic state of a huge number of elements. Once more, this emphasizes the major role played by quantity in complexity.

If we acknowledge that complexity can be regarded as uncertainty [7], it appears once again that: increase in quantity leads to more complexity. Indeed, the more parameters that exist lead to more possibilities and thus more uncertainty.

Complexity can also be seen as the probability of success of achieving the functional requirements [8]. The more interconnected elements you need to assembly in order to achieve the requirement, the higher the probability of failure and therefore the higher the complexity.

Referring to the definition of complexity as "a quality of an object with many *interwoven* elements", it clearly appears that complexity has at least two attributes which are the elements and the links. Understanding the quantity of interactions that occur within a system is also a key to quantify complexity.

As a conclusion, both the quantity of elements and the quantity of their interactions matters in complexity.

Quality is another aspect of complexity. Here, by "quality" we mean both: specificity and intensity. These two notions also apply to elements and links.

The particularity of the elements is accurately described by the set of five basic functions (vocabulary) used to characterize every function performed by an element. The intensity of the elements is accurately described by the layers (hierarchy) where the relevant elements are in the system decomposition. Hierarchy and vocabulary are the two dimensions which we propose to use to give a good insight into the quality of the elements. They assess complexity describing both the function of the elements (vocabulary) and their strength (hierarchy).

Up to a certain extent, hierarchy also determines the quality (intensity and specificity) of the links because it gives both an idea of their strength and their spatial distribution.

As noted relative to quantity, the quality of elements and the quality of their interactions matters in complexity. We deeply believe that hierarchy is the best way to describe the qualitative aspect of complexity. This statement will be detailed in B.III.1 (p. 18).

Quantity and quality need to be combined appropriately to accurately describe complexity.

b. Characterization of Internal Complexity metrics

Following Joshua D. Summers and Jami J. Shah [9] we divide complexity into two components: Scale Complexity ($\mathcal{C}_S$) and Link Complexity ($\mathcal{C}_L$). These two dimensions of complexity are studied separately because they allow different insights into the overall internal complexity.

- Scale Complexity

Summers and Shah [9] identify "size" as the first of the "three fundamental aspects of complexity". Our formulation of Scale Complexity takes into account the number of elements to describe the "horizontal" size, the number of levels to describe the "vertical" size and the number of basic functions to account for the "functional" complexity.

- Link Complexity

Summers and Shah [9] identify "degree of coupling" as the second of the "three fundamental aspects of complexity". In our formulation, Link Complexity takes into account the number of links and the number of elements because it is believed that the density of links or the connectivity of elements directly infers the degree of coupling.

Neither Scale Complexity nor Link Complexity alone can give a balanced assessment of complexity because they describe different features of complexity. The two components of complexity need to be combined. Besides, the definition of complexity used in this paper also apposes these two components when it states that complexity is the "quality of an object with many interwoven elements": "many" referring to scale and interwoven to links.

Braha and Maimon [8] adapt two concepts from software complexity: *structural complexity* and *functional complexity*. Both complexities are functions of the *information content* of the design. *Structural complexity* is the complexity that is based on the representation of the information. The appeal of this complexity is that its valuation is facilitated using decomposition diagrams that describe the elements and interfaces of the system. Here, Scale Complexity which takes into account elements, hierarchy and functions refers both to *structural* and *functional* complexity. But a major aspect of structure i.e., the links between the elements is missing in this approach to structural complexity. The coupling is brought by Link Complexity which refers to structure.

Our approach combines both components of complexity: Scale Complexity ($\mathcal{C}_S$) and Link Complexity ($\mathcal{C}_L$) in an attempt to obtain an overall assessment of complexity.

II. Representation of complex systems

1. Definition of the Reference Decomposition

The Reference Decomposition is a representation of systems which shows interconnected *mono-functional elements* belonging to different levels of decomposition. This representation will be used in this thesis to assess and to compute the Internal Complexity of a complex system.

2. Methodology to obtain the Reference Decomposition

The decomposition process to obtain the Reference Decomposition is a top-down recursive process illustrated in Figure 3.

- We start at the first level of decomposition of the system (level 1).
- At any given level of decomposition (i.e., a layer) of the system we identify each subsystem with the function it performs. For each element, either of the two cases can occur.
  - The element performs only one of the five basic functions and the decomposition process does not go further for this element. This element is then represented in the Reference Decomposition with the basic function it achieves.
  - The element performs a more complex function (i.e., a function described with more than one of the five basic functions). Then the decomposition process for this element goes one step further to the next level of decomposition.
- We apply the process to each subsystem until every element achieves *only one basic function.*
- We connect all the elements that appear in the Reference Decomposition with the appropriate links attributing to them their direction (directional or bidirectional) and the substance flowing (matter, energy, information or value).

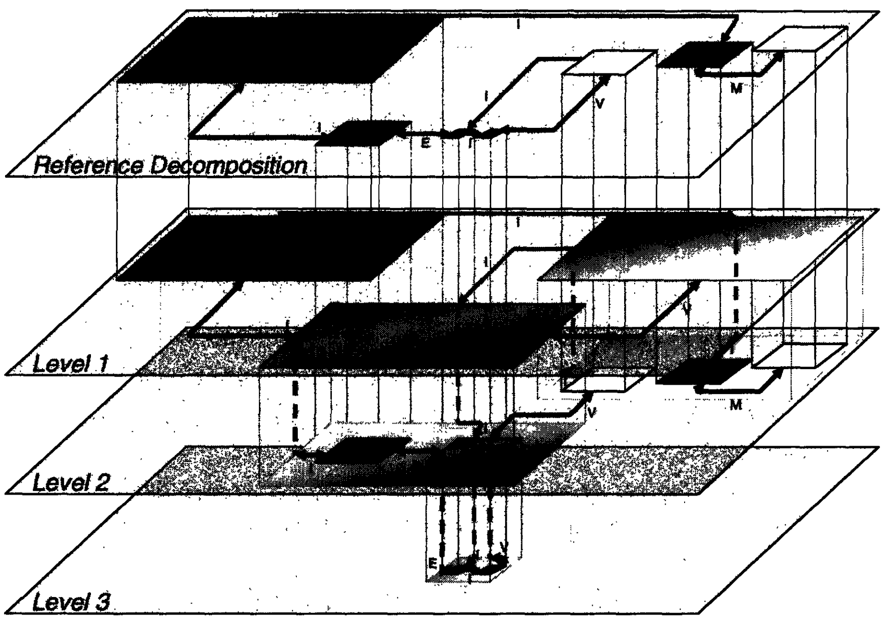
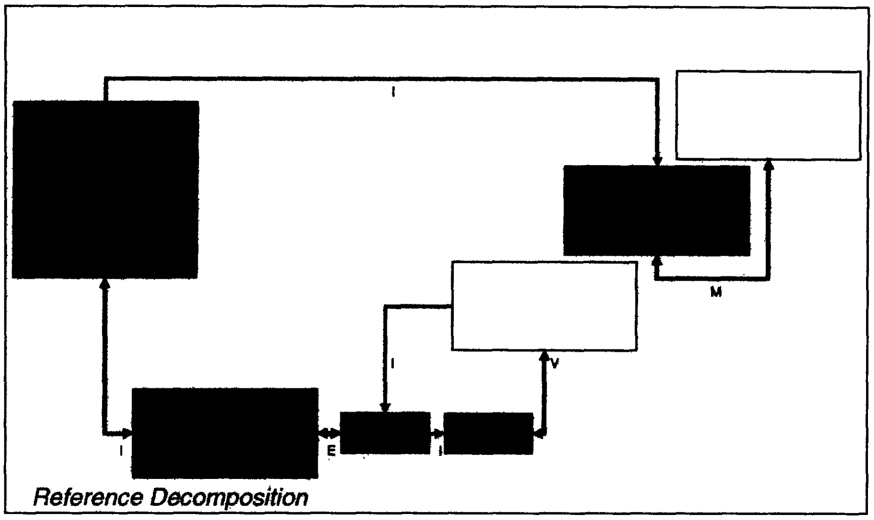


Figure 3 – Illustration of the methodology to obtain the Reference Decomposition

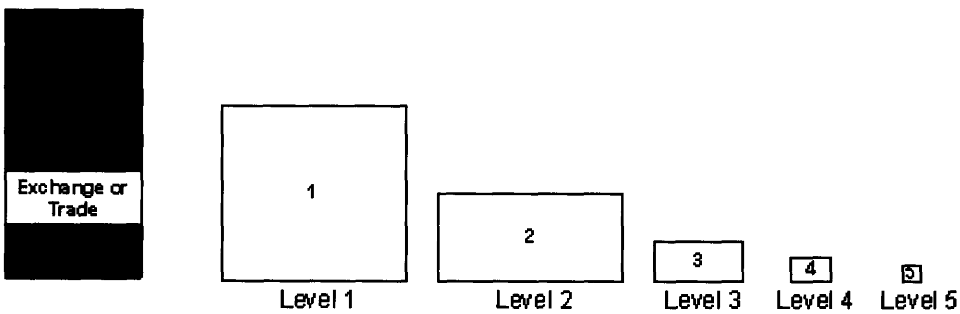
The outcome of this process is the Reference Decomposition (Figure 4). It is a map of interconnected mono-functional elements.

Figure 4 – Illustration of the Reference Decomposition



3. Assessment of complexity through the Reference Decomposition

In the representation of the Reference Decomposition, elements are represented with boxes. The size of the box represents the hierarchy of the element (i.e., the layer to which it belongs). The higher the layer, the bigger the box; it is as if you were looking at the system from above. The color of the box represents the operation performed by the element. A different color is associated with each basic operation; so a multicolor box represents a multifunctional element.



In the representation, links are represented with arrows of two different colors depending upon their directionality.

The substance flowing stands by the arrow in a red capital letter to further qualify a link: M for matter, E for energy, I for information and V for value.

As a summary, we believe that the Reference Decomposition gives a good insight into the complexity of the system since it brings out the main features of complexity: quantity and quality thanks to the number of elements, their hierarchy (box size), the number of functions performed (color) and the coupling (links).
Moreover, this representation of a system as if it were seen from the top elucidates the sources of emergent behavior. The more small connected boxes of different colors, the more complex: the behavior of the system tends to be more complex if it requires the conjunction of many intertwined and deeply integrated basic functions.
The Reference Decomposition is thus a useful representation of complexity.

III. Internal Complexity metrics

1. Drivers

In this part, we present the four main drivers of complexity that need to be taken into consideration in Internal Complexity metrics.

Quite obviously, a main driver of complexity is the number of **elements**. Y. Bar-Yam [10] grounding his thoughts on complexity science identifies "elements (and their number)" as the first "central property of complex systems".

Y. Bar-Yam [10] also identifies "interactions (and their strength)" as the second "central propriety of complex systems". **Links** are also a critical driver of complexity.

The quantification of the number of elements needs to be further qualified to assess complexity. We believe that the function performed by an element must be taken into consideration to compute complexity. Functionality is an interesting feature of complexity and it can be compared to another "central property of complex systems" "activities (and their objectives)" [10]. **Vocabulary** (i.e., the number of basic functions in a system) is another driver of complexity.

As for vocabulary, **hierarchy** relates the "central property of complex systems" which Y. Bar-Yam [10] calls "diversity and variability". Behind the idea of hierarchy, there are two different important notions. The first one is the number of layers and the second one the depth of the layers. These two dimensions are proposed in this work to be factors increasing complexity. Since, there are several ways to understand the importance of hierarchy in complexity, some of the most relevant views on the role of hierarchy in complexity are presented here.
First, it is noted that hierarchy is the way human systems tend to organize to better manage themselves. When systems become too complex, they adopt a hierarchical structure in order to be more manageable. This organization reduces complexity and allows people to understand the system. Nonetheless, it is to be noted that this organization may increase apparent understanding but it may not decrease the underlying

complexity. A system with many different layers tends to denote higher underlying complexity. Joel Moses [11] highlights that hierarchy matters when he writes about the hierarchical decomposition in the Bible, the Greek or Middle Ages philosophies...What is true for "human" systems also seems to be true for complex systems.

Kauffman [12] contends that self-organization is a great undiscovered principle of nature. He claims that complexity itself triggers self-organization (and that if enough different molecules pass a certain threshold of complexity, they begin to self-organize into a new entity: a living cell). Kauffman extends his biological paradigm to economic and cultural systems, showing that they may evolve according to similar laws. We believe that complex engineering systems tend to be (self)-organized: the number and the depth of their hierarchical levels reflects their complexity.

Herbert Simon [13] proposed that complex systems only arise as hierarchical combinations of sub-systems.

Joel Moses [11] also details three approaches to design large-scale engineering systems and underlines the importance of "hierarchical decomposition" and "layered design" to cope with complexity in many fields such as the "Human Mind and Body", "Mathematics", "Philosophy", "Manufacturing"... So, it appears that the effort to design a device that achieves a complex behavior is more challenging if the Reference Decomposition shows deeper basic functions (and more layers).

Based upon the former statements, the reasons for including hierarchy in Internal Complexity metrics are now detailed.

An element that combines two or more basic functions is more complex than one with any single basic function. We assess this by representing the system as a set of mono-functional elements, in which it is necessary to go deeper in the decomposition to represent a multifunctional element than to represent a mono-functional one. Since a system made of multifunctional elements is arguably more complex, a system that has deeper layers (i.e., whose hierarchy is higher) is also similarly more complex.

The systemic emergent behavior (i.e., the complex function achieved by the system which may not be fully understood) is the combination of many basic functions. The deeper the basic functions in the decomposition, the more potentially complex the behavior. Indeed, these basic functions need to be combined and integrated more intricately to achieve their part of the resulting behavior than functions which are at a higher level. Since the set of five basic functions is limited and straightforward, every complex function needs to be represented by a complex combination of these five basic functions. The overall emergent behavior of a system is key in complexity. This combined behavior has often been labeled "emergent" and in the sense of combining mono-functional elements from various levels, such description is appropriate. Now, this behavior being the integration on different levels of all the basic functions of the system, the number of levels needed to achieve single-function representation is an important indicator of complexity.

Complexity correlates with depth. In the Reference Decomposition, the higher an element, the more important its basic function in the system. Conversely, when an element is broken further down to discover its basic functions, there are generally more elements in the Reference Decomposition. These elements have lesser influence because they belong to inferior layers and because it is only together that they can achieve a

function as important as the one achieved by elements of the layer above. Complexity is higher when hierarchy increases because it is the sum of many smaller influences that generates the resulting behavior.
The latter idea can be brought into closer alignment with Linda Beckerman's thoughts on emergent behavior [14]. She identifies system's goals as the emerging sum of functions and functions as the emerging sum of characteristics and techniques implemented to achieve them. From small interacting pieces at deep levels one can "build" a complex system with its emergent behavior.

To conclude on a quantitative basis, the drivers of complexity are: the number of mono-functional elements, their hierarchy (vertical position), the vocabulary (number of basic functions in the system) and the links between these elements. We will now pursue quantitative approaches to describe these effects.

2. Metrics

a. Characteristics

In this thesis, we define three functions and four variables.

Functions

C: Internal Complexity (also noted $\mathrm{C_{INT}}$ mainly in comparison to $\mathrm{C_{EXT}}$ and $\mathrm{C_{INTERFACE}}$)

$\mathrm{C_S}$: Scale Complexity

$\mathrm{C_L}$: Link Complexity

Variables

E: number of Elements

L: number of Links (directional links count as one link and bidirectional links as two)

V: Vocabulary (number of basic functions)

H: Hierarchy (level of decomposition)

The Reference Decomposition which succinctly summarizes the drivers of complexity is the basis to assess the influence of the four variables. Algorithmic complexity [15] argues that the length of the shortest algorithm which fully describes the artifact is the complexity of the artifact. Thus, the "length" of the Reference Decomposition infers complexity.

Complexity is a function of the number of elements, simple links, basic functions and hierarchy as they appear in the Reference Decomposition. It is noted:

$$\mathrm{C}(E, L, V, H): \quad \begin{array}{c} \aleph^4 \rightarrow \Re \\ (E, L, V, H) \rightarrow \mathrm{C}(E, L, V, H). \end{array}$$

Complexity is a dimensionless number.

b. Components

i. Scale Complexity

From the preceding arguments, Scale complexity is derived from Shannon entropy except that the number of elements (E) is *modulated* both by the number of functions necessary to describe the system (V) and the number of the highest level in the system decomposition (H_M) in order to give appropriate attention to quality as well as quantity. Scale Complexity increases faster with the number of elements (E) than with the two other variables V and H_M in order to emphasize the prevailing role of "quantity" in complexity.

Thus, the basic form for Scale Complexity is:

$$E \times \mathrm{Ln}\,(V \times H_M)$$

This expression is very close to the "size complexity" based upon information content defined by Summers and Shah [9]. In a similar expression, their complexity "includes both the total number of primitive modules (variables) and the total number of possible relation between these modules". In the formulation here the number of primitive modules is obviously E and the number of qualitative relations between them is obtained by multiplying V and H_M (because they are independent dimensions).

Since Scale Complexity should not be null when the system is only composed of mono-functional elements achieving one function (V = 1) in the first level (H_M = 1), it is more appropriately described by:

$$E \times \mathrm{Ln}\,((V + 1) \times (H_M + 1))$$

Finally, Scale Complexity is normalized so that it equals the number of elements when they all achieve the same function (V = 1) in the first level (H_M = 1). After normalization, the algorithm to compute Scale Complexity is:

$$C_s\,(E, V, H) = E \times \frac{\mathrm{Ln}\,((V+1) \times (H_M+1))}{\mathrm{Ln}\,(4)}$$

This algorithm allows one to compute Scale Complexity from a macroscopic point of view.

A second way to compute Scale Complexity is to compute the complexity of each level of decomposition, with the previous metric, and then to sum the different values to obtain Scale Complexity. This microscopic way of computing Scale Complexity emphasizes the role of the number of layers in complexity.

$$C_s\,(E, V, H) = \sum_j E_j \frac{\mathrm{Ln}\,((V_j+1) \times (H_j+1))}{\mathrm{Ln}\,(4)} \quad \text{(j: the level)}$$

Here, for each level j, E_j is the number of elements, V_j the vocabulary and H_j the hierarchy of the level (i.e., the number of the level: $H_j = j$).

Moreover, each of the two variants described above can be further divided into three sub-metrics depending upon the element counting procedure. Up to this point, E or E_j represent all the elements (in general or in a layer). It seems logical to explore alternative ways of taking the elements into account. First, we only count the number of non-redundant elements (E^{NR} and E_j^{NR}) or second only the number of non-identical elements (E^{NI} and E_j^{NI}). The reduction to the non-redundant elements is a logical way of proceeding because we can argue that doubling an existing element will not add Scale Complexity. The further reduction to the non-identical elements is also coherent because we can argue that adding an element which already exists in the system and performs the same function may not add significant Scale Complexity. However, we should note that adding these elements will affect Link Complexity (because they need to be connected) and therefore the overall Internal Complexity of the system.

As a conclusion, six algorithms to compute Scale Complexity ($\mathbf{c_s}$) are proposed: macroscopic versus microscopic and counting all the elements, only the non-redundant ones or only the non-identical ones. Table 1 recalls and names these six algorithms.

	Macroscopic	Microscopic
All the elements	(MACRO, All)	(micro, All)
Non-redundant elements	(MACRO, NR)	(micro, NR)
Non-identical elements	(MACRO, NI)	(micro, NI)

Table 1 – Six useful Scale Complexity algorithms

To better understand the behavior of the various Scale Complexity ($\mathbf{c_s}$) metrics it is worth determining whether each is intensive (independent of the size of the system) or extensive (depend on the size of the system).

- $\mathbf{c_s}(E, V, H) = E \times \frac{\mathrm{Ln}((V+1) \times (H_M+1))}{\mathrm{Ln}(4)}$ is **extensive**

Indeed, $\mathbf{c_s}(\text{system S} \cup \text{system S}) = \mathbf{c_s}(2E, V, H) = 2E \times \frac{\mathrm{Ln}((V+1) \times (H_M+1))}{\mathrm{Ln}(4)}$

$$= E \times \frac{\mathrm{Ln}((V+1) \times (H_M+1))}{\mathrm{Ln}(4)} + E \times \frac{\mathrm{Ln}((V+1) \times (H_M+1))}{\mathrm{Ln}(4)}$$

$= \mathbf{c_s}(E, V, H) + \mathbf{c_s}(E, V, H) = \mathbf{c_s}(\text{system S}) + \mathbf{c_s}(\text{system S})$

So, $\mathbf{c_s}(\text{system S} \cup \text{system S}) = \mathbf{c_s}(\text{system S}) + \mathbf{c_s}(\text{system S})$

- $\mathbf{c_s}(E, V, H) = \sum_j E_j \frac{\mathrm{Ln}((V_j+1) \times (H_j+1))}{\mathrm{Ln}(4)}$ (j: the level) is **extensive**

Indeed, $\mathbf{c_s}(\text{system S} \cup \text{system S}) = \mathbf{c_s}(2E, V, H) = \sum_j 2E_j \frac{\mathrm{Ln}((V_j+1) \times (H_j+1))}{\mathrm{Ln}(4)}$

$$= 2 \sum_j E_j \frac{\mathrm{Ln}((V_j+1) \times (H_j+1))}{\mathrm{Ln}(4)}$$

$= \mathbf{c_s}(\text{system S}) + \mathbf{c_s}(\text{system S})$

So, $\mathbf{c_s}(\text{system S} \cup \text{system S}) = \mathbf{c_s}(\text{system S}) + \mathbf{c_s}(\text{system S})$

• Scale Complexity ($\mathbf{C_s}$) becomes **intensive** when we consider the two <u>alternatives</u>: NR and NI because the elements are only counted once when two identical systems are juxtaposed. So, the following four metrics:

$$E^{NR}\frac{Ln((V+1)\times(H+1))}{Ln(4)},\ E^{NR}\frac{Ln((V+1)\times(H+1))}{Ln(4)},\ \sum_j E_j^{NR}\frac{Ln((V_j+1)\times(H_j+1))}{Ln(4)},$$

$$\sum_j E_j^{NR}\frac{Ln((V_j+1)\times(H_j+1))}{Ln(4)}$$ are **intensive**.

ii. Link Complexity

Firstly, Link Complexity can be conceptualized as the density of interactions.

The number of interactions between E elements all connected to one another is:

$$\frac{E\times(E-1)}{2}$$

Since these connections can be double links, the number of simple links between E elements all connected to one another is:

$$E\times(E-1)$$

Thus the number of simple links varies from 0 to E(E-1):

$$L\in[0\,,\,E(E-1)]$$

So, the density of links is:

$$\mathbf{D}\,(E, L)=\frac{L}{E\times(E-1)}$$

Then, regarding Link Complexity ($\mathbf{C_L}$) as a density, we compute it as follow:

$$\mathbf{C_L}\,(E, L)=\frac{L}{E\times(E-1)}$$

A variant can be the normalized density where Link Complexity equals 1 for a system having each of its elements linked once and only once.

Since $\mathbf{D}\,(E_0, L_0)=\frac{L_0}{E_0\times(E_0-1)}=\frac{1}{E_0-1}$ (because $L_0=E_0$),

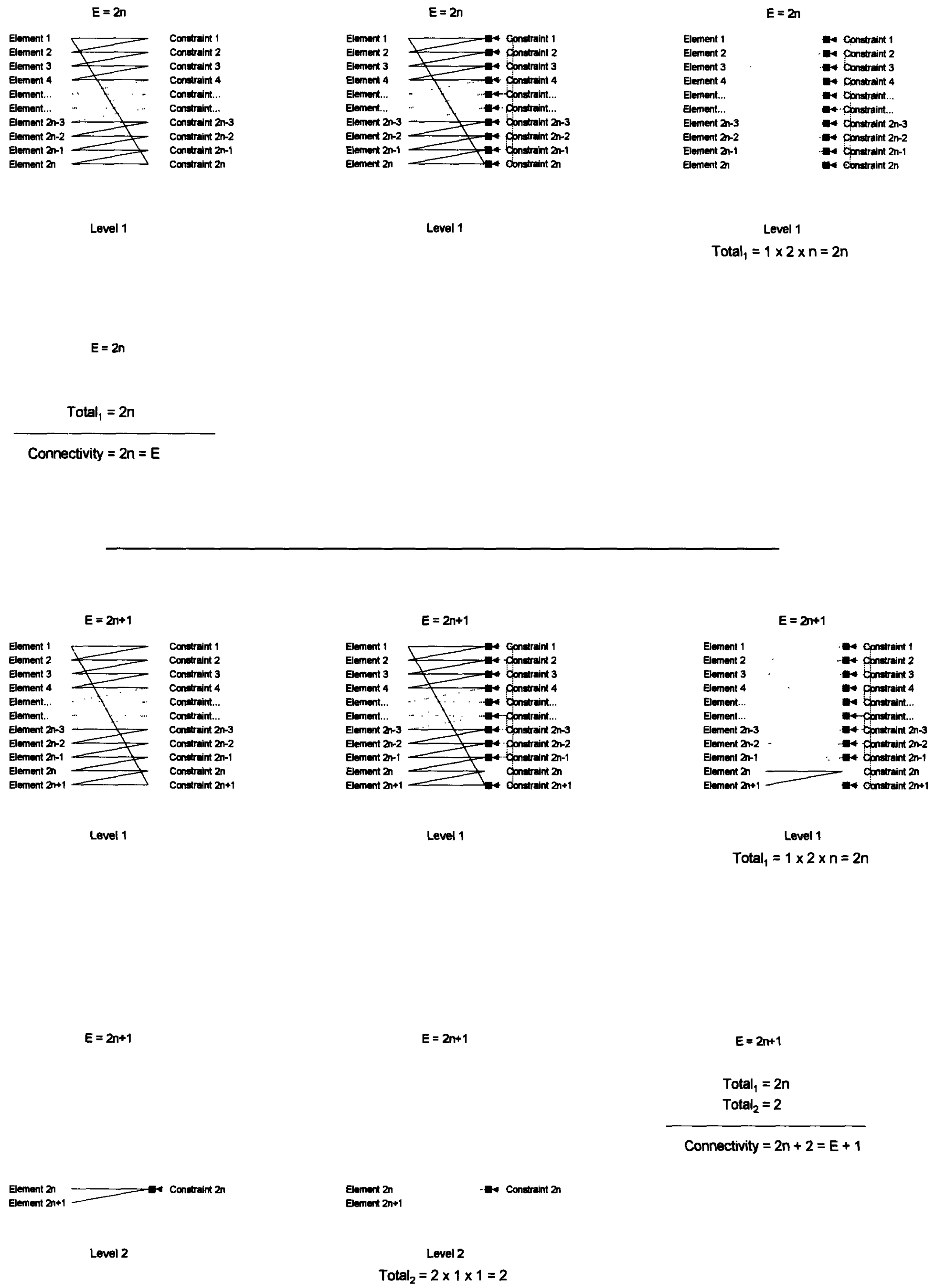
it finally comes that the Link Complexity ($\mathbf{C_L}$) is:

$$\mathbf{C_L}\,(E, L)=\frac{L}{E}$$

Secondly, Link Complexity can be regarded as Connectivity (as defined and explained by Summers [16]).

As for density, the ultimate variant is normalized Connectivity where Link Complexity equals 1 for a system having each of its elements linked once and only once. This algorithm for Link Complexity is obtained for the former by dividing it by the number of elements (E) when it is even or by the number of elements plus one (E+1) if it is not. Figure 5 illustrates the derivation of this result.

Figure 5 – Normalization of Connectivity



As a conclusion, four algorithms to compute Link Complexity (C_L) are proposed: Density normalized or not and Connectivity normalized or not. Table 2 summarizes and names these four algorithms.

	Density	Connectivity
Non normalized	(1)	(Cty)
Normalized	(2)	(Cty')

Table 2 – Four useful Link Complexity algorithms

As for Scale Complexity, it is important to determine whether Link Complexity (C_L) is intensive or extensive.

- $C_L(E, L) = \dfrac{L}{E \times (E-1)}$ is **extensive** (but not proportional to the size of the system)
- $C_L(E, L) = \dfrac{L}{E}$ is **intensive**

Indeed, $C_L(\text{system S} \cup \text{system S}) = C_L(2L, 2E) = \dfrac{2L}{2E} = \dfrac{L}{E} = C_L(L,E) = C_L(\text{system S})$

So, $C_L(\text{system S} \cup \text{system S}) = C_L(\text{system S})$

- $C_L(E, L)$ defined as Connectivity is **extensive** because the algorithm to compute connectivity is applied independently to the two juxtaposed systems.

- $C_L(E, L)$ defined as normalized Connectivity is **intensive** <u>only</u> when the number of elements is even. The non-normalized Connectivity of two juxtaposed systems which is twice the Connectivity of the one system divided by 2E equals the Connectivity of the system S divided by E (but not by E+1).

c. Configuration of the metrics

Since the two components of complexity previously defined emphasize different aspects of complexity that we want to appropriately combine, we propose two metrics to compute Internal Complexity (C) combining these components differently. The metrics proposed are fully ordered.

i. Norm

Scale (C_s) and Link (C_L) Complexities can be regarded as the two components of a vector in an orthonormal base. Complexity (C) is thus the norm of this vector:

$$C = (C_s^2 + C_L^2)^{1/2}$$

Now, we apply the 48 metrics identified in Table 3 to STAR 2000 in order to select the most appropriate one. Then, we also apply this metrics to SCOUT in order to compare both complex systems.

1. Application to the ATC radar

The ATC radar STAR 2000 is a high performance, fail safe, affordable S-band primary radar designed to deal with dense air traffic situations, within approach or extended approach control area. STAR 2000 is a pulsed radar that supports reduced separation between aircraft and features high processing capacity.

The level 1 of decomposition of STAR 2000 is represented in Figure 6 with the convention we use for the decomposition (B.II.3 p. 17).

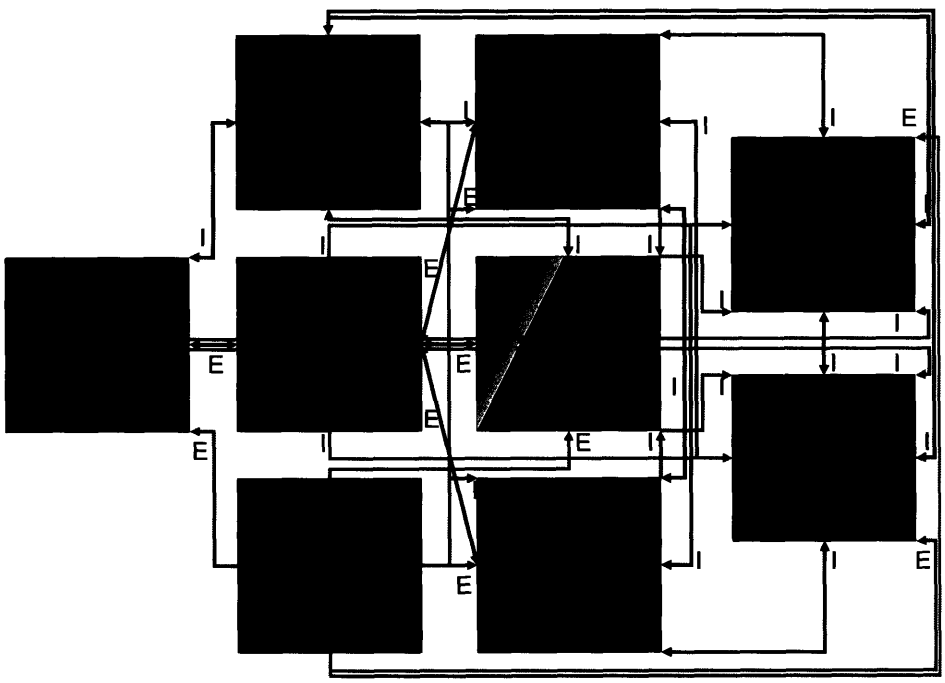
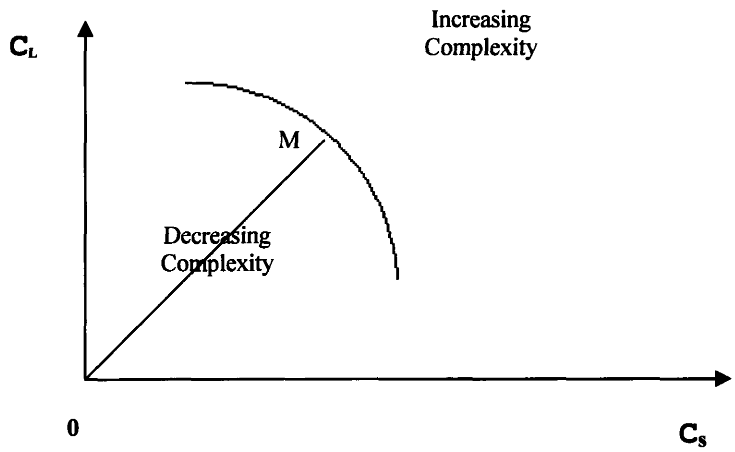


Figure 6 – Level 1 of decomposition of STAR 2000

STAR 2000 is composed of 9 subsystems:

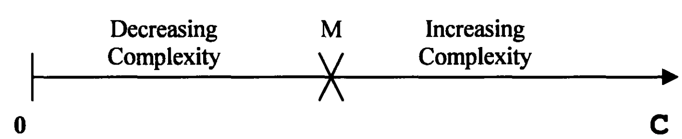
- Aerial System
- RCMS (Remote Control and Monitoring System)
- MWA (Microwave Assembly)
- AE 2000 (Main Distribution)
- GRA A (Generation Reception Assembly)
- GRA B (Generation Reception Assembly)
- SST(8) (Solid-State Transmitter)
- TR-2000 A (Aircraft and Weather Processor)
- TR-2000 B (Aircraft Processor)



ii. Product

We can also access to complexity by multiplying the two components.

$$C = C_s \times C_L$$



As a conclusion, Table 3 summarizes the 48 different possibilities to compute Internal Complexity.

C		Norm						Product					
C_s		MACRO			micro			MACRO			micro		
		All	NR	NI	All	NR	NI	All	NR	NI	All	NR	NI
C_L	(1)												
	(2)												
	(Cty)												
	(Cty')												

Table 3 – The 48 possible Complexity metrics

IV. Applications and choice of the relevant metrics

In this section we compute the Internal Complexity of two systems: STAR 2000 and SCOUT. Both systems are radars used in different transportation systems. STAR 2000 is used in Air Transportation System while SCOUT is used in the Maritime Transportation System. We choose these two test bed systems because they are complex systems which both have commonalities (because they are both radars) and differences (different environment and principles (pulse versus continuous wave)). Therefore, their comparison is more direct and thus potentially valuable for studying the sources of complexity.

Two of these subsystems (RCMS and AE 2000) are mono-functional and will not be further decomposed to obtain the Reference Decomposition and to compute Internal Complexity. One subsystem (MWA) is bi-functional, another one (SST(8)) is tri-functional and the five others (Aerial System, GRA A and B, TR-2000 A and B) are tetra-functional. All these seven subsystems will need to be further decomposed in order to compute Complexity.

a. The Reference Decomposition

First, we illustrate the methodology to obtain the Reference Decomposition on the example of the Aerial System, a subsystem of STAR 2000 (i.e., an element which belongs to the first level (level 1) of STAR 2000 decomposition).
Figure 7 represents the first level of the decomposition of the subsystem "Aerial System" which is only the part of the level 2 of STAR 2000 decomposition focusing on the Aerial System. The colored rectangles represent the elements of this layer and the diamonds represent the other subsystems of STAR 2000 liked with the Aerial System (Main Distribution Unit AE 2000, MWA 2000 S, RCMS...) which belong to the level 1 of decomposition and will not be decomposed in this example because we only focus on the Aerial System.

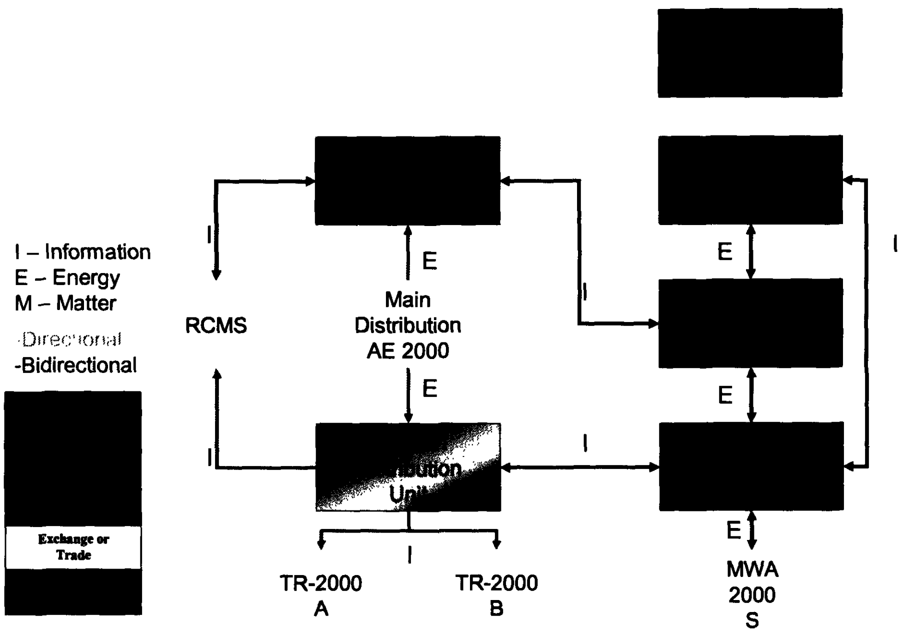


Figure 7 – Aerial System: level 2 of decomposition of STAR 2000

Figure 8 does not add any information: it is the same representation as Figure 7 but with bigger boxes to make the representation of the further decomposition easier to understand.

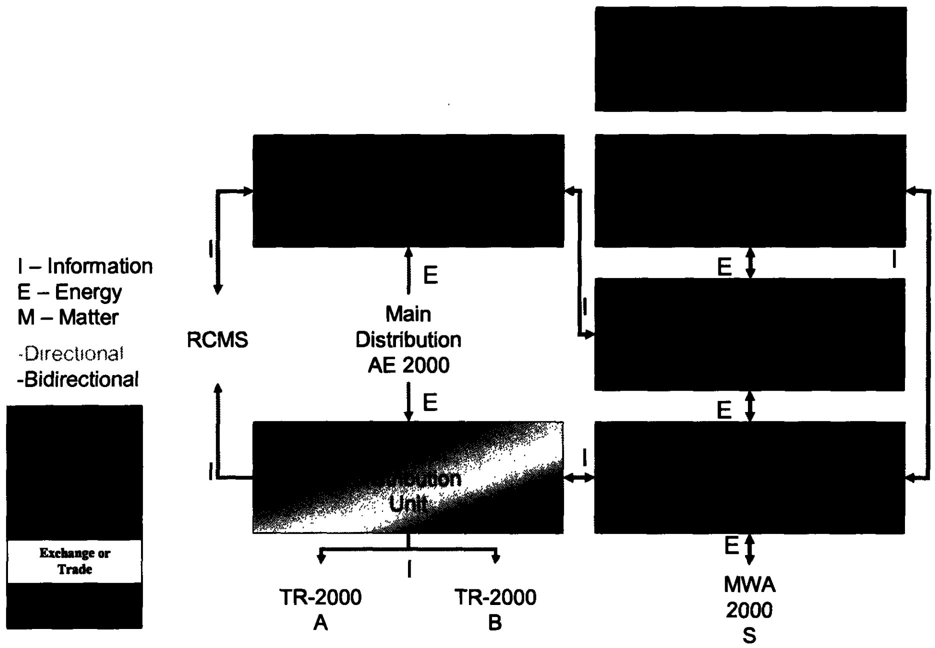


Figure 8 – Redrawing of Figure 7

In Figure 9, each element of the level 2 which are not mono-functional (i.e., represented with monochromatic boxes) are further decomposed into elements. These elements consequently belong to the Level 3 of decomposition and are represented, according to our convention, with smaller boxes.

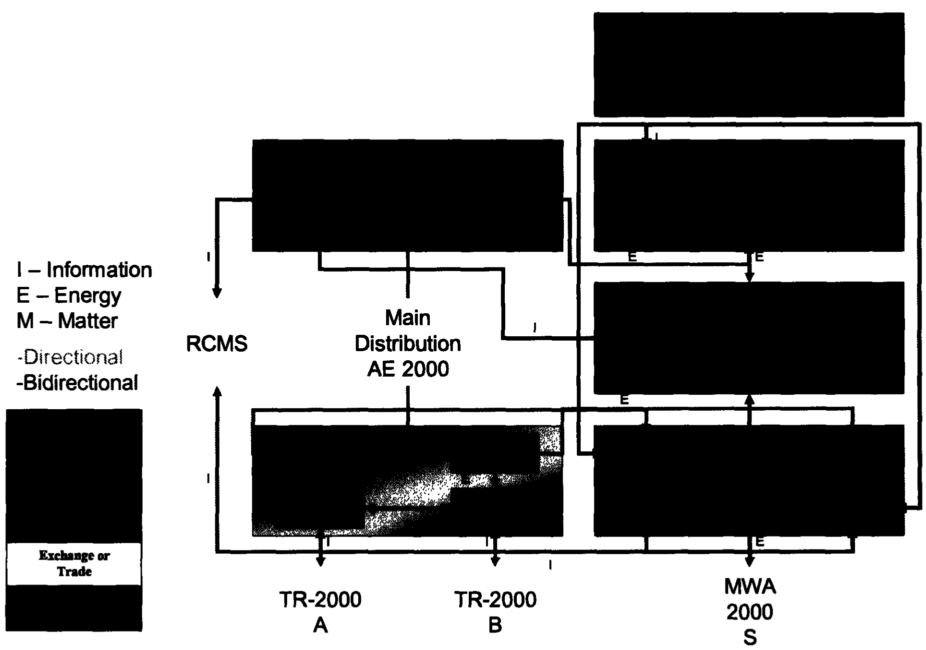


Figure 9 – Aerial System: level 3 of decomposition of STAR 2000

Finally, Figure 10 shows the Reference Decomposition of the Aerial System. It is obtained by keeping only the mono-functional elements of Figure 9. The Pedestal and the MSSR Antenna (opt) which was enlarged earlier for convenience purposes are also rescaled according to the convention.
The multifunctional element: "MSSR Antenna (optional)" is represented in the Reference Decomposition while it should not be, just to remind us that it has to be considered for decomposition if it were included into the system (which is not the case here as we consider only the basic configuration of STAR 2000).

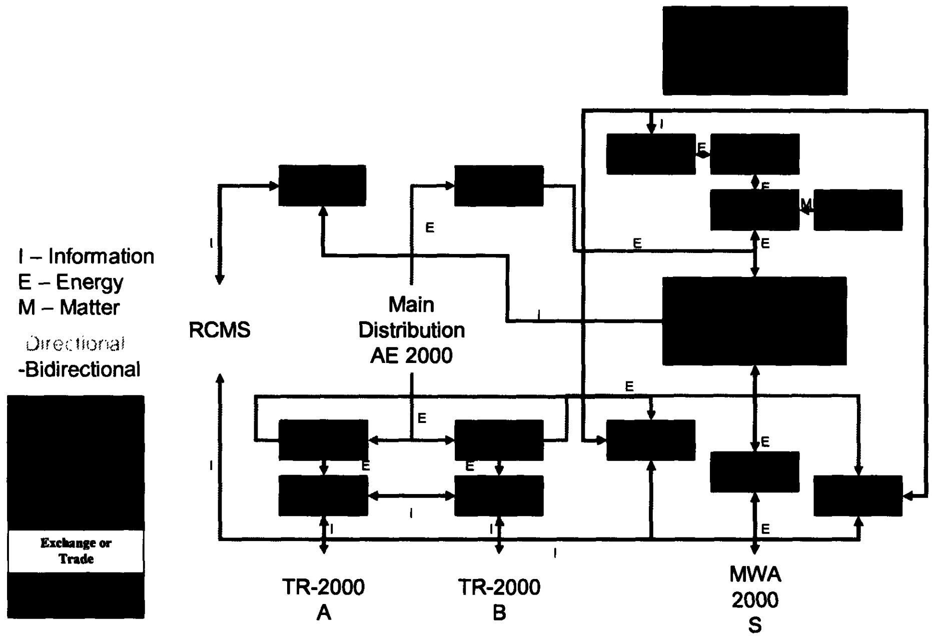


Figure 10 – Reference Decomposition of the Aerial System (with an optional element)

The figures detailing the methodology to obtain the Reference Decomposition of the other STAR 2000 subsystems are in Appendices. Appendix A is for the GRA, Appendix B for the MWA, Appendix C for the TR-2000 A, Appendix D for the TR-2000 B and Appendix E for the SST(8).

Finally, following the methodology in section B.II.2 (p. 16), the Reference Decomposition of STAR 2000 (Figure 11) is obtained by linking appropriately the Reference Decompositions of all its subsystems.

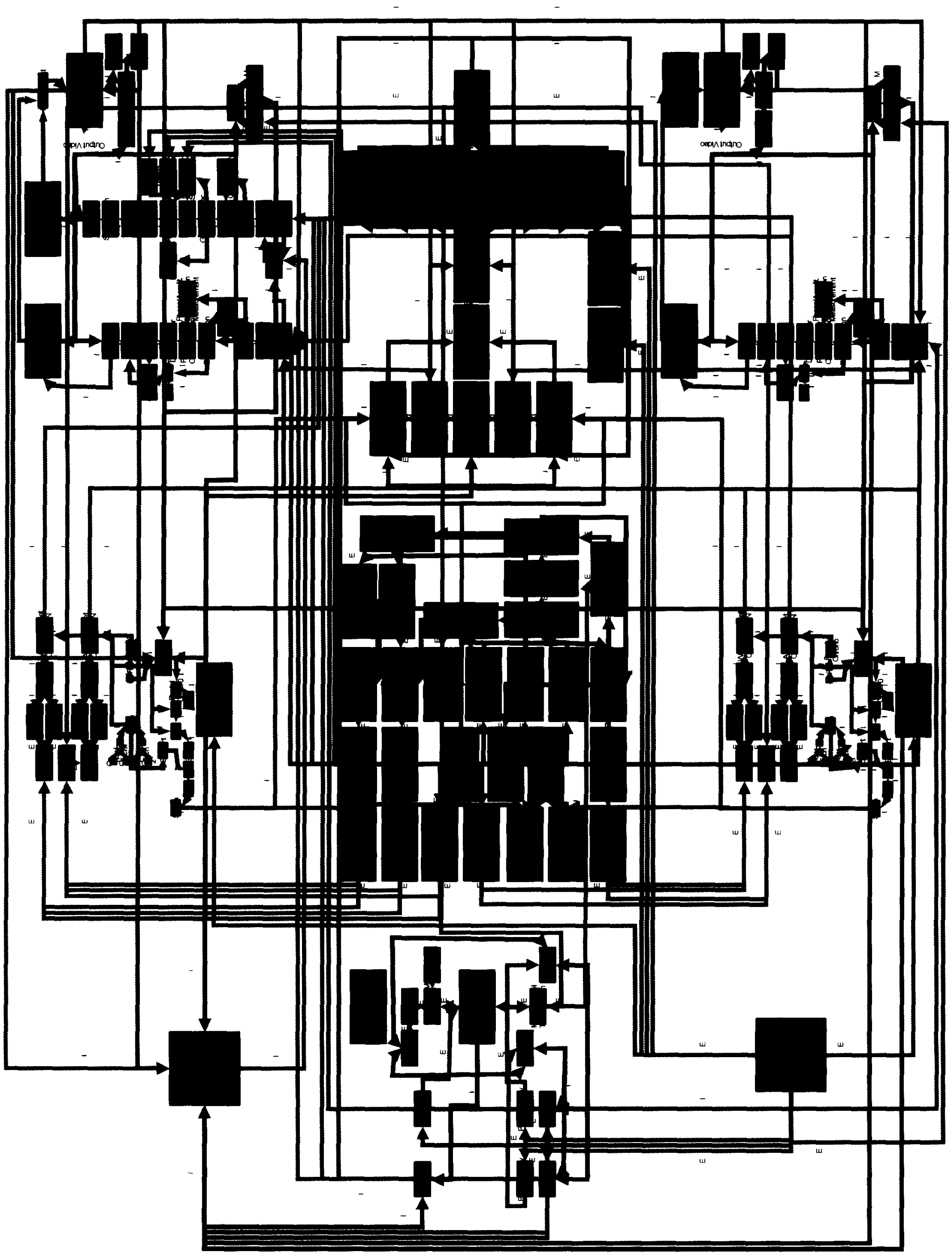


Figure 11 – Reference Decomposition of STAR 2000 (scale: 45%)

b. Computation of Internal Complexity

First, we illustrate the process to compute Internal Complexity on a STAR 2000 subsystem: the processor TR-2000 B.

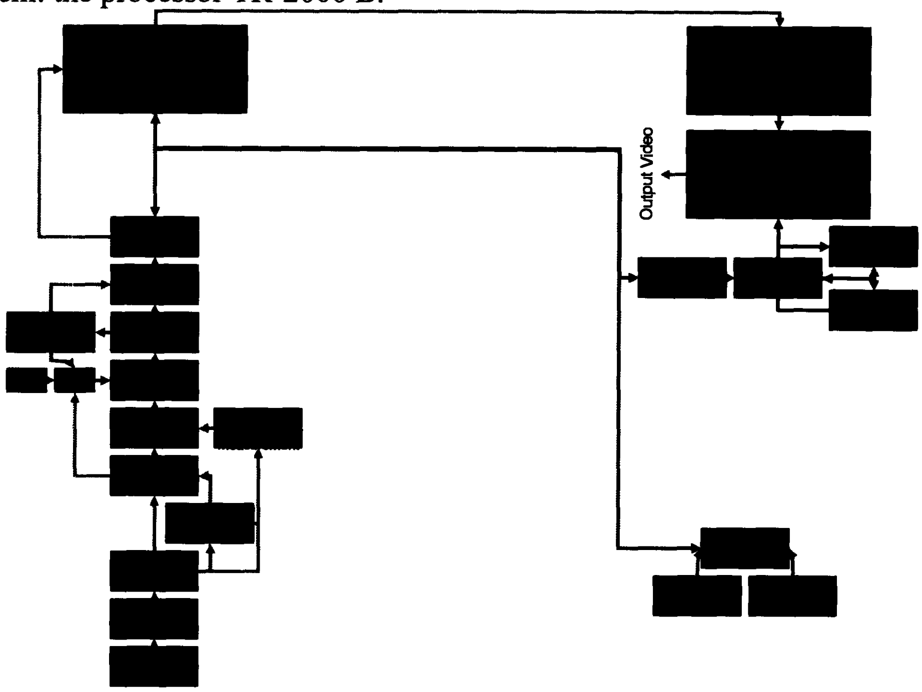


Figure 12 – Reference Decomposition of TR-2000 B

The computation of TR-2000 B Internal Complexity is based on its Reference Decomposition (Figure 12) to assess the number of elements (E_j), the vocabulary (V_j) and the hierarchy (H_j) for each level and the total number of links (L).

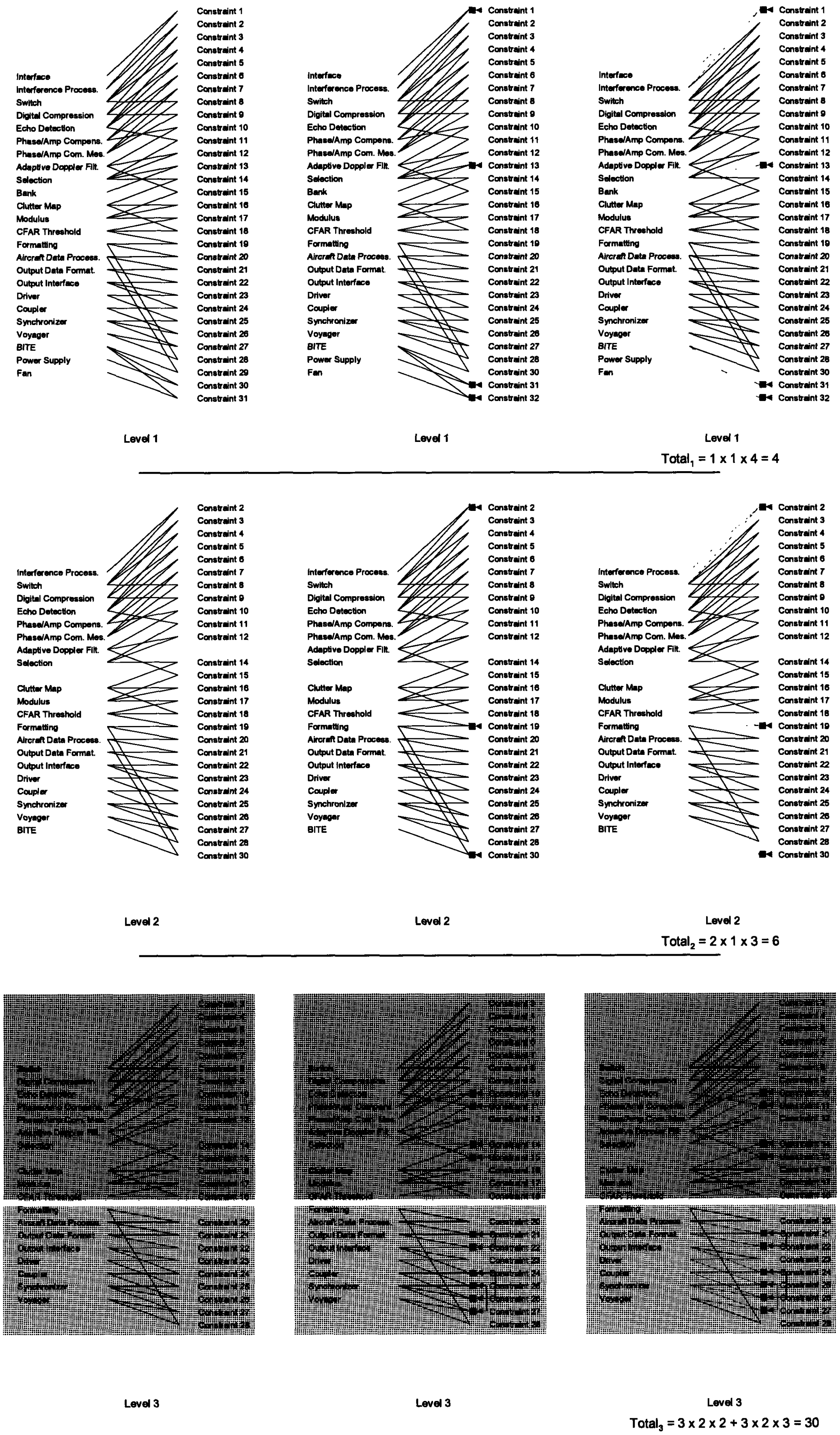
Level	E_j	V_j	H_j	L
2	3	2	2	35
3	19	4	3	
4	2	2	4	
Total	24	4	4	

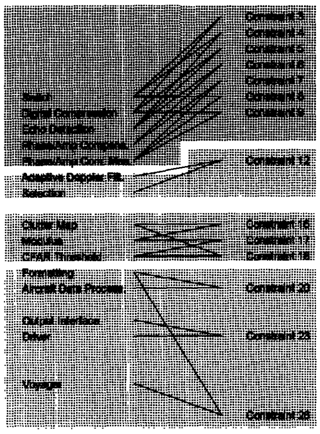
Table 4 – Quantitative description of TR-2000 B

Since, in this subsystem, there are no redundant or identical elements, the number of elements (E_j) equals the number of non-redundant elements (E_j^{NR}) and the number of non-identical elements (E_j^{NI}): $E_j=E_j^{NR}=E_j^{NI}$. This property reduces the number of possible values of Internal Complexity given by alternative metrics because Scale Complexity (c_s) is identical counting all the elements, counting only the non-redundant ones or counting only the non-identical ones.

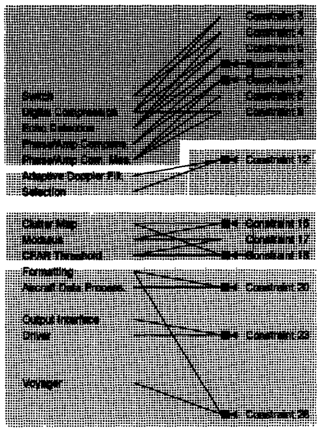
Some Internal Complexity metrics are based on Connectivity. We detail in Figure 13 the computation of Connectivity for TR-2000 B. The Connectivity of TR-2000 B is 103 (= 4+6+30+32+25+6).

Figure 13 – Computation of the Connectivity of TR-2000 B

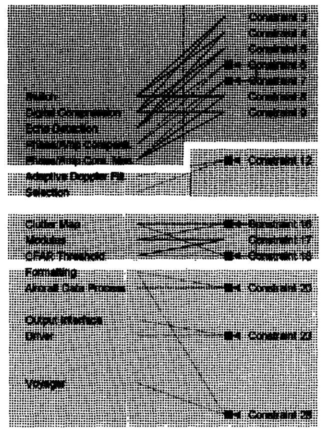




Level 4

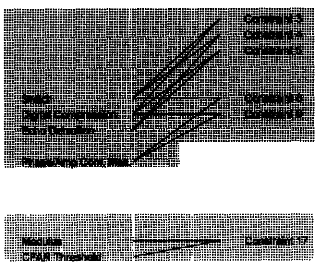


Level 4

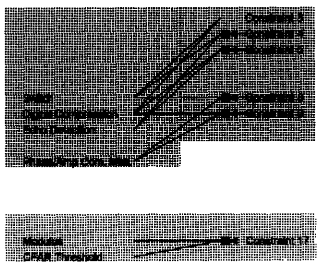


Level 4

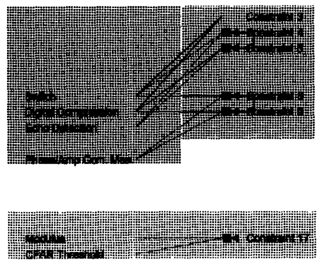
$Total_4 = 4 \times 2 \times 1 + 4 \times 1 \times 1 + 4 \times 2 \times 1 + 4 \times 1 \times 2 + 4 \times 1 \times 1 = 32$



Level 5

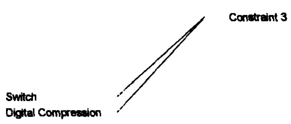


Level 5

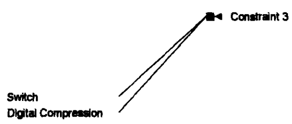


Level 5

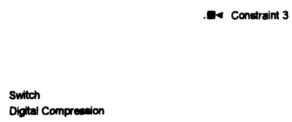
$Total_5 = 5 \times 2 \times 2 + 5 \times 1 \times 1 = 25$



Level 6



Level 6



Level 6

$Total_6 = 6 \times 1 \times 1 = 6$

Finally, Table 5 presents the different components: Scale Complexity (C_S) and Link Complexity (C_L) of TR-2000 B. Table 6 summarizes the value of the different Internal Complexity metrics (C) for this processor. For the components as well as for Complexity itself, the different available alternatives (Micro/Macro for C_S, (1)/(2)/(Cty)/(Cty') for C_L and in each case, norm/product for C) are computed.

Components					
Scale Complexity (C_S)		Link Complexity (C_L)			
MACRO	micro	1	2	Cty	Cty'
55.73	**49.72**	**0.06341**	**1.458**	**103**	**4.292**

Table 5 – Components of the Internal Complexity of TR-2000 B

	Complexity Metrics							
	$C=(C_S^2+C_L^2)^{1/2}$				$C=C_S \times C_L$			
C_S \ C_L	1	2	Cty	Cty'	1	2	Cty	Cty'
MACRO	**55.73**	**55.75**	**117.1**	**55.89**	**3.533**	**81.27**	**5740**	**239.2**
micro	**49.72**	**49.74**	**114.4**	**49.90**	**3.153**	**72.51**	**5121**	**213.4**

Table 6 – Internal Complexity metrics for TR-2000 B

Following the process just detailed on the example of TR-2000 B, Table 7 summarizes the different Scale Complexity (C_S) and Link Complexity (C_L) metrics possible for the different STAR 2000 subsystems.

	Components					
	Scale Complexity (C_S)		Link Complexity (C_L)			
	MACRO	micro	1	2	Cty	Cty'
Aerial System	**30.25**	**29.39**	**0.1593**	**2.071**	**74**	**10.57**
TR-2000 A	**92.92**	**85.40**	**0.03488**	**1.465**	**187**	**4.250**
GRA	**83.42**	**69.50**	**0.03565**	**1.176**	**139**	**4.088**
SST(8)	**32.26**	**32.26**	**0.1699**	**2.889**	**192**	**9.600**
SST(16)	**46.60**	**46.60**	**0.1292**	**3.231**	**352**	**12.57**
MWA	**44.38**	**44.38**	**0.04233**	**1.143**	**55**	**1.964**
TR-2000 B	**55.73**	**49.72**	**0.06341**	**1.458**	**103**	**4.292**

Table 7 – Complexity components of the main STAR 2000 subsystems

Finally, Table 8 shows the values of the 48 alternatives of Internal Complexity (C) for the main STAR 2000 subsystems. All the parameters necessary to do the computation are in Appendix G.

Complexity (C)	$C = (C_s^2 + C_L^2)^{1/2}$				$C = C_s \times C_L$				$C = (C_s^2 + C_L^2)^{1/2}$				$C = C_s \times C_L$			
Scale (C_s)	MACRO								micro							
Link (C_L)	1	2	Cty	Cty'	1	2	Cty	Cty'	1	2	Cty	Cty'	1	2	Cty	Cty'
Aerial System	30.25	30.32	79.95	32.05	4.821	62.67	2239	319.8	29.39	29.46	79.62	31.23	4.682	60.87	2174	310.6
TR-2000 A	92.92	92.93	208.8	93.02	3.241	136.1	17376	394.9	85.40	85.41	205.6	85.50	2.979	125.1	15970	362.9
GRA	83.42	83.43	162.1	83.52	2.974	98.14	11595	341.0	69.50	69.51	155.4	69.62	2.478	81.76	9660	284.1
SST(8)	32.27	32.39	194.7	33.66	5.483	93.21	6195	309.7	32.27	32.39	194.7	33.66	5.483	93.21	6195	309.7
SST(16)	46.60	46.72	355.1	48.27	6.023	150.6	16405	585.9	46.60	46.72	355.1	48.27	6.023	150.6	16405	585.9
MWA	44.38	44.39	70.67	44.42	1.878	50.72	2441	87.17	44.38	44.39	70.67	44.42	1.878	50.72	2441	87.17
TR-2000 B	55.73	55.75	117.1	55.89	3.533	81.27	5740	239.2	49.72	49.74	114.4	49.91	3.153	72.51	5121	213.4
Aerial System NR	23.77	23.86	77.72	26.02	3.788	49.24	1759	251.3	22.90	23.00	77.46	25.22	3.649	47.44	1695	242.1
TR-2000 A NR	92.92	92.93	208.8	93.02	3.241	136.1	17376	394.9	85.40	85.40	205.6	85.51	2.979	125.1	15970	363.0
GRA NR	73.60	73.61	157.3	73.72	2.624	86.59	10231	300.9	60.33	60.34	151.5	60.47	2.151	70.97	8386	246.6
SST(8) NR	14.34	14.63	192.5	17.26	2.437	41.43	2753	137.7	14.34	14.63	192.5	17.26	2.437	41.43	2753	137.7
SST(16) NR	14.34	14.70	352.3	19.07	1.853	46.33	5048	180.3	14.34	14.70	352.3	19.07	1.853	46.33	5048	180.3
MWA NR	33.28	33.30	64.29	33.34	1.409	38.04	1831	65.38	33.28	33.30	64.29	33.34	1.409	38.04	1831	65.38
TR-2000 B NR	55.73	55.75	117.1	55.89	3.533	81.27	5740	239.2	49.72	49.74	114.4	49.91	3.153	72.51	5121	213.4
Aerial System NI	23.77	23.86	77.72	26.02	3.788	49.24	1759	251.3	22.90	23.00	77.46	25.22	3.649	47.44	1695	242.1
TR-2000 A NI	64.83	64.85	197.9	64.97	2.261	94.98	12123	275.5	59.40	59.42	196.2	59.55	2.072	87.03	11108	252.4
GRA NI	63.79	63.80	152.9	63.92	2.274	75.05	8867	260.8	53.16	53.17	148.8	53.31	1.895	62.54	7389	217.3
SST(8) NI	14.34	14.63	192.5	17.26	2.437	41.43	2753	137.7	14.34	14.63	192.5	17.26	2.437	41.43	2753	137.7
SST(16) NI	14.34	14.70	352.3	19.07	1.853	46.33	5048	180.3	14.34	14.70	352.3	19.07	1.853	46.33	5048	180.3
MWA NI	23.77	23.80	59.92	23.86	1.006	27.17	1308	46.70	23.77	23.80	59.92	23.86	1.006	27.17	1308	46.70
TR-2000 B NI	55.73	55.75	117.1	55.89	3.533	81.27	5740	239.2	49.72	49.74	114.4	49.91	3.153	72.51	5121	213.4

Table 8 – 48 alternative Internal Complexity metrics for the STAR 2000 subsystems

2. Choice of the relevant Internal Complexity metrics

From all the 48 possible Complexity metrics we have identified and applied to STAR 2000, we will now pick the most relevant one. Following a top-down process, we will eliminate the metrics which are not in accordance with our conception of Internal Complexity. We will also assess the metrics taking into account the views of an expert of STAR 2000 [17]. The subsystems listed by increasing order of complexity are:

- MWA (Microwave Assembly)
- Aerial System
- GRA (Generation Reception Assembly)
- SST(8) (Solid-State Transmitter)
- SST(16) (Solid-State Transmitter)
- TR-2000 B (Aircraft Processor)
- TR-2000 A (Aircraft and Weather Processor)

C(MWA) is low, C(SST(16)) is comparable to C(SST(8)) but slightly higher and C(TR-2000 A) is a bit higher than C(TR-2000 B).

Even though these considerations are based upon deep practical knowledge based upon experience, they may not be blindly followed in the discussion to assess the metrics.

The mathematical properties of the Internal Complexity metrics defined as Norm ($\mathsf{C}=(\mathsf{C}_s^2+\mathsf{C}_L^2)^{1/2}$) make it inapplicable when the two components are quite different. In our example, Link Complexity (1), (2) and (Cty') are very different from Scale Complexity (one or two order of magnitude). The resulting Complexity is thus arbitrarily dominated by the higher component. We eliminate these metrics because they do not properly balance the two components of Complexity we believe essential to internal complexity - scale and link. Table 9 summarizes the remaining possible metrics where the black boxes are eliminated metrics.

C		Norm						Product					
C_s		MACRO			micro			MACRO			micro		
		All	NR	NI	All	NR	NI	All	NR	NI	All	NR	NI
C_L	(1)												
	(2)												
	(Cty)												
	(Cty')												

Table 9 – 30 remaining metrics

The Complexity metrics defined as a Product ($\mathsf{C} = \mathsf{C}_s \times \mathsf{C}_L$) using Link Complexity (1) is also not valid (Table 10). When we apply it to the different subsystems and to the overall system, the complexity of STAR 2000 is lower than the complexity of some of its subsystems (e.g., when all the elements are counted, C(STAR)=4.669 (for the MACRO and 3.900 for the micro) while C(SST(8))=5.483 (for both the MACRO and the micro). This contradiction arises because STAR is less dense than some of its subsystems. The product relationship equally weights the two components of complexity: the Internal Complexity emerging from the scale of the overall system does not compensate for the reduced coupling due to the non-normalized density.

C		Norm						Product					
C_s		MACRO			micro			MACRO			Micro		
		All	NR	NI	All	NR	NI	All	NR	NI	All	NR	NI
C_L	(1)												
	(2)												
	(Cty)												
	(Cty')												

Table 10 – 24 remaining metrics

In order to further examine the 24 remaining metrics, we now study each remaining possible alternative metric individually. In Figure 14 and Figure 15, we represent for each alternative metric the relative Internal Complexity of the different subsystems. From now, the nomenclature for the metrics is:

- MACRO / micro for: C_s Macro or micro
- (1) / (2) / (Cty) / (Cty') for: C_L (1), (2), (Cty) or (Cty')
- (+) / (x) for: C Norm or Product

- Ø / NR / NI — for: C_s counting all the elements, only the non-redundant or only the non-identical ones

(e.g., MACRO1xNR: stands for the Product metrics based on the Link Complexity (1) and on the Macro Scale Complexity counting only the non-redundant elements, micro2+NI: stands for the Norm metrics based on the Link Complexity (2) and on the micro Scale Complexity counting only the non-identical elements...)

In the figures presenting the different alternatives some metrics that have previously been eliminated are shown to provide a better basis for comparison and understanding. The data are presented in three sets (Ø / NR / NI) of two graphs (MACRO and micro) and are analyzed following the order of appearance. Below is the list of all the metrics and for each one eliminated the reason why it is not retained as a valuable metrics is shown. Conversely, the metrics that can be potentially used as Internal Complexity metrics are underlined in this list (restrictions appear in comments).

For Ø:

- micro1x: C(SST(16)) >> C(TR-2000 A) not retained
- micro2x: C(SST(16)) ≈ C(TR-2000 A) not retained
- microCtyx: C(TR-2000 A) >> C(TR-2000 B) not retained
- microCty+: C(SST(16)) >> C(TR-2000 A) not retained
- MACRO1x: C(SST(16)) >> C(TR-2000 A) not retained
- MACRO2x: C(SST(16)) ≈ C(TR-2000 A) and C(SST(16)) >> C(SST(8)) not retained
- MACROCtyx: C(TR-2000 A) >> C(TR-2000 B) not retained
- MACROCty+: C(SST(16)) so high not retained

For NR:

- micro1xNR: C(TR-2000 B) > C(TR-2000 A) not retained
- <u>micro2xNR</u>: C(TR-2000) >> C(SST) seems to be a bit exaggerated
- microCtyxNR: C(TR-2000 A) and C(GRA) so high not retained
- microCty+NR: C(SST(16)) so high not retained
- MACRO1xNR: C(SST(8)) > C(SST(16)) not retained
- <u>MACRO2xNR</u>: C(TR-2000) >> C(SST) seems to be a bit exaggerated
- MACROCtyxNR: C(TR-2000 A) >> C(TR-2000 B) not retained
- MACROCty+NR: C(SST(16)) >> C(TR-2000 A) and C(TR-2000 B) >> C(GRA) not retained

For NI:

- micro1xNI: C(TR-2000 B) > C(TR-2000 A) not retained
- <u>micro2xNI</u>: C(GRA) > C(SST(8)) a little questionable
- microCtyxNI: C(TR-2000 A) and C(GRA) so high not retained
- microCty+NI: C(Aerial System) so low not retained
- MACRO1xNI: C(SST(8)) > C(SST(16)) and C(TR-2000 B) > C(TR-2000 A) not retained

- MACRO2xNI: C(GRA) seems too high
- MACROCtyxNI: C(TR-2000 A) and C(GRA) so high not retained
- MACROCty+NI: C(Aerial System) so low not retained

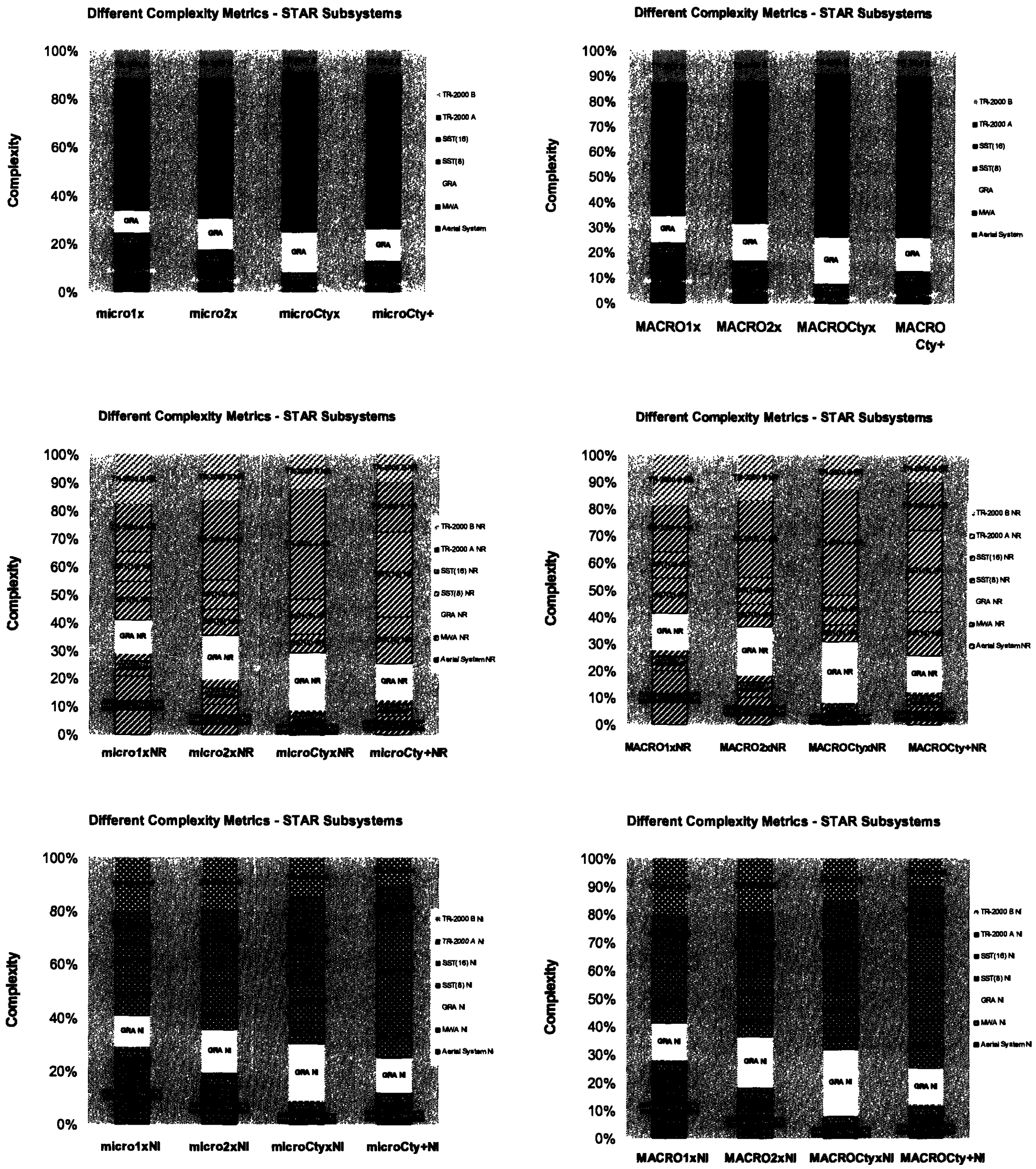


Figure 14 – Complexity of the STAR 2000 subsystems for different metrics (A)

The only parameter that has not been yet taken into account is the Link Complexity defined as normalized Connectivity (Cty'). To evaluate it, we now present the data under the same format as earlier with Cty' instead of Cty. The analysis will only focus on the metrics featuring Cty'. Internal Complexity metrics defined as norm and featuring Link

Complexity (1) or (2), having already been studied earlier, are presented only to provide a better basis for comparison.

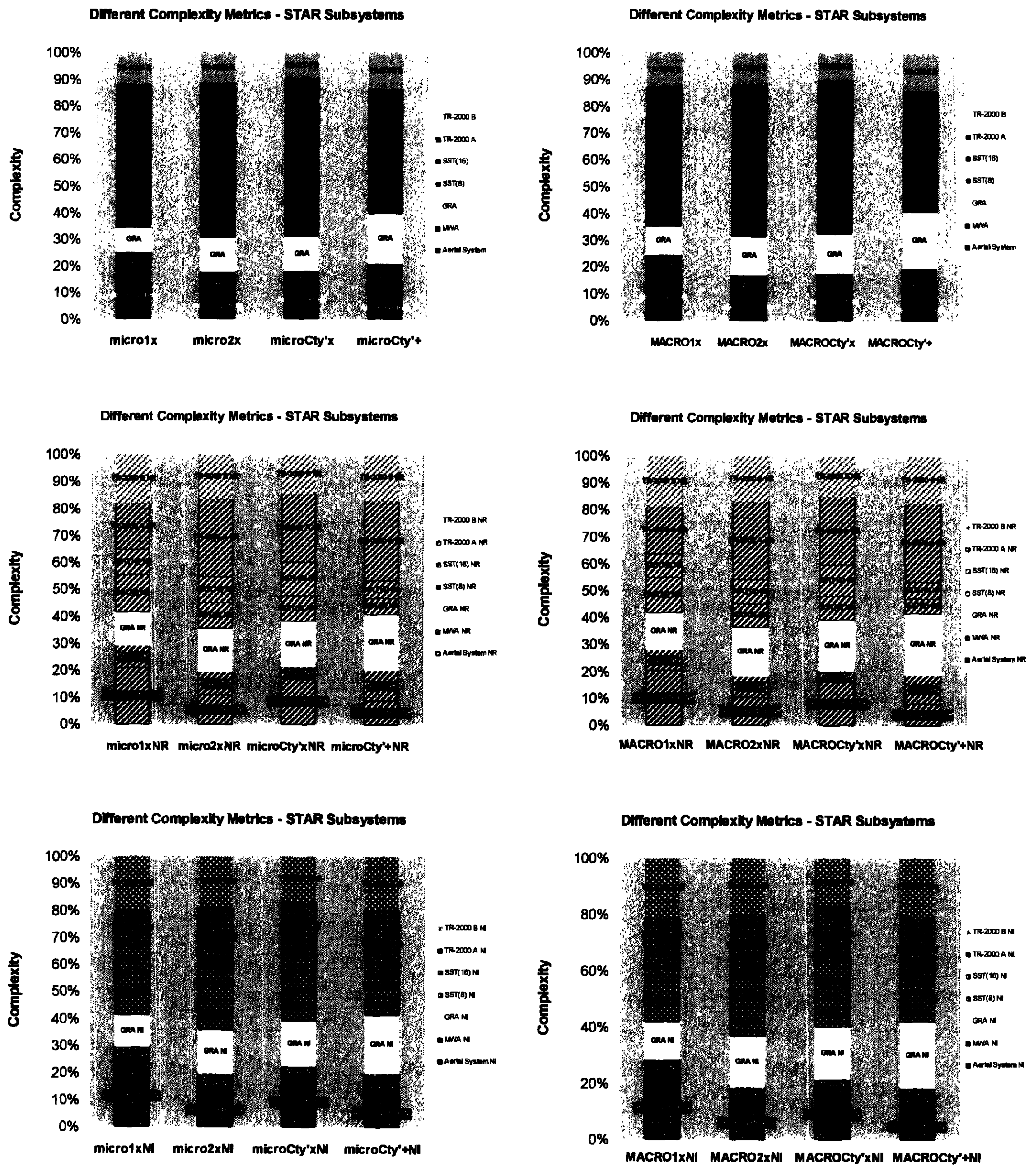


Figure 15 – Complexity of the STAR 2000 subsystems for different metrics (B)

For Ø:

- microCty'x: C(SST(16)) so high not retained
- microCty'+: C(GRA) >> C(Aerial System) not retained
- MACROCty'x: C(SST(16)) >> C(SST(8)) not retained
- MACROCty'+: C(GRA) >> C(Aerial System) and

C(TR-2000 A) >> C(TR-2000 B) not retained

For NR:

- microCty'xNR: C(TR-2000 A) >> C(TR-2000 B) not retained
- microCty'+NR: C(SST) so low not retained
- MACROCty'xNR:C(TR-2000 A) >> C(TR-2000 B) a little questionable
- MACROCty'+NR:C(GRA) >> C(Aerial System) and C(TR-2000 A) >> C(SST(8)) not retained

For NI:

- microCty'xNI: C(GRA) > C(TR-2000 B) a little questionable
- microCty'+NI: C(SST(8)) so low not retained
- MACROCty'xNI: C(SST(8)) maybe a little too low
- MACROCty'+NI: C(GRA) so high not retained

Table 11 maps the remaining possible metrics.

C		Norm						Product					
C_S		MACRO			micro			MACRO			micro		
		All	NR	NI	All	NR	NI	All	NR	NI	All	NR	NI
C_L	(1)												
	(2)												
	(Cty)												
	(Cty')												

Table 11 – 7 remaining metrics

Regarding Link Complexity (2), micro metrics appear superior to the Macro ones because they give higher importance to the Aerial System; micro2xNI is also better than micro2xNR because it gives more importance to the SST and the Aerial System in accordance with the expert rating.

Regarding, Link Complexity (Cty'), MACROCty'xNI appears superior to MACROCty'xNR because it better agrees with the expert ranking of the SST and the Aerial System; microCty'xNI is even better than MACROCty'xNI because it emphasizes again the former argument and also decreases the importance of the GRA in agreement with the expert rating.

C		Norm						Product					
C_S		MACRO			micro			MACRO			micro		
		All	NR	NI	All	NR	NI	All	NR	NI	All	NR	NI
C_L	(1)												
	(2)												
	(Cty)												
	(Cty')												

Table 12 – The 2 possible metrics

Table 12 identifies the two remaining possible metrics for valuable Internal Complexity metrics: microCty'xNI and micro2xNI. Even if they are both useful, they do not give exactly the same result: the ranking of two subsystems is inverted and the respective importance of the subsystems in complexity is not exactly the same (even if it is very close). Finally, we prefer the second one, not for the result it gives, but only because it is easier to compute. Indeed, while Connectivity is an algorithm which needs to be computed manually, Density is just a simple calculation.
The Internal Complexity metrics we recommend is microCty'xNI:

$$C_{INT}(E,L,V,H) = \frac{L}{E} \times \sum_j E_j^{NI} \frac{\mathrm{Ln}\,((V_j+1) \times (H_j+1))}{\mathrm{Ln}\,(4)} \quad \text{(j: level)}$$

We now note some of its properties. As the product of two intensive components, C_{INT} is intensive.

$C_{INT}(0, L, V, H) = 0$

$\aleph \rightarrow \Re$

$L \rightarrow C_{INT}(E, L, V, H)$ is an increasing function: $C_{INT}(E, L+1, V, H) > C_{INT}(E, L, V, H)$

Moreover $C_{INT}(E, 0, V, H) = 0$

$\aleph \rightarrow \Re$

$V \rightarrow C_{INT}(E, L, V, H)$ is an increasing function: $C_{INT}(E, L, V+1, H) > C_{INT}(E, L, V, H)$

Moreover $C_{INT}(E, L, 0, H) = 0$ (also because $V=0 \Rightarrow E=L=0$)

$\aleph \rightarrow \Re$

$H \rightarrow C_{INT}(E, L, V, H)$ is an increasing function: $C_{INT}(E, L, V, H+1) > C_{INT}(E, L, V, H)$

Moreover $C_{INT}(E, L, V, 0) = 0$ (also because $H=0 \Rightarrow E=L=0$)

Further properties of this metric are specific to the system studied and it is difficult to say more about its general behavior because it depends upon the relative evolution of the elements (number, vocabulary and hierarchy) and the links.

3. Computation of STAR 2000 Internal Complexity

Since we have now identified the best Complexity metric, we can compute the Internal Complexity of STAR 2000. Table 13 summarizes all the parameters necessary to apply the metrics to this complex system.

Level	E_j^{NI}	V_j	H_j
1	2	2	1
2	28	3	2
3	41	4	3
4	15	3	4
5	6	3	5

E
197

L
373

Table 13 – Parameters to compute STAR 200 Internal Complexity

So, $\mathbf{C}_{\mathrm{INT}}(\mathrm{STAR\ 2000}) = \frac{L}{E} \times \sum_j E_j^{NI} \frac{\mathrm{Ln}((V_j+1) \times (H_j+1))}{\mathrm{Ln}(4)} = 350.20$

We note that the sum of the Internal Complexities of each STAR 2000 subsystem is appropriately lower than the complexity of the latter. (The computation of the sum of the complexities of the different subsystems in order to be compared to the complexity of the overall system is done counting only the non-identical elements (i.e., only counting one GRA (and not two) and not counting the TR-2000 B because it is fully equivalent to a part of TR-2000 A). The discrepancy between these two computations of Internal Complexity reflects the increase in Link Complexity due to the connection between the different subsystems.

4. Application to the maritime radar

In comparison to STAR 2000 (whose description is given p.28), SCOUT is an I-Band, short-to-medium range, surface surveillance and navigation radar. It is a fully solid-state system with high reliability, low weight and small dimensions. As a frequency modulated continuous wave radar it is used in littoral, coastal and harbor surveillance applications.

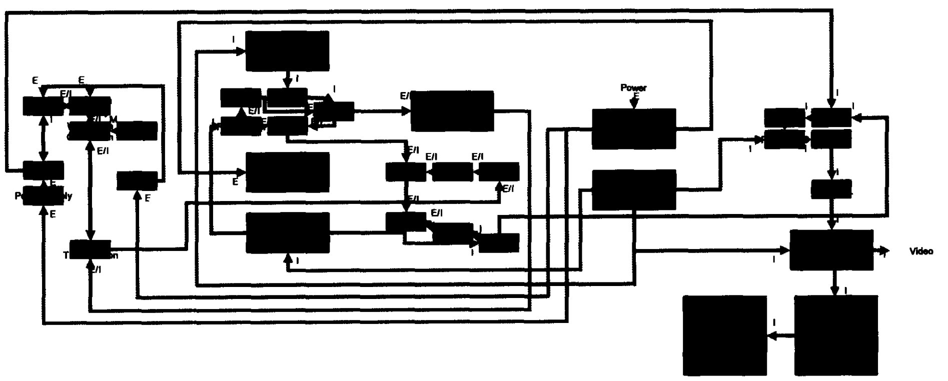


Figure 16 – Reference Decomposition of SCOUT (scale: 45%)

From the Reference Decomposition in Figure 16 we can infer the data presented in Table 14 in order to compute the Internal Complexity of SCOUT.

Level	E_j^{NI}	V_j	H_j
1	2	2	1
2	7	4	2
3	23	4	3

E
32

L
45

Table 14 – Parameters to compute SCOUT Internal Complexity

So, $C_{INT}(\text{SCOUT}) = \frac{L}{E} \times \sum_j E_j^{NI} \frac{\text{Ln}((V_j + 1) \text{ x } (H_j + 1))}{\text{Ln}(4)} = 89.123$

Page intentionally left blank

C. Interface Complexity

Between a device and its environment lies, by definition, the interface. To understand the influence of the external complexity on the internal complexity, the interface is an obvious aspect to study.
Before focusing on interface complexity more precisely we want to identify its role in the general complexity framework detailed in Figure 17. Requirements attempt to capture what is asked of the system by its environment; therefore, interface complexity can be studied in terms of requirements. Interface complexity can also be studied in terms of performances. Indeed, performances echo requirements because they express how the system acts on its environment. Requirements are imposed by the environment to the system: they stem from the outside of the system and go inside the system; performances, the counterpart of requirements, are the answer of the system to its environment: they stem from the inside of the system and go outside the system.

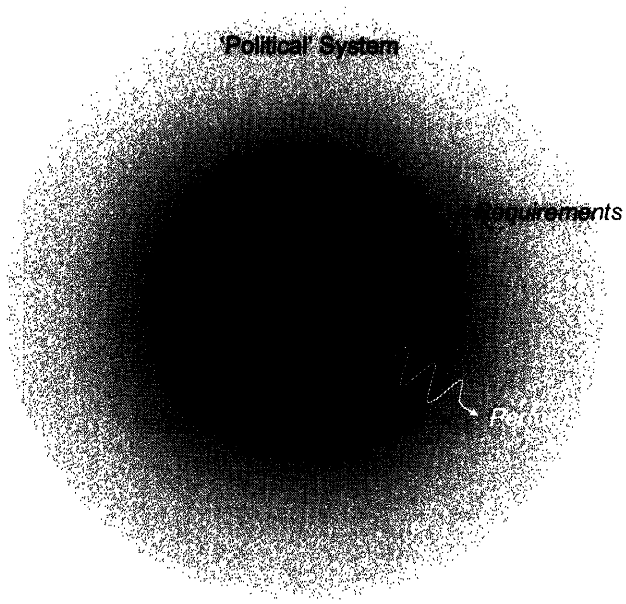


Figure 17 – Interface Complexity in the framework to study complexity

In this thesis, we first study interface complexity qualitatively in terms of performances. Then, we study it quantitatively in terms of requirements.

I. Qualitative approach to Interface Complexity

One of our aims is to explore the relationship between internal and external complexity. To assess the complexity at the interface between the system and its environment we compare the respective performances of the two test bed systems. A qualitative framework is developed to inform the complexity of the interface and to help understanding how external and internal complexities interact.

1. Method

The framework consists in making a comparative table of the respective performances of both systems. In Table 15, the performance category is named in the first column and the value of each performance for both systems is given in the two others. For each performance category, a red cell signifies the system whose value is more stringent (i.e., hard to achieve) and a green cell, the system whose value is less stringent. Light red and light green are used to compare performances which do not seem to weigh a lot in complexity. Yellow is used for equal performances. A comparison of the dominant color of the two columns will inform the relative complexity at the interface of the two systems in their respective environment. Thanks to such table, we can assess comparatively the Interface Complexity of the two systems in their environment.

2. Application to the ATC and maritime radar

In fact the comparative table (Table 15) presents two sub-columns for each test bed system to increase the scope of comparison comparing two typical configurations of each system.

3. Conclusion

Simply looking at the two main columns for the ATC and the maritime radar, one can easily tell that red is the dominant color for STAR 2000 and green the dominant one for SCOUT.
To conclude, Interface Complexity assessed in this qualitative performance comparison is higher for STAR 2000 than for SCOUT:

$$C_{INTERFACE}\text{ (STAR 2000)} > C_{INTERFACE}\text{ (SCOUT)}$$

4. Drawbacks of this approach

We acknowledge that this way of proceeding works because we have chosen two comparable systems: two radars. We also acknowledge, by way of consequence, that this method cannot be generalized to systems whose performances are not comparable. Moreover, we know that this method is not quantitative and we knew in advance that it would help support our hypothesis because we expected much more “red” in one column than in the other. In case of outcomes less distinguishable, it would be hard to draw a conclusion because we do not know the respective weight of each typical performance in complexity. Finally, while useful, this approach to interface complexity is not the only one. We now present a quantitative approach to interface complexity.

[redacted]					
Minimum range		[redacted]			
Instrumented range		[redacted]			
Transmitter	Modules	[redacted]			
Antenna	Gain	[redacted]			
	Polarization	[redacted]			
	Rotation rate	[redacted]			
	Beamwidth (Hor.)	1.4°	1.4°	1.4°	1.4°
	Beamwidth (Vert.)	[redacted]			
	Side lobe level	[redacted]			
Range cell		[redacted]			
Accuracy (range/azimuth)		[redacted]			
Resolution (range/azimuth) (Pd = 80%)		[redacted]			
Improvement Factor	ground clutter	[redacted]			
	rain clutter	[redacted]			
Sub-clutter visibility	ground clutter	[redacted]			
	rain clutter	[redacted]			
Pulse compression side lobes		-50 dBc	-50 dBc	-50 dBc	-50 dBc
Target	Velocity range	[redacted]			
	Manœuvre	[redacted]			
	Handling	[redacted]			
	Output delay	1.2 sec	1.2 sec	+/- 2sec	+/- 2ses
MTBCF		[redacted]			
MTTR		[redacted]			
Pd(Pfa=10^{-6}, typical range and section)		[redacted]			
Environment	Temperature	[redacted]			
	Humidity	93(or 80)% at +40°C	93(or 80)% at +40°C	95% at +30°C	95% at +30°C
	Wind (operation)	[redacted]			
	Wind (survival)	[redacted]			
	Solar Radiation	1 kW/m^2 at +45°C	1 kW/m^2 at +45°C	1120 W/m^2 at +40°C	1120 W/m^2 at +40°C
	Salt Atmosphere	Yes	Yes	Yes	Yes
	Vibrations	[redacted]			
	Shock (Processor)	[redacted]			

Table 15 – Comparative table to assess Interface Complexity

*: 16 modules version

II. Quantitative approach to Interface Complexity

A potentially valuable way to look at the interface complexity is to see it as the tension that reigns on the "membrane" that separates the system from its environment. This tension mainly depends upon two variables which are the resistance of the system and the pressure of the environment. When both the pressure and the resistance are high, interface complexity is high. Interface complexity is seen as the result of what is asked of the system and what the system can offer.

1. Method and theory

According to Nam Pyo Suh [18], "complexity is defined as a measure of the uncertainty in achieving the specified Functional Requirements. Therefore, complexity is related to information content, which is defined as a logarithmic function of the probability of achieving the Functional Requirements. The greater the information required to achieve the Functional Requirements of a design; the greater is the information content, and thus the complexity". Thus,

$$I = -\log_2(P)$$

The information content that the interface imposes on the system: $I_{sys} = -\log_2(P_{sys})$ is a function of the joint probability to satisfy all the Functional Requirements (P_{sys}). Therefore, the complexity defined as systemic information content depends upon P_{sys}: when P_{sys} is low, the information content is high and the complexity is high too.

There are two reasons why P_{sys} can be low. Firstly, the level of the Functional Requirements which is aimed to be reached is high. Secondly, this level is hard to reach. It is all the more improbable that one reaches one's aim if the aim is remote and if the track to go there is steep. P_{sys} depends upon the "where" to go and the "how" to go there.

The quantitative approach to Interface Complexity will take these two points into account to justify the nature of the complexity emerging from the system and its environment. Interface Complexity reflects the stringency of what is asked of the system by its environment and how the system handles it.

From a system standpoint, Suh defines complexity as a logarithmic function of the probability of achieving the Functional Requirements. Now, since we are interested in interface complexity as the result of what is asked of the system and what it can offer, we need to look at the problem from the other side and formulate Interface Complexity as a function of the probability of not achieving what is required. All the Functional Requirements are driven by the idea that the system must achieve its main functions without failure. Ideally, every complex system should achieve what it is asked to perform (i.e., its primary function: what it is "basically" designed for) without failing. Therefore, the information content of the probability of failure for the system as it is used (P) can be seen as what drives the Internal Complexity of the system in its environment. The lower the probability of failure of a system in its configuration as used, the higher the Interface Complexity.

Therefore, Interface Complexity is defined as:

$$C_{INTERFACE} = -\log_2(P)$$

P is the probability of failure of the system *as it is used*. The "as it is used" is essential in the concept of Interface Complexity since it is what takes into account the resulting trade-off between the external requirements and the internal capability of the system. The

Interface Complexity metric proposed measures the degree of achievement of this trade-off.

When P = 1, $C_{INTERFACE}$ = 0: the Interface Complexity of a system is null when it cannot achieve the function it has been designed for.
When P = 0.5, $C_{INTERFACE}$ = 1: the Interface Complexity of a system equals one when it equally fails or succeeds in achieving its usual function.
When $P \to 0$, $C_{INTERFACE} \to \infty$: the Interface Complexity of a system tends to infinity when the system fully achieves what it has been designed for.

To conclude, Interface Complexity defined as the information content of the probability of failure of the system under its normal conditions of use measures how hard it is to achieve what the system achieves. Indeed, what the system achieves stems form the trade-off between what can be done easily and what should be done ideally. Therefore the degree of achievement of the primary function takes into account both the characteristics of the system and the ones of its environment. According to this quantitative approach, Interface Complexity lies in the degree to which the system achieves what it must achieve.

2. Application to the ATC and maritime radar

Most of the performances of the two radars can be summarized into a "probability of correct detection over time". For their typical range, their typical probability of false alert (Pfa) and typical cross-section to be detected, both radars can be characterized by their:

- Probability of detection: Pd
- Availability: Av

Therefore, P' = Pd x Av is literally the probability that a target is correctly detected at anytime, i.e., the "probability of correct detection over time".
Conversely, P = 1-P' = 1-(Pd x Av) is the probability of failure of the radar as it is used.

Thus, for a radar Interface Complexity becomes:

$$C_{INTERFACE} = -\log_2(P) = -\log_2(1-(Pd \times Av))$$

Here, Pd and Av are the values as the radar is used.

For, STAR 2000, the usual range for the 8 modules version is 83.96 nmi and 97.79 nmi for the 16 modules version. For a typical Pfa = 10^{-6} on a typical cross-section = $2m^2$, Pd and Av are:
Pd = 80%
Av = 0.99999

So, the Interface Complexity of STAR 2000 in its environment is:

$C_{INTERFACE}$= - $\log_2$ (1-(Pd x Av)) = 3.321
This value of Interface Complexity is both for the usual range of 83.96 nmi (for the 8 modules version) and 97.79 nmi (for 16 modules) and for typical parameters.

$C_{INTERFACE}$ is the same when STAR 2000 is built with 8 or 16 modules. C_{INT} is also nearly the same for these two versions: 350.20 for 8 modules and 350.97 for 16 modules.
$C_{INTERFACE}$ is the same for the two configurations of STAR 2000 (range 60 nmi or 80 nmi). C_{INT} is also the same in these two cases.

For SCOUT, the usual range is 5.2 nmi. For the typical Pfa = 10^{-6} on a typical cross-section = $1m^2$, Pd and Av are:
Pd = 50%
Av = 0.99995

So, the Interface Complexity of SCOUT in its environment is:
$C_{INTERFACE}$= - $\log_2$ (1-(Pd x Av)) = 0.9999

3. Conclusion

$C_{INTERFACE}$ (STAR 2000) = 3.321
$C_{INTERFACE}$ (SCOUT) = 0.9999

So, $C_{INTERFACE}$ (STAR 2000) > $C_{INTERFACE}$ (SCOUT)

D. External Complexity

To fully explore the relationship between internal and external complexity, we also need to assess external complexity as we have internal complexity. Figure 18 which applies the framework to study complexity to the two test bed large-scale systems, reminds that external complexity is the complexity of the transport system. The transport system is differentiated from the "Political" system but they both constitute the environment of the radar. We propose metrics for external complexity that are to be computed in the transportation system (in order to be fully quantitative) and that also take into account the overall political system (in order for the metrics to be fully representative of the environment).

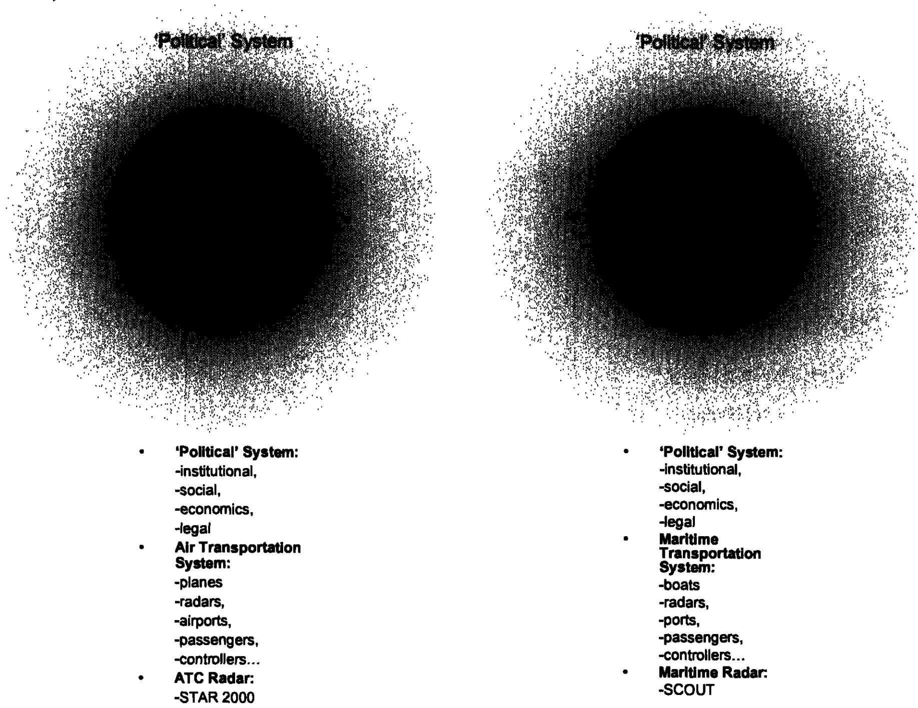


Figure 18 – Representation of the two radars in their respective environment

Focusing on the complexity of large-scale systems, we first implement a way to characterize their complexity. Then, from this characterization, we define two possible External Complexity metrics. Finally, we apply the characterization and the metrics to the two test bed large-scale systems (plus another one in order to have a further basis for comparison).

I. Characterization of complex systems

Failure is a characteristic common to all systems. No system is perfect: every system presents a certain level of risk and a certain configuration of risk. We believe that the configuration of risk, as a characteristic of systems, is a relevant indicator of the complexity of large-scale systems.

In order for a system to be a useful complex system, it must be reliable. Complex systems only exist because they fulfill human needs and because these needs tend to be more numerous and more stringent. Since there is no discontinuity in the fulfillment of human needs, complex systems cannot be discontinuous: they must be reliable. A complex system cannot be punctuated with many mistakes, as small as they could be. If it were so, nobody would use the system anymore. Reliability is a necessary condition to the existence of complex large-scale systems: the frequency of all failures (and specifically small failures) of complex systems is low.

Conversely, J.M. Carlson and J. Doyle state that complex systems are "robust, yet fragile" [19]. Extending their ideas, we assert that complex systems are reliable but also catastrophic. Besides reliability, the "tendency to catastrophe" seems to be a second key characteristic of complex systems. Complex large-scale systems present relatively high degree of tendency to catastrophe: they entail non negligible occurrence of high magnitude failures. On the other hand, "simple" systems are not complex enough to achieve functions that may lead to high magnitude failures: "simple" systems do not have high magnitude failures.

From these characteristics of complex large-scale systems: higher reliability and higher tendency to catastrophe, one can draw a conclusion on their risk configuration: complex large-scale systems are more extreme than "simple" systems – they are more reliable yet more catastrophic. As illustrated in Figure 19, they present relatively low frequencies for low magnitude failures as well as occurrences of high magnitude failures.

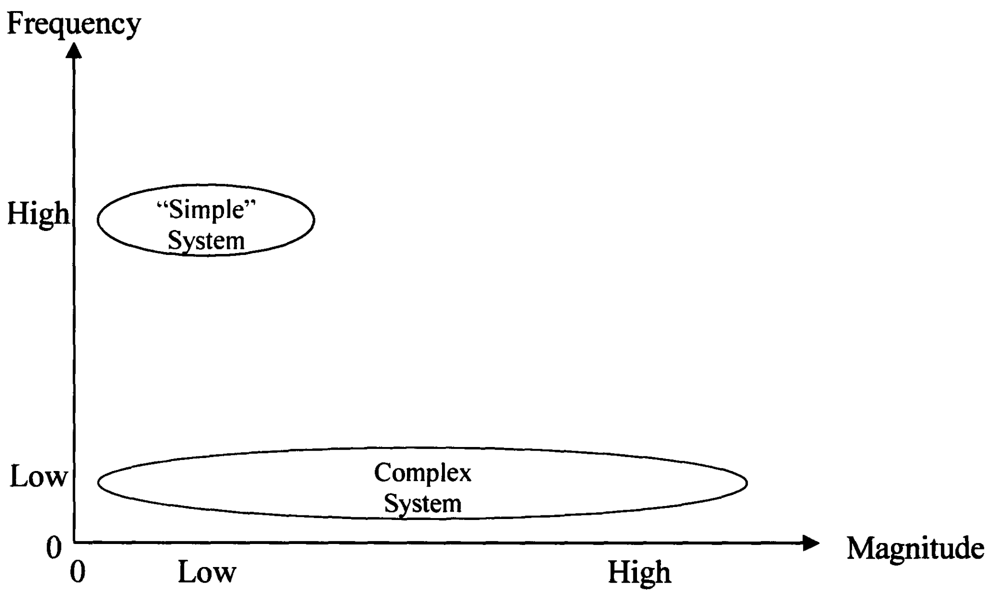


Figure 19 – Risk configuration of complex large-scale systems

"Reliable, yet catastrophic" seems to be one of the characteristics of complex large-scale systems. Acknowledging that other aspects of complex systems may be neglected, we would nonetheless recognize this characteristic as the basis for characterization of complex large-scale systems. Therefore: the more "reliable, yet catastrophic" a large-scale system, the more complex.

II. External Complexity metrics

1. General theory

Based upon the former statement which characterizes the complexity of large-scale systems with an ordered relationship, External Complexity metrics should simply relate failures of large magnitude with those of small magnitude. Now, we explore two such possibilities: C^1_{EXT} and C^2_{EXT}.

The notations used to define the metrics are:

i: magnitude

p_i: frequency associated with the failures of magnitude i

N: the highest magnitude possible

C^1_{EXT} is defined as the ratio of the risk associated with all failures of higher than average magnitude over the risk associated with all failures of lower than average magnitude. Thus, if m is the average magnitude, the risk associated with the failures of higher than average magnitude is: $\sum_{i=[m]}^{N} ip_i$ and the risk associated with the failures of lower than average magnitude is: $\sum_{i=1}^{[m]} ip_i$.

So, $C^1_{EXT} = \dfrac{\sum_{i=[m]}^{N} ip_i}{\sum_{i=1}^{[m]} ip_i}$ where m, the average magnitude of failures (i.e., the average risk per failure) is, by definition, the total risk (i.e., the risk associated with all the failures) ($\sum_{i=1}^{N} ip_i$) over the total number of failures ($\sum_{i=1}^{N} p_i$): $m = \dfrac{\sum_{i=1}^{N} ip_i}{\sum_{i=1}^{N} p_i}$

Here are some examples of the basic behavior of C^1_{EXT}:

$C^1_{EXT} = 0$ when the risk is constant. This occurs when the frequency of each failure (and for all magnitudes) is inversely proportional to its associated magnitude.

$C^1_{EXT} = 1$ when the risk associated with the failures of higher than average magnitude equals the risk associated with the failures of lower than average magnitude.

C^1_{EXT} = 3 when the risk linearly increases with magnitude. This occurs when the frequency of all the failures is constant (i.e., the frequency does not depends upon the magnitude). (The computation of these three typical behaviors is detailed in Appendix H)

C^2_{EXT} is defined as the risk associated with the failures with the top 10% magnitude ($\sum_{i=[0.9N]}^{N} ip_i$) over the risk associated with the failures with the bottom 10% magnitude ($\sum_{i=1}^{[0.1N]} ip_i$).

$$C^2_{EXT} = \frac{\sum_{i=[0.9N]}^{N} ip_i}{\sum_{i=1}^{[0.1N]} ip_i}$$

By construction, C^2_{EXT} emphasizes much more than C^1_{EXT} the prevailing reliability and the tendency to catastrophe of complex large-scale systems.

Here are examples of the basic behavior of C^2_{EXT}:
C^2_{EXT} = 1 when the risk is constant. This occurs when the frequency of each failure (and for all magnitudes) is inversely proportional to its associated magnitude.
C^2_{EXT} = 1 also when the risk associated with the failures with the top 10% magnitude equals the risk associated with the failures with the bottom 10% magnitude.
C^2_{EXT} = 19 when the risk linearly increases with magnitude. This occurs when the frequency of all the failures is constant (i.e., the frequency does not depends upon the magnitude).
(The computation of these three typical behaviors is detailed in Appendix H)

2. Application of the general theory to transportation systems

a. Transcription of the concepts

In order to apply these metrics to the respective environment of the two test bed systems, one must define the systemic failure of the large-scale systems in which the radars are embedded. In order to apply the metrics, one needs to choose the right type of failure and then to identify its frequency, magnitude and associated risk.

The aim of transportation systems is to transport safely people form one point of to another in time. A failure can be characterized by the non fulfillment of this function, i.e., that a person does not reach safely its destination. A failure may be a crash. But, if we consider safety in its most restrictive meaning, a failure is a fatal crash. As a conclusion, for transportation systems:

- a failure is a fatal accident
- the magnitude of a failure is the number of fatalities in a given fatal accident

- the risk is the total fatality (because $\text{Risk} = \sum_i \text{Frequency}_i \times \text{Magnitude}_i$)

b. Transcription of the metrics

Applied to transportation systems the first External Complexity metrics C^1_{EXT} is defined as the total fatality occurring in the higher than average magnitude fatal accidents over the total fatality occurring in the lower than average magnitude fatal accidents.

$C^1_{EXT} = \dfrac{\sum_{i=[m]}^{N} ip_i}{\sum_{i=1}^{[m]} ip_i}$ where $m = \dfrac{\sum_{i=1}^{N} ip_i}{\sum_{i=1}^{N} p_i}$ is the average fatality per fatal accident (i.e., the total fatality over the total number of accidents).

C^1_{EXT} has several advantages. This metrics, as the ratio of numbers of same dimension, is independent of the number of passengers traveling, the time or the distance traveled. Moreover, as the ratio of the fatality in accidents above and below the average fatality per fatal accident, is independent of the carrying vehicle size.
These two characteristics allow pure comparison of the different transportation modes.

Applied to transportation systems the second External Complexity metrics C^2_{EXT} is defined as the fatality occurring in the top 10% fatal accidents over the fatality occurring in bottom 10% fatal accidents.

$$C^2_{EXT} = \frac{\sum_{i=[0.9N]}^{N} ip_i}{\sum_{i=1}^{[0.1N]} ip_i}$$

C^2_{EXT} also has several advantages. As the first metrics, being the ratio of numbers of same dimension, it is independent of the number of passengers, the time or the distance traveled. Moreover, as the ratio of fatality in the top 10% fatal accidents over the one in the bottom 10% is, contrarily to the first one, dependent on the carrying vehicle size.
These two characteristics allow a good measure of the complexity of the different transportation systems if we consider that the potential magnitude of the event is deeply linked to the complexity of the system.

III. Application to the test bed large-scale systems

In addition to the two test bed complex large-scale systems: the air and maritime transportation system, we also study a supplementary system: the land transportation system. To have an even better basis for comparison, we will look at these large-scale systems at three different scales. We study these systems in the case of France, the United States of America and the world. (However, due to the lack of reliable data on the worldwide land transportation system, this system will not be studied at this global level).

1. Characterization of the complexity of test bed large-scale systems

Figure 20, Figure 21 and Figure 22 illustrate the characterization of the complexity of the large-scale systems for France, the United States and the world. The data used to obtain these figures are respectively in Appendix I, Appendix J and Appendix K. Data on air and water transportation focus on public transportation because the test bed radars are designed for this system. Thus, from the sources on air transportation only plane accidents in the category passenger and cargo (e.g., pleasure is excluded) are considered. Just like for air transportation, the water transportation data do not consider recreational boating.

In France, as shown in Figure 20, the water and the air transportation systems are very reliable, mainly in comparison to the land transportation system. Conversely, the air transportation system is more catastrophic than the water transportation system mainly due to the high magnitude accident (113 deaths).
Since reliability is a necessary condition of complex large-scale systems, land transportation which is not very reliable tends to drive relatively less complexity into lower-level elements. On the other hand, the water and the air transportation systems tend to drive more complexity into these lower-level elements. But, the only real complex system tends to be air transportation because of its tendency to catastrophe which the water transportation has not. This analysis can be summarized in:

C_{EXT} (air transportation) > C_{EXT} (water transportation) > C_{EXT} (land transportation)

Studying the case of the United States of America is even more striking. Figure 21 shows that water and air transportation once again meet the first criteria of complex large-scale systems: they are reliable; while land transportation is not. Conversely, air transportation is very catastrophic while maritime transportation only shows small tendency to catastrophe.
So, we come up with the same conclusion as for France:

C_{EXT} (air transportation) > C_{EXT} (water transportation) > C_{EXT} (land transportation).

While the French data may not be sufficiently significant for air and water transportation, the American data for air transportation are better. Besides this restriction, the American air transportation system tends to be much more complex than the French one because of higher tendency to catastrophe. Thanks to this characterization we can identify, in the United States, three distinct levels of complexity in the systems examined.

The relative order of complexity brought out is confirmed in Figure 22 studying the air and the water transportation system on a more significant scale: the world. Both systems are quite reliable. They are also quite catastrophic. Nonetheless, the air transportation system presents higher occurrences of very fatal accidents and therefore would tend to be more complex than the water transportation system.
So, we would say that both systems are highly complex and that:

C_{EXT} (air transportation) > C_{EXT} (water transportation)

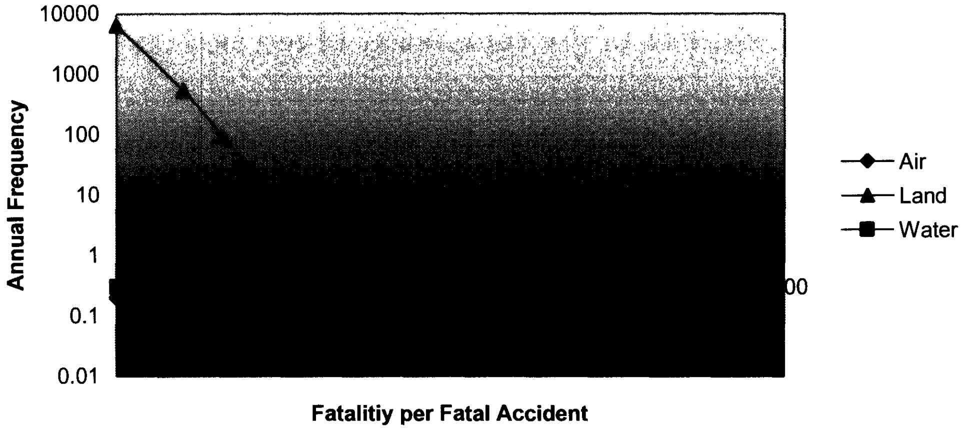


Figure 20 – France: characterization of the complexity of transportation systems

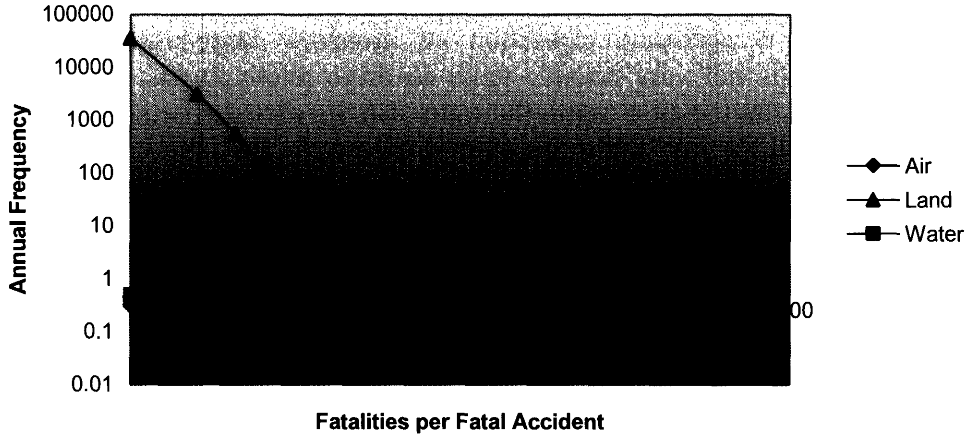


Figure 21 – United States: characterization of the complexity of transportation systems

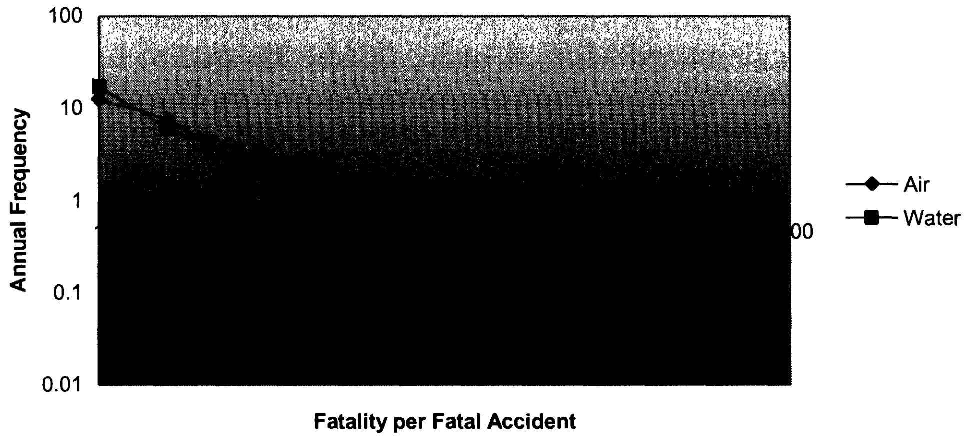


Figure 22 – World: characterization of the complexity of transportation systems

2. Measurement of the complexity of test bed large-scale systems

The values of External Complexity computed with the two metrics for the different large-scale systems studied are presented in Table 16 and plotted in Figure 23 and Figure 24.

C_{EXT}

	Air	Water	Land
World	3.611	3.529	
U.S.	5.299	1.579	0.2429
France	3.093	3.000	0.2373

C_{EXT}

	Air	Water	Land
World	118.5	71.92	
U.S.	228.9	11.63	2.232
France	113.0	8.000	2.206

Table 16 – External Complexity of the large-scale systems studied

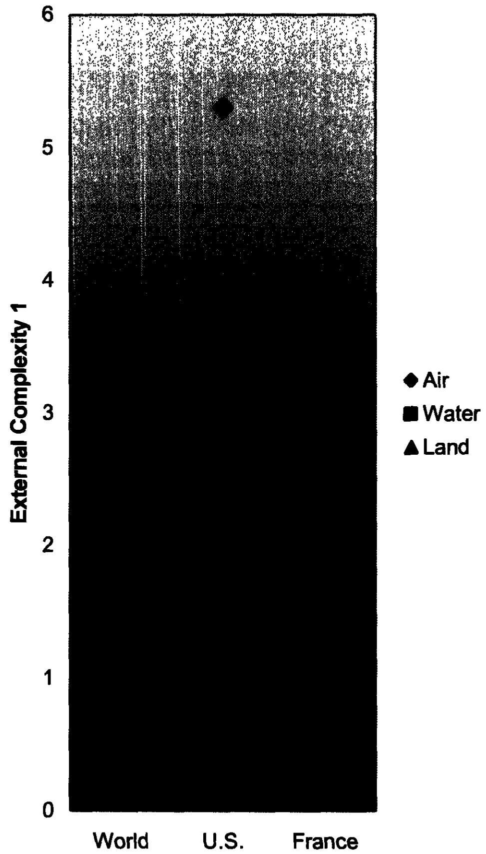


Figure 23 – External Complexity metric 1

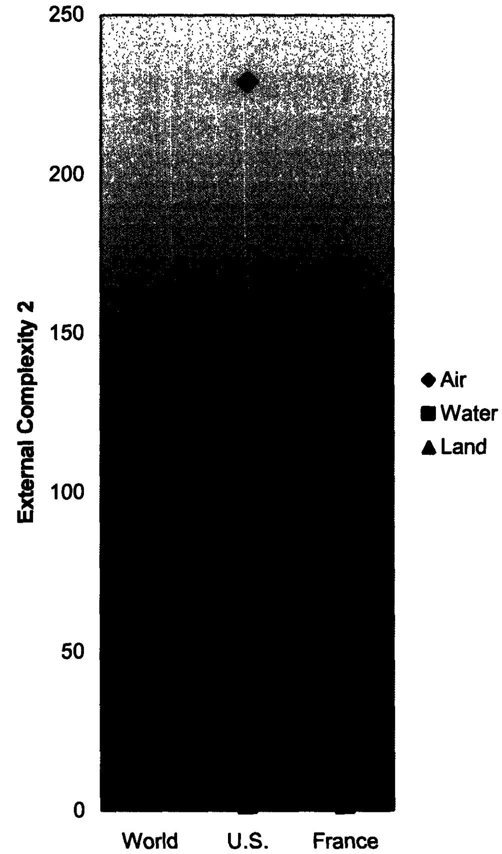


Figure 24 – External Complexity metric 2

The two metrics C^1_{EXT} and C^2_{EXT} confirm the qualitative results for all the geographical scales obtained form the characterization of complex large-scale systems presented in the former section:

C_{EXT} (air transportation) > C_{EXT} (water transportation) > C_{EXT} (land transportation)

These metrics also emphasize that air transportation tends to be a complex system while land transportation tends to be a "simple" system.

Page intentionally left blank

E. Results and discussion

This section presents the results and some comments on complexity. The discussion aims at more abstractly developing and challenging the concepts and the methodology used in this thesis. It also aims at providing a holistic understanding of complexity.

I. Remarks

1. Remarks on the metrics

Regarding Internal Complexity, it is important that the process to obtain the Reference Decomposition is reproducible. For any system, the Reference Decomposition (and therefore Internal Complexity) is unambiguously determined following the methodology. The allocation of functions is the critical step: the top down approach enables the uniqueness of the outcome. The basic function attributed to a higher-level element is its macroscopic function which does not take into account the potentially different functions performed by its sub-elements. This occurs firstly because these different functions are usually scarce and secondly because the holistic basic function of the elements prevails in the complex configuration of the system. The complex integration of these basic functions is what we claim is responsible for the emergent behavior.
A good understanding of the process also clarifies the reason why the Reference Decomposition is built from functional block diagrams. The functional block diagrams of the system, the subsystems and the elements are used to attribute the function to each element and, if necessary, to develop the decomposition further down in order to identify the mono-functional elements. These diagrams are also used to identify the links, their directionality and the substance flowing.

Regarding External complexity, C^1_{EXT} and C^2_{EXT} seem to be useful metrics for complex large-scale systems, but one has to remember that they are based on only one characteristic of these systems (higher reliability and higher tendency to catastrophe). Even though we believe that this characteristic is sufficient to characterize complex large-scale systems, the application of the two metrics to the test bed systems already allows us to bring out potential differences in the relevant aspects of complexity. While the two metrics are fully coherent, they do not have exactly the same behavior. The French water transportation system is complex for C^1_{EXT} (Figure 23 p. 59). Since it is a reliable system (Figure 20 p. 58), this metric seems to emphasize reliability as a major feature of complex systems. Conversely, for this system, C^2_{EXT} is low (Figure 24 p. 59); the system not being prone to catastrophes (Figure 20 p. 58), this metric seems to emphasize the tendency to catastrophe as a major feature of complex systems.
C^1_{EXT} is quite similar for the worldwide water and air transportation systems while C^2_{EXT} is much higher for the worldwide air transportation system than for the other system. Now, both systems tend to be equally reliable (mainly in comparison to the land transportation system) but the air is more prone to catastrophes than the water. So, C^2_{EXT} seems to emphasize more the discrepancy between reliability and tendency to catastrophe. Finally, we note from Table 16 (p. 59) that C^2_{EXT} clearly separates, for all scales, the three different systems attributing to them very distinct complexities.

2. Remarks on the approach to complexity

Since people do not fully agree on the characteristics and the metrics of complexity, the framework which identifies three sets (Figure 1 p. 11) offers the opportunity to define different kinds of complexity focusing on different characteristics in each set. This division also offers the opportunity to develop different metrics to compute the identified complexities. Indeed, this framework enabled us to propose an Internal Complexity metrics taking into account the number of links, the number of elements, the function and hierarchy of the elements, an Interface Complexity metric based upon the information content of the probability of failure of the system as it is used in its environment and External Complexity metrics which deal with the risk configuration of large-scale systems emphasizing their reliability and their tendency to catastrophe.
Moreover, one of the rationales for identifying three sets and focusing on three complexities is the distinction between the complexity directly created by man and the one indirectly created. Designers generate directly the complexity of the radar while the complexity of the air transport system is exogenous in that it is the result of many human actions undertaken somewhat independently.
This framework also allows us to put aside from quantification one aspect of complexity. This thesis does not tackle the complexity of the political system because it is difficult to define and quantify it meaningfully. Some people argue that the highest complexity exactly lies where we are not able to compute it yet, i.e., in the political system. Here, the complexities computed are closely linked with the ability to quantify them. Further studies may well expand the frontiers of knowledge and then the areas where complexity can be computed. Nonetheless, not to be so restrictive, it is proposed in E.III.2 (p. 68) a speculative framework to attempt to better perceive the role of the socio-political complexity.

Internal complexity which stems from the network (i.e., hierarchy and links) of functions in the system must be unambiguously defined. Now, the Reference Decomposition and then internal complexity heavily depend upon the choice of the basic elements. But we believe that the set of the five basic functions we use is fully representative of the functions elements can achieve. There may be other generator sets of basic functions but they are certainly very close to the one we use; the resulting complexity must be very similar too because the alternative internal complexities would capture the same idea.

Interface Complexity has to be computed in the normal condition of use of the system. For radars, the usual range is a trade-off between what could be done ideally and what can be done reasonably. As the range is decreased to account for the performance of the possible radar, more radar units are needed to achieve the required overall system needs. The higher numbers of radar are the way that the system meets its overall complexity while controlling the internal complexity of individual systems. Multiplying the radars and increasing their practical range (while keeping an acceptable level of detection): both actions impact external complexity. Considering the argument on maneuverability and self-organization [12], the transportation system shaped by the political one tends to find its optimal complexity. Once external complexity is set, interface and internal complexity are largely determined. What is true for the range is also true for the probability of false alert. The typical Pfa is also a trade-off because increasing the accuracy of detection also increases the occurrence of false alert. The threshold being lower to improve detection, noise is more frequently interpreted as a false target. As a summary, typical data take into account what is required of the system from its environment and what can be guaranteed by the designers.

Finally, regarding external complexity, Table 16 – External Complexity of the large-scale systems studied (p. 59) shows that for the two metrics C^1_{EXT} and C^2_{EXT}, the worldwide water transportation system is much more complex than any of the two national water transportation systems studied. The reason may be that the world is the "right" scale for the system to be complex. Scale seems to be a driver of external complexity. However, for both metrics, the U.S. air transportation is more complex than the two other air transportation systems. The U.S. system has an intensity that the two others may not have: the French one is too small and the worldwide one is too dispersed. Intensity seems to be also a driver of external complexity. This intuition is confirmed studying the land transportation system. Its complexity is low but it is slightly higher for the U.S. than for France: while the U.S. system is larger than the French one, their intensity may be comparable.

II. Results and findings

Table 17 summarizes the three complexities of the two test bed systems where the External Complexity shown is C^2_{EXT} applied to the world. Internal, Interface and External Complexities of the ATC radar are clearly higher than the ones of the maritime radar.

	ATC Radar	Maritime Radar
C_{INT}	350.2	89.12
$C_{INTERFACE}$	3.321	0.9999
C_{EXT}	118.5	71.92

Table 17 – The three complexities of the two test bed systems

Plotting the three complexities in the three dimension diagram of Figure 25 one can notice that the triangles linking the three complexities of each system do not cross. This illustrates that the ATC radar and its environment are more complex than the maritime radar and its environment.

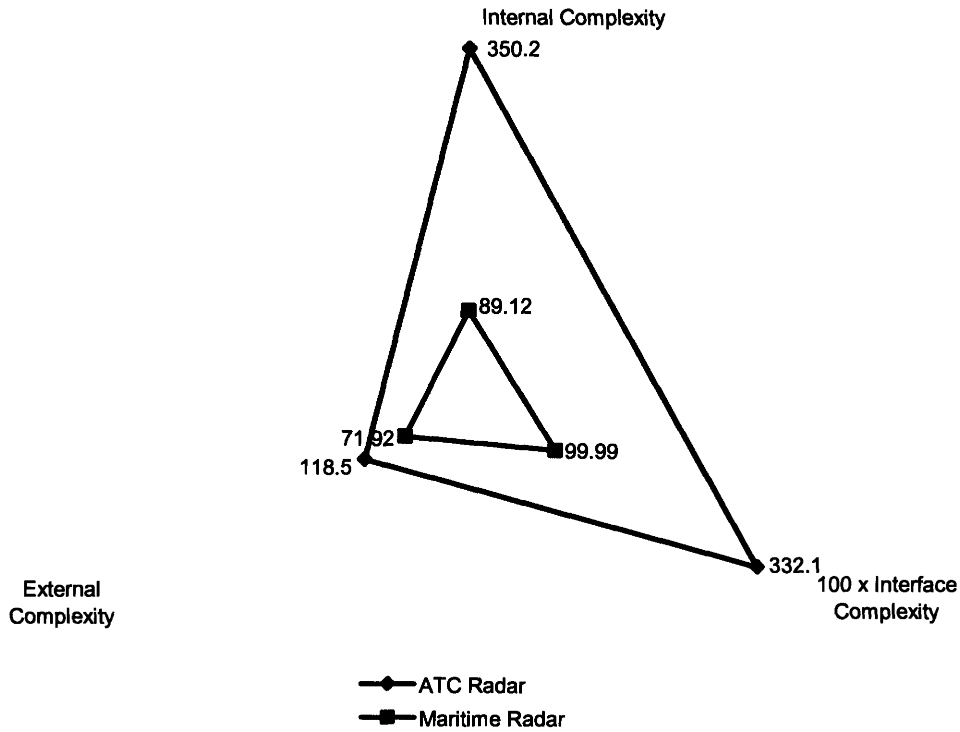


Figure 25 – Interaction between Internal, Interface and External Complexity metrics

From Figure 26 we note that Internal and Interface Complexities are somewhat proportional. Indeed, the ratio of Interface Complexity (3.3) and Internal Complexity (3.9) between the two set bed systems is quite similar. If it were confirmed on a wider variety of systems, a simple linear relationship between interface and internal complexity would appear to be factual.
Interface complexity which was first introduced as a tool to understand the propagation of complexity from the environment to the system now appears to be much more than that. The combination of interface complexity and internal complexity seems to give a full and accurate insight into the complexity of a complex system. An attempt to analyze qualitatively this relationship is presented in section E.III.1 (p.67).

Plotting in the 3D diagram of Figure 27 the three complexities where the External Complexity is C'_{EXT} applied to United States, we note that the blue triangle for the ATC radar and the pink one for the maritime radar are nearly homothetic. The dotted purple triangle is the theoretical extrapolation from the maritime radar to the ATC radar. The extrapolated shape of the complexity of the ATC radar is very close to the real one.

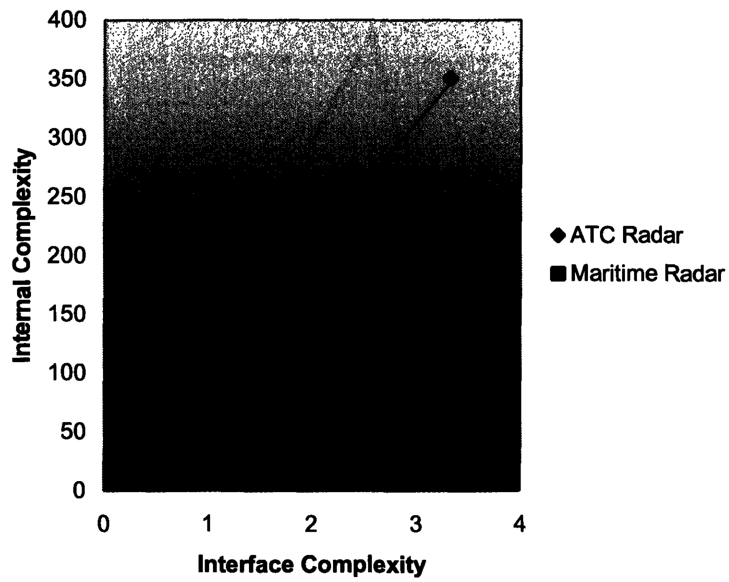


Figure 26 – Proportionality between Internal and Interface Complexity

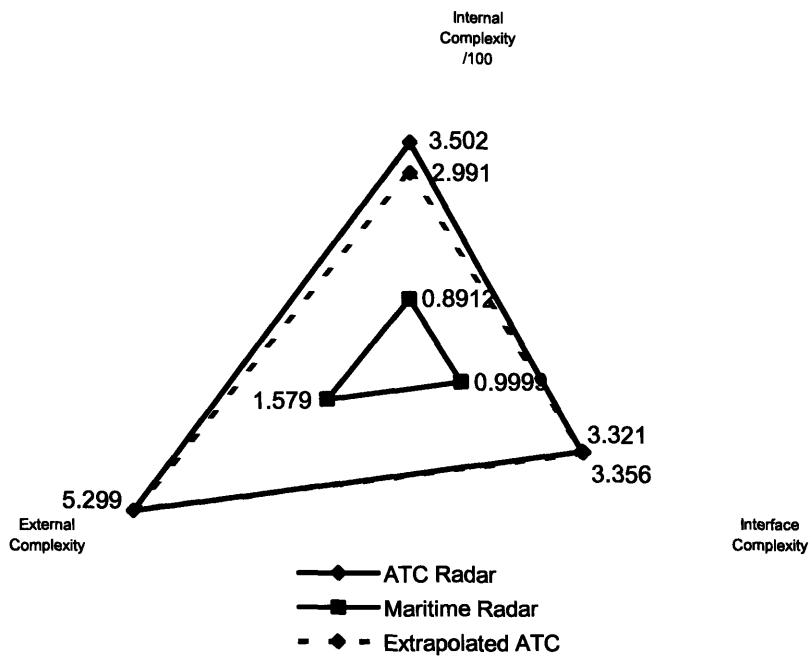


Figure 27 – Homothetic complexities

Even if this result may not be generalized, at least it supports the idea already illustrated in Figure 25 that the three complexities are closely linked.
These complexities were identified on the complexity framework because it was both relevant and convenient to do so. Now, it seems that they really need to be brought together to describe an overall complexity. This multifaceted overall complexity seems to be accurately described integrating the three dimensions of complexity identified (or even more). More than a propagation of complexity from the environment to the complex system itself, it seems that there is a whole dynamics of complexity. The following section (E.III.1) attempts to illustrate qualitatively this dynamics.

A last finding deals with the relationship between complexity and price. The ratio of Internal Complexity for the two set bed systems is 3.93 while the price ratio is roughly 15. The price seems to increase much faster than complexity. Indeed, the price may depend more upon the incremental complexity (the increase in complexity form the old to the new version of a system) than upon the absolute complexity. Conversely, from the result which emphasizes the proportionality between internal and interface complexity, the absolute complexity seems to be responsible for the systemic behavior.

III. Interplay of the different complexities

1. Relationship between internal, interface and external complexity

The link between internal, interface and external complexity may be clearly understood through the analysis of a particular case. The Interface Complexity of a system tends to infinity ($C_{INTERFACE} \rightarrow \infty$) when the system fully achieves what it was designed for ($P \rightarrow 0$). Now, "what the system was designed for" is the outcome of the trade-off between the external requirements and the internal capability. It is, up to a certain extent, the potentially maximum complexity. When Interface Complexity which measures the degree of achievement of this trade-off (i.e., how close you are from it) tends to infinity, internal complexity of the complex system tends to the maximum complexity it is required to have by its environment.

A more theoretical way to understand the mechanism is to look at the complexity defined by Suh. This complexity is high because the Functional Requirements are "far" and "hard" to reach. The stringency of the Functional Requirements is the source of internal complexity. The set of Functional Requirements being "far", designers must design a complex system in order to reach them. Internal complexity tends to be high and it becomes harder to achieve the Functional Requirements. The outcome being less probable, the systemic information content increases. The "hard", as the consequence of the "far", is in fact tightly linked with internal complexity. The Interface Complexity metric proposed attempts to catch the dynamics previously described.

Interface and internal complexity are also tightly linked. The stringency of the requirements which is partly responsible for high interface complexity comes from external complexity. A complex large-scale system fulfilling higher functions may either ask more of its composing systems or have more composing systems (according to the "divide and conquer" rule). A balance must be found between the internal and external complexity so that, at the end, the emergent behavior of the complex large-scale system is correctly achieved. This balance results in the proper allocation of complexity between internal and external complexity. Complexity has both positive and negative consequences; complex systems require a certain amount of complexity to achieve their function. The balance is reached when that complexity is properly allocated so that the overall function is correctly fulfilled. This Pareto equilibrium imposes the level of internal and external complexity and, consequently, the level of interface complexity.

More specifically, following this line of thought, two complexities: internal and interface complexities are directly related. Internal complexity attempts to quantify how the sum of functions creates the emergent behavior of the system. Conversely, interface complexity attempts to quantify how well the emergent behavior is created. Looking at these two complexities together seems to give a pretty accurate view of the overall complexity of a complex system.

The example of the ATC radar may clarify this argument. The Internal Complexity metric may underplay certain aspects of complexity because it only takes redundancy into account through Link Complexity. Conversely, Interface Complexity emphasizes the importance of redundancy. Indeed, the accepted outcomes of redundancy which are higher availability (if one channel breaks down, the other allows the system to work) and higher accuracy (if one channel does not perform well, the other one is used) are taken into account in the probability of failure and then in Interface Complexity. So, looking at these two complexities in parallel may help to better describe the overall complexity of the ATC radar.

The need for two measures to properly assess complexity is also made in the literature. Seth Lloyd [20] criticizing complexity metrics based on Logical Depth and Breadth [21] argues that information alone is not sufficient to describe complexity. Thus, Interface Complexity based on information content cannot be self sufficient. From a systems standpoint, he argues that what is missing in Logical Depth is the measuring of complexity based on the function of the system. Interestingly, our formulation of Internal Complexity exactly achieves this last feature.

From a different point of view, the reason why external and internal complexities (and more broadly the three complexities) are linked is simply because large-scale complex systems are recursive. The external complexity of a complex system is part of the internal complexity of the large-scale system in which it is embedded; the internal and external complexities of a complex system correspond to the internal complexity of the large-scale system.

2. Attempt to link the external and the socio-political complexity

Let's start with an example. The argument made in D.I (p.52) stating that the tendency to catastrophe characterizes complex large-scale systems may be contested. Some people may argue for instance that, in the air transportation system, high magnitude failures tend to have very high magnitude, which are much higher than the most catastrophic failures in land transportation because planes are bigger than cars and therefore no conclusion on complexity can be drawn from our analysis. While their argument is of course true, their conclusion is false. The reason why the system is reliable is the same reason why it is catastrophic. The reason why planes can be big is the complexity and the reliability of the air transportation system. If this system was not that complex and that reliable, no plane manufacturers would have ever built such big planes (e.g., Boeing 747-400 or Airbus A-380) for fear that repetitive accidents would hinder the confidence people put in this plane, the development of the plane and their profitability. It is precisely because the system is complex and reliable that big planes may exist. It is also precisely because of the complexity, the reliability and the existence of big planes that catastrophic events exist in this complex system. Finally, we come back to the initial conclusion that complex systems are reliable, yet catastrophic.

A hypothetical more complex land transportation system would be, for instance, a transportation system with faster semi-automatically driven cars and buses which maximizes the passenger flow. Seldom catastrophic failures such as the crash of hundreds of cars due to a major failure in the overall positioning system (despite its redundancy, reliability...) might be of high magnitude because people would not have been able to avoid the collision by mere lack of experience due to general over-confidence in the reliability of the semi-automatic driving system. Such a crash, even very seldom would be associated with very high fatality. This hypothetical land transportation system would be more complex and much more costly (which is probably part of the reason why it does not exist).

The latter argument on cost is non negligible in the overall dynamics of complexity. Wealth allocated to the design or to the improvement of a system is what allows the system to be more complex (since complexity has a cost), to have its complexity better allocated and therefore to be more reliable.

The importance of the political system can be described using a Systems Dynamics model. We remind that in Systems Dynamics [22], arrows between two variables represent the influence from a variable to another and the polarity of the arrows is the sign of the correlation between the variables linked. A double bar on arrows represents a delay in the influence.

Figure 28 illustrates the socio-political dynamics of external complexity. External complexity has two opposite consequences on risk: reliability which tends to decrease risk and tendency to catastrophe which tends to increase it. Reliability and tendency to catastrophe also impact risk perception in the same way. Nonetheless, regarding risk perception, the impact of catastrophes is much stronger than the influence of reliability. Risk perception also takes into account the effective risk. Then, risk perception and risk are responsible for policies that shape the large-scale system while increasing its complexity.

This Systems Dynamics model unfolds over three periods of time with different relative strength of loops. First, the Catastrophe-Perception loop prevails, then the Reliability-Perception loop and finally the Reliability-Risk loop.

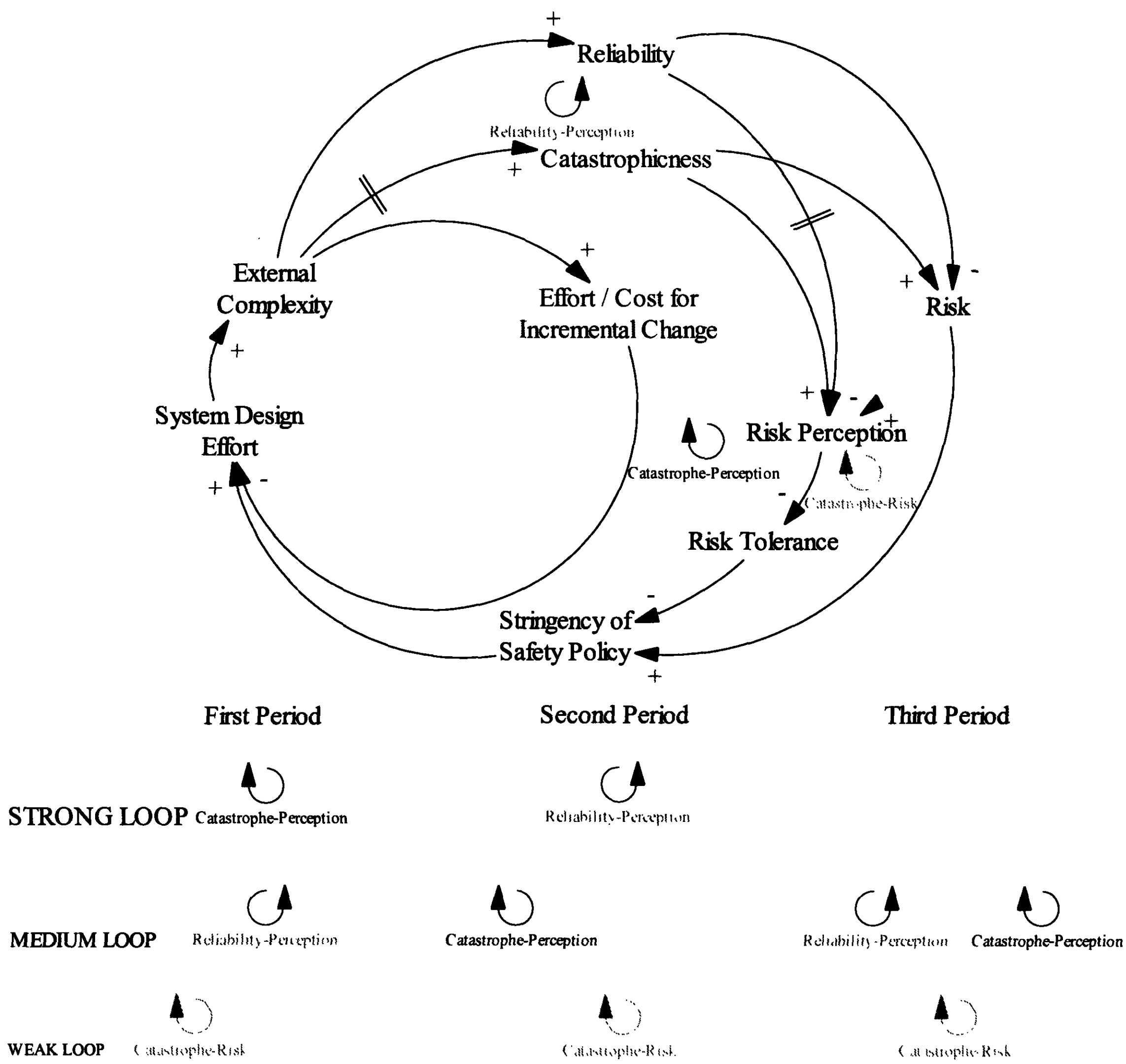


Figure 28 – Socio-political dynamics of external complexity

Figure 29 summarizes the influence of risk perception and risk in the political decision-making process and may clarify the lower right part of the Systems Dynamics model (Figure 28). The determination of the level of risk perception is based upon the lexicographic consideration of, firstly, the magnitude and secondly the level of risk (which is the magnitude of events times their frequency). Magnitude is the main driver of risk perception.

Magnitude of events, due to its huge impact on people's mind is the first determinant of risk perception. If it is high, risk perception will be high too. If it is low, risk perception will be low too. (This is why the color used to qualify risk perception in Figure 29 is the same as the one used to qualify magnitude).

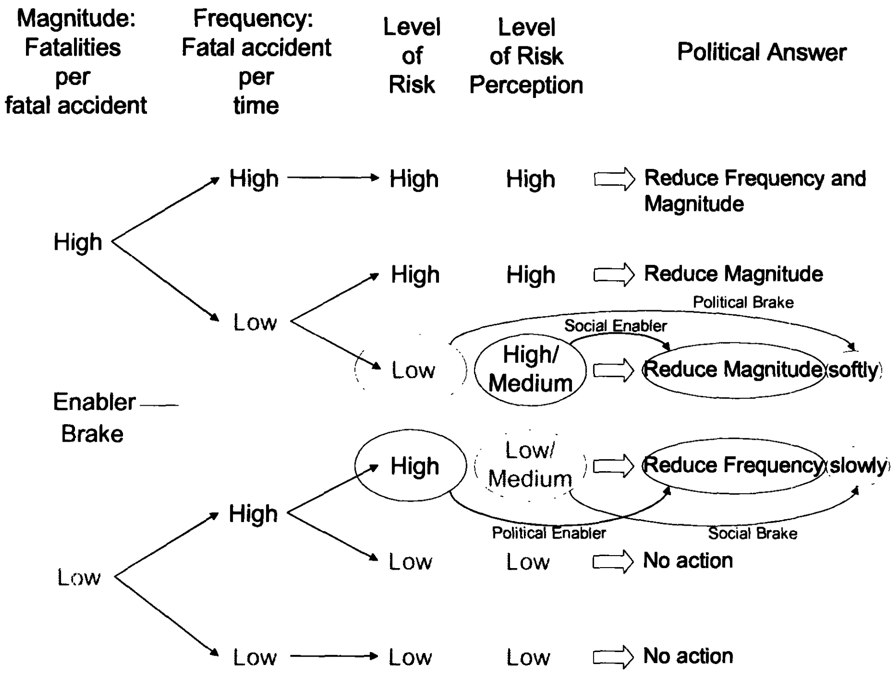


Figure 29 – Political decision-making process to deal with risk

The effective level of risk can simply moderate this first categorization of risk perception. If the magnitude of an event and the level of its risk are opposite (i.e., if one is high and the other is low), the level of risk perception formerly categorized by the magnitude will be moderated by the level of risk. For example, the outcome of high magnitude and low level of risk is high/medium level of perception of risk ; the outcome of low magnitude and high level of risk is low/medium level of risk perception. This process is illustrated by the pink and blue arrows in Figure 29.
Moreover, we mention that risk perception encloses something broader than mere magnitude and that falls under the concept of tendency to catastrophe. Even if it is ethically hard to acknowledge, the "importance" of the person who dies in an accident clearly influences people's mind and therefore the political decision-making process. The death of a star is much more striking than the death of an anonymous person.

Besides, policy actions shape the complexity of the transportation systems. The implementation of political decisions is at the source of the socio-technical complexity of transportation systems.

Thus, the transportation system and the political system are in fact tightly linked in a feedback loop. This is the reason why we can easily group them under the name of "environment". The Systems Dynamics model indicates that, at the end, a dynamic equilibrium is reached. The complexity of the environment emerges from this dynamic equilibrium between the two systems.

3. Holistic approach to complexity

The higher the socio-political complexity is, the more numerous and the more stringent the requirements on the complex large-scale system are. Higher external complexity results in higher interface complexity due to necessary better fulfillment of the requirements. Now, to correctly fulfill the numerous stringent requirements, internal complexity must be high. The environment dictates the requirements to be correctly fulfilled by the system. External and internal complexities echo each other thanks to interface complexity. On the ATC radar example, the goals of the large-scale system which are safety, efficiency and cost-effectiveness are shaped by the Political system. These goals, to be fulfilled, require complex functions performed by the complex systems which need to implement a variety of characteristics and techniques to achieve them. Part of this holistic dynamics of the ATC radar in its environment is illustrated in Figure 30.

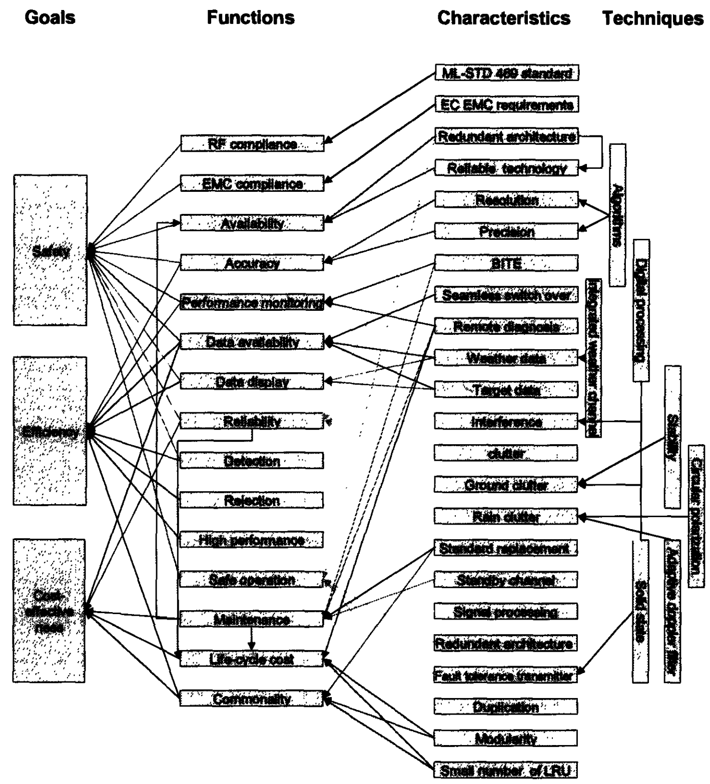


Figure 30 – Partial illustration of the holistic approach to complexity

IV. Future work

The first future work would be to continue attempting to improve the metrics. Even if the Internal Complexity metric identified in B.IV.2 (p. 37) is a good metric, we may propose some improvements. Quality is a fundamental concept in internal complexity. For this purpose, C. Magee and O. de Weck [2] distinguish four different types of links: flows of energy, matter, information and value but we do not really take advantage of this distinction in the computation of complexity. A way to integrate it in the algorithm is to subdivide each basic function into four supplementary functions depending upon the operand on which it operates. Thus, for instance, instead of having simply the basic function: "transform", we would have the four basic functions: "transform energy", "transform matter", "transform information" and "transform value". This will increase the number of vocabulary available (V_{max}) from 5 to 20. The drawback of this method is that the representation of complexity with the Reference Decomposition will be less readable because the representation would contain much more different and difficult to distinguish colors.

Integrating heterogeneity among layers may improve the metric. However, it is hard to identify the heterogeneous configuration which is responsible for higher complexity: is it 10/4/10 or 12/8/1 for instance? Nonetheless, it worth noticing that Link Complexity modestly integrates the idea of heterogeneity since it is normalized for homogeneous systems (i.e., it equals 1 for systems having each of its elements linked once and only once).

Weighing the links may also improve the metric. The normalized density used for Link Complexity (C_L) could become a normalized weighted density. The weighted density, which is the sum of the links weighed by their intensity over the maximum number of links weighted by the higher intensity found in the system, is normalized to 1 for a system having each of its elements linked once and only once by links of equal intensity.

Thus: $C_L(E,L) = \frac{\sum_{i=1}^{L} w_i}{w_{Max} \times E}$. However, it is hard to identify the intensity of each link and even harder to combine weighted links of different nature (energy, information, matter...).

Loops would also worth being introduced in Link Complexity but there are hard to define meaningfully (a system may present loops but some are irrelevant for complexity).

External Complexity metrics may also be improved. One of the drawbacks of the two metrics proposed is that they are based upon statistical data and that their accuracy heavily depends upon the amount of data collected. France not being really a large-scale system (due to its small geographic span), the data for the air and the water transportation system (Appendix I p. 92) may not be significant enough. This may explain the reason why the values of the two metrics C^1_{EXT} and C^2_{EXT} are very different for the French water transportation system. Noise may be responsible for this apparent inconsistency. The External Complexity for those two systems may not be fully reliable. Nonetheless, one can state that complex systems generate a lot of data and the lack of significance of data may only mean that the system is not a complex system.

More conceptually, complexity and mainly external complexity seem to be plural. The differences between the two External Complexity metrics and the variety of features external complexity may have (E.I.2) make us realize that complexity is multi-faceted. Having this plurality in mind, External Complexity metrics would by highly improved attempting to assess the variety of complexity instead of focusing on only one aspect as we have done.

A second future work may be the improvement of the framework. What may be missing in the framework to study complexity is a global approach. To facilitate conceptualization, quantification and understanding, we first identified three sets and three complexities and then we tried to link them. Since the conclusions drawn seem to indicate that the three complexities are linked, it would worth trying to study complexity in a more holistic way. Of course, this work will require a huge conceptualization and formalization effort.

A third work would be to confirm the results of this thesis studying a wider variety of complex systems. The relationships between interface and internal complexity as well as the one between internal/interface complexity and other variables such as cost, price or engineering effort would be interesting to further investigate.

CONCLUSION

The quantification and the analysis of the three complexities identified in the proposed framework to study complexity allow us to draw several conclusions.

The Internal, Interface and External Complexity metrics calculated for the Air Traffic Control radar are higher than those calculated for the maritime radar. In the case of the United States (for C^1_{EXT}) the three complexities of the two test bed systems are even nearly homothetic. These results highlight the close relationship between the three complexities, the influence of external complexity on internal complexity and the need for a holistic approach to complexity.

For the two test bed systems the more rigorous and quantitative complexity metrics: Interface and Internal Complexity are approximately linearly related. This relationship should be further investigated and these two dimensions of complexity should also be studied in parallel in order to infer an overall complexity of a complex system.

REFERENCES

[1] Engineering Systems Division, MIT, "Appendix A: ESD Terms and Definitions (Version 13)", *ESD Symposium Committee Overview*, May 2002, MIT website: http://esd.mit.edu/WPS/ESD Internal Symposium Docs/ESD-WP-2003-01.20-ESD Internal Symposium.pdf

[2] Christopher L. Magee and Olivier L. de Weck, *An Attempt at Complex System Classification*, MIT website: http://esd.mit.edu/WPS/ESD Internal Symposium Docs/ESD-WP-2003-01.02-ESD Internal Symposium.pdf (Magee, de Weck 2002)

[3] Mark W. Maier and Eberhardt Rechtin, *The Art of Systems Architecting*, CRC Press, Boca Raton, 2002.

[4] Basem El-Haik and Kai Yang, *The Components of Complexity in Engineering Design*, IIE Transactions, 1999.

[5] Joe Sussman and Rebecca Dodder, *The Concept of a CLIOS Analysis: illustrated by the Mexico City Case*, 2003.

[6] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, 1949.

[7] Nam P. Suh, *Axiomatic Design: Advances and Applications*, Oxford University Press, New York, 2000.

[8] Dan Braha and Oded Maimon, *The measurement of a design structural and functional complexity*, IEEE Transactions on Systems, Man, and Cybernetics–Part A: Systems and Humans, 28:527–535, 1998.

[9] Joshua D. Summers and Jami J. Shah, *Developing Measures of Complexity for Engineering Design*, ASME 2003 Design Engineering Technical Conferences and Computer and Information in Engineering Conference, Chicago, 2003.

[10] Y. Bar-Yam, *Dynamics of Complex Systems*, Addison-Wesley, Reading, MA, 1992.

[11] Joel Moses, *Three Design Methodologies, their Associated Organizational Structures and their Relationship to Various Fields*, March 2004.

[12] Stuart Kauffman, *At Home in the Universe, The Search for the Laws of Self-Organization and Complexity*, 1995.

[13] Hebert A. Simon, *The Architecture of Complexity*, Proceedings of the American Philosophical Society, Vol. 106, No. 6, December 1962, p. 467-482.

[14] Linda P. Beckerman, *Application of Complex Systems Science to Systems Engineering*, John Wiley & Sons, Inc. Systems Engineering 3: 96-102, 2000.

[15] G. J. Chaitin, *Algorithmic Information Theory*, Cambridge University Press, Cambridge, 1987.

[16] Joshua D. Summers, *Algorithm for Complexity Connectivity*, Technical Report, Clemson Research in Engineering Design and Optimization Lab, Clemson University, Clemson, SC, CREDO/AID/complexity/04-01, 2004.

[17] Personal correspondence with Alain Johann.

[18] Nam P. Suh, *Axiomatic Design: Advances and Applications*, Oxford University Press, New York, 2000, p. 473-474.

[19] J.M. Carlson and J. Doyle, *Complexity and Robustness*, Proceedings of the National Academy of Science of the United States of America, p. 2538-2545, vol. 99, 2002.

[20] Seth Lloyd, *The Calculus of Intricacy*, The Sciences, 30(5):38–44, 1990.

[21] Seth Lloyd and Heinz Pagels, *Complexity as Thermodynamic Depth*, Ann. Phys., 1988.

[22] John D. Sterman, *Systems Thinking and Modeling for a Complex World*, Irwin McGraw-Hill, 2000.

[23] Bureau d'Enquêtes et d'Analyses pour la sécurité de l'aviation civile (BEA), 1997-2002: http://www.bea.aero/francais/rapports/rap.htm, 1994-1996, 2003: additional data directly provided.

[24] L'Observatoire National Interministériel de Sécurité Routière (O.N.I.S.R.), 1993-2003

[25] International Maritime Organization (IMO), Database: ACCESS, 1991-2003

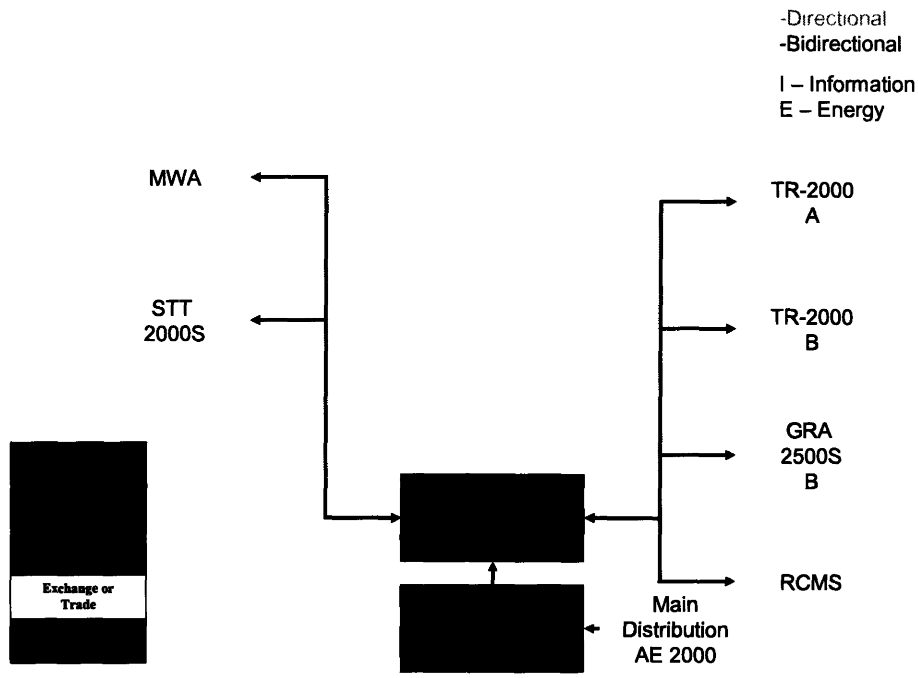
[26] National Safety Board, http://www.ntsb.gov/aviation/Paxfatal.htm, 1982-2003

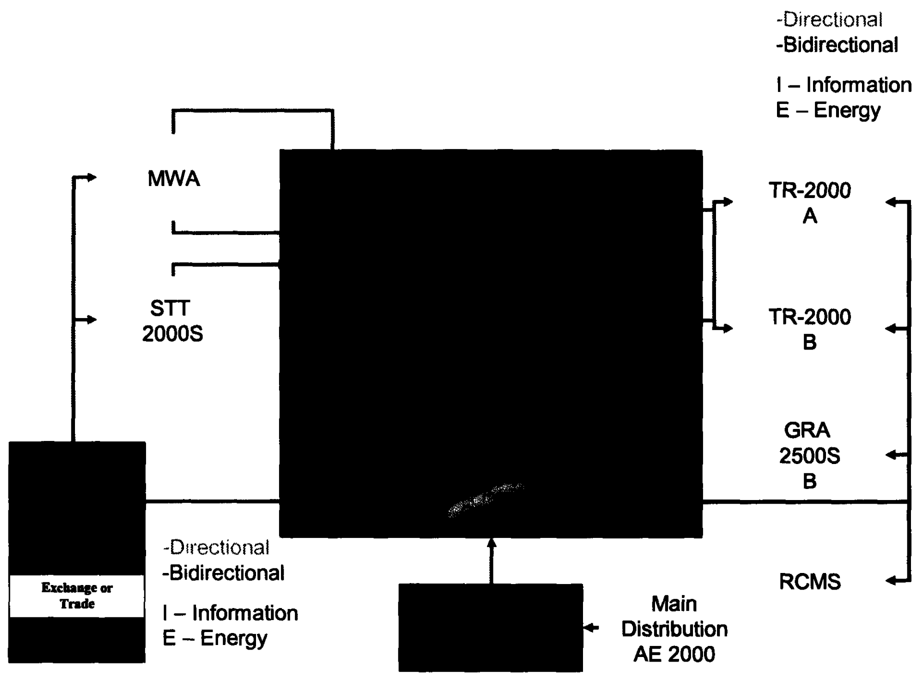
[27] National Center for Statistics and Analysis (NCSA), National Highway Traffic Safety Administration (NHTSA), Report generated by the Information Services Branch, NCSA, CMS# 2004.00686; 4_00686.SAS; B.K.; August 3, 2004 10:36, 1976-2002

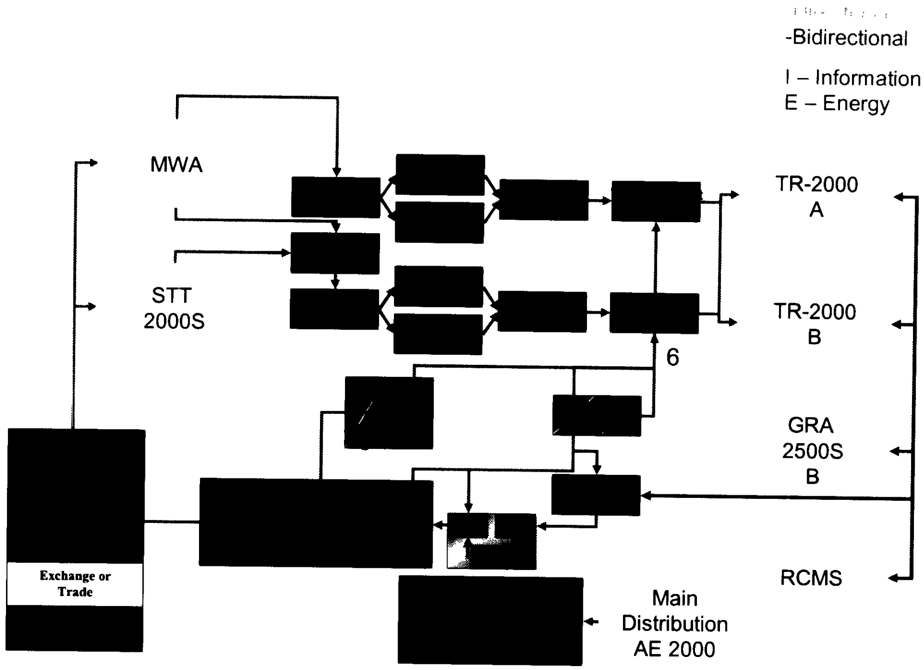
[28] International Civil Aviation Organization (ICAO), Database: ECCAIRS, 1994-26/7/2004

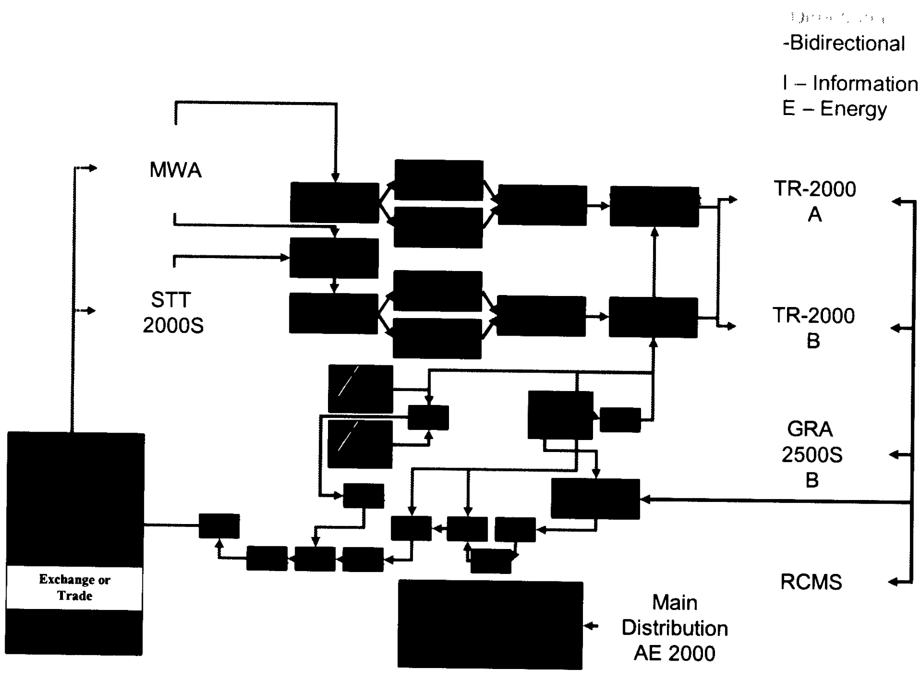
APPENDICES

Appendix A: Complexity Diagrams of the GRA

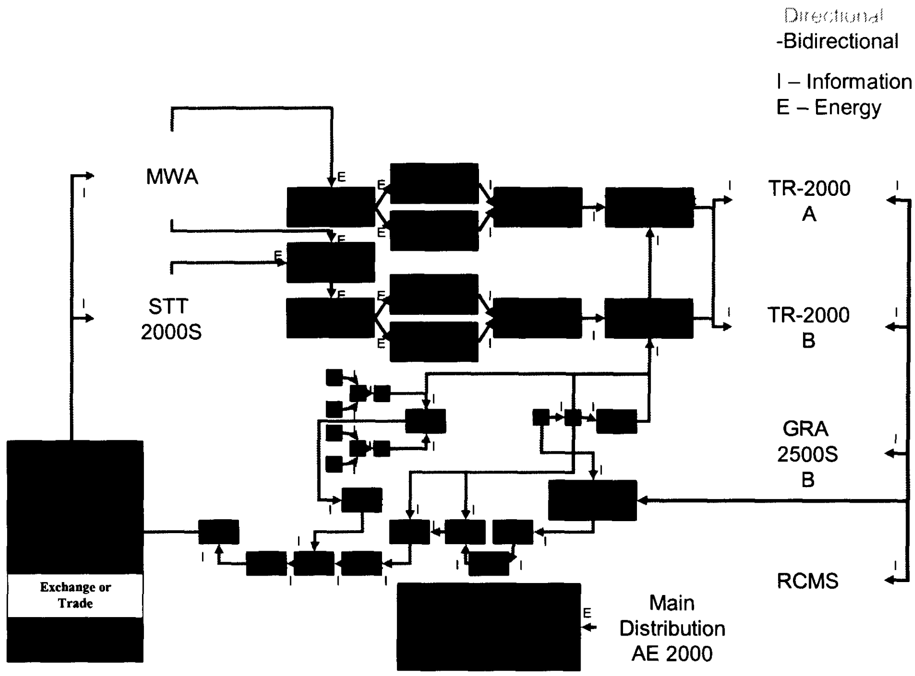




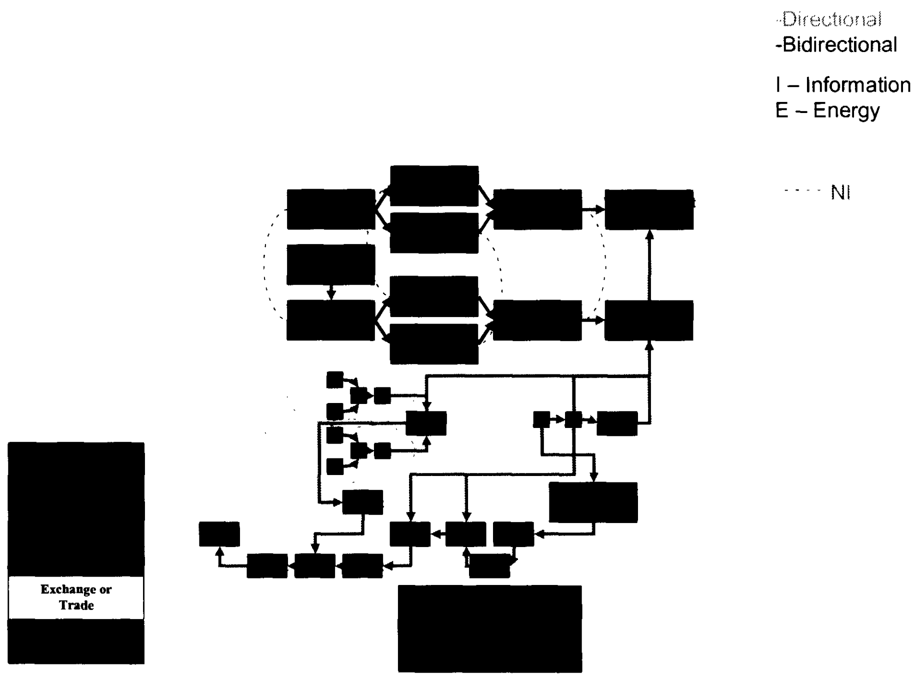

-Bidirectional
I – Information
E – Energy
MWA
STT
2000S
TR-2000
A
TR-2000
B
6
GRA
2500S
B
RCMS
Exchange or
Trade
Main
Distribution
AE 2000


-Bidirectional
I – Information
E – Energy
MWA
STT
2000S
TR-2000
A
TR-2000
B
GRA
2500S
B
RCMS
Exchange or
Trade
Main
Distribution
AE 2000

The Reference Decomposition is:



The diagram used to compute GRA Complexity is:

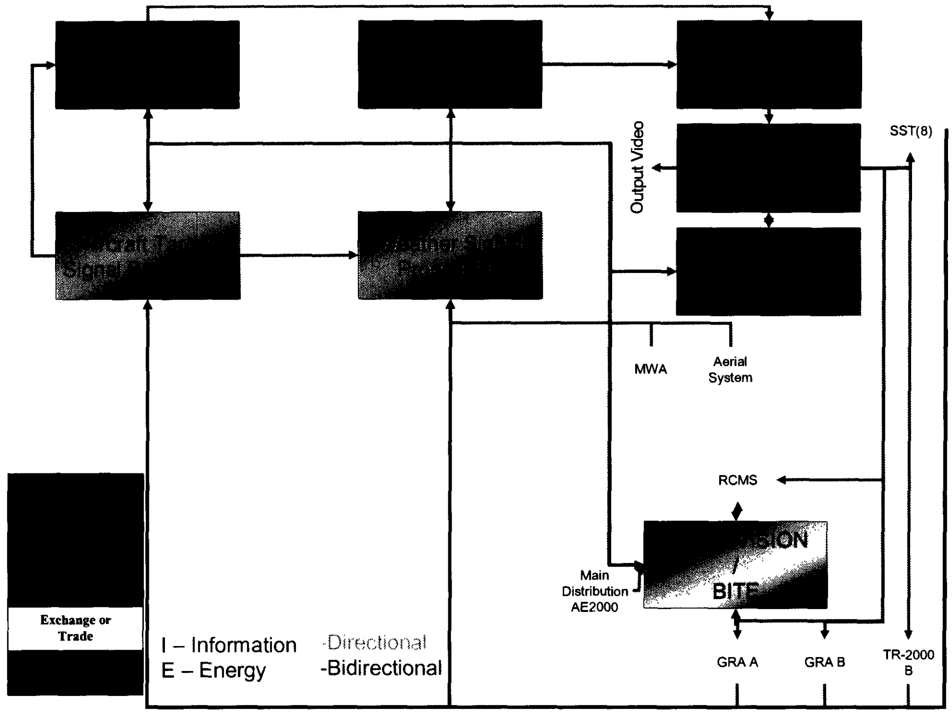


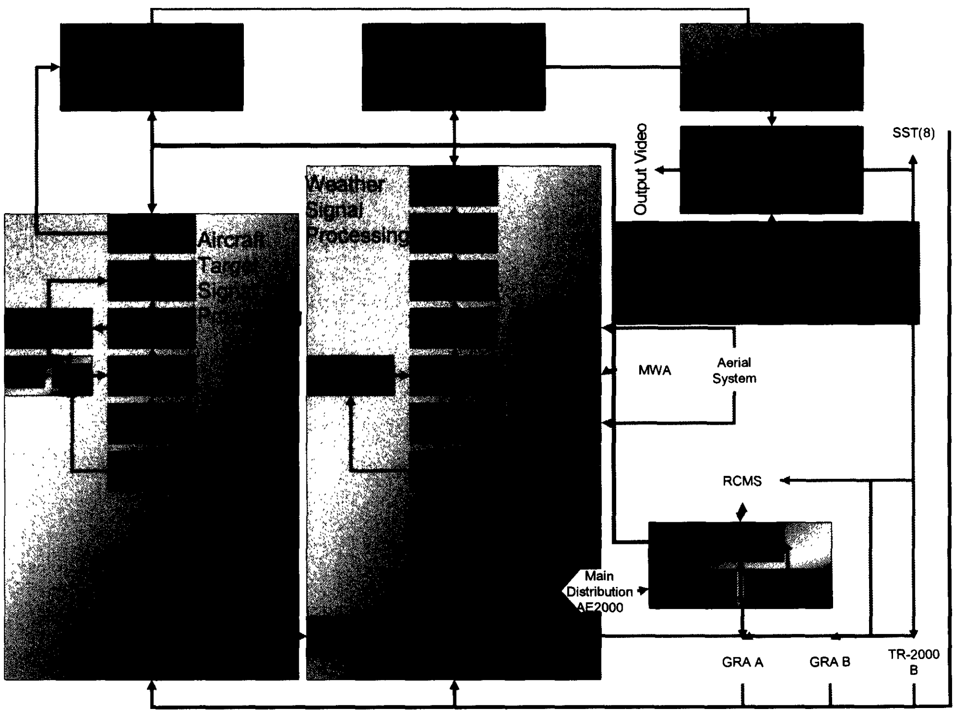
Appendix B: Complexity Diagrams of the MWA

The Reference Decomposition is:

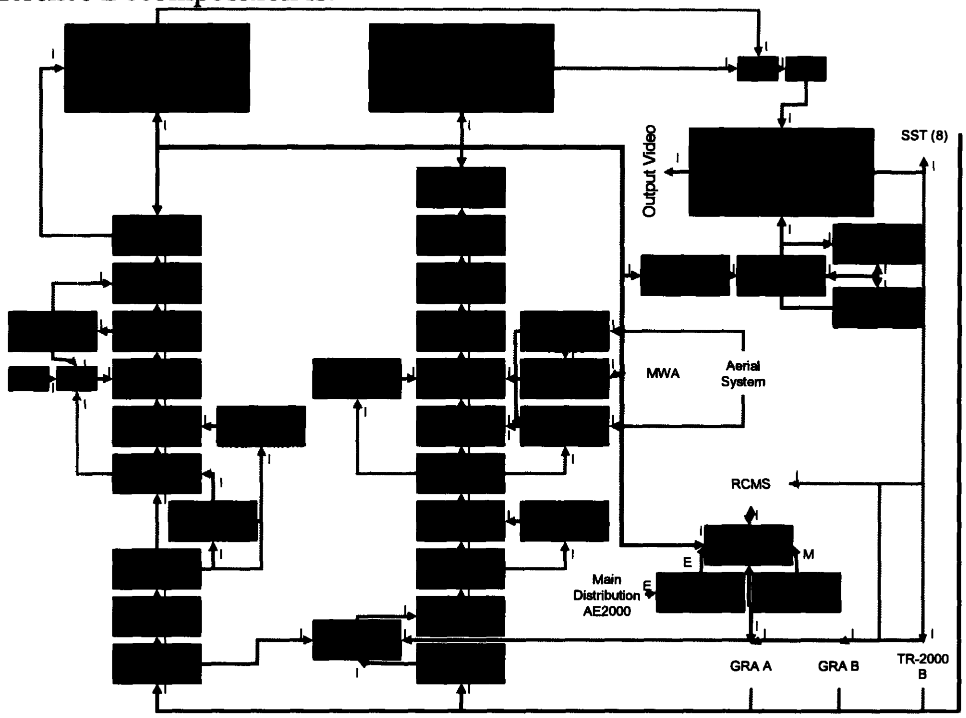
The diagram used to compute MWA Complexity is:

Appendix C: Complexity Diagrams of the TR-2000 A

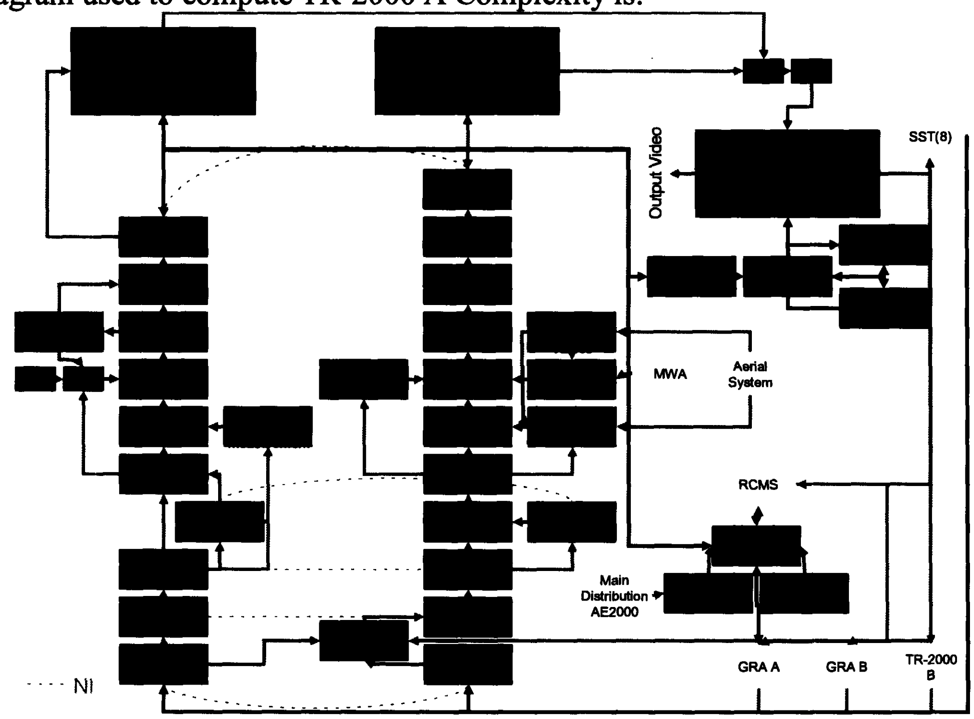




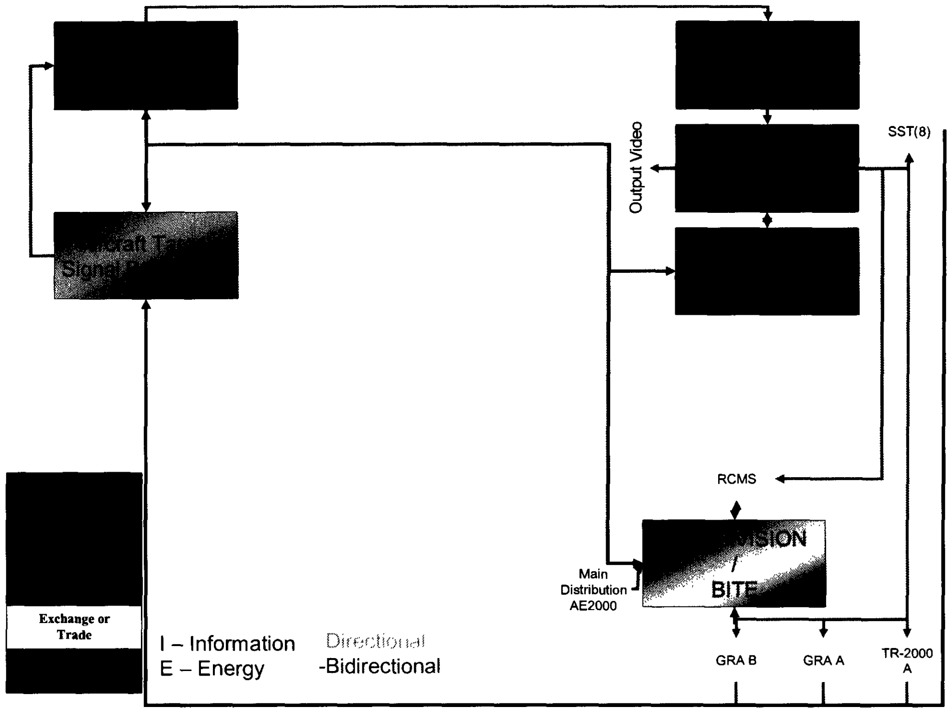
The Reference Decomposition is:

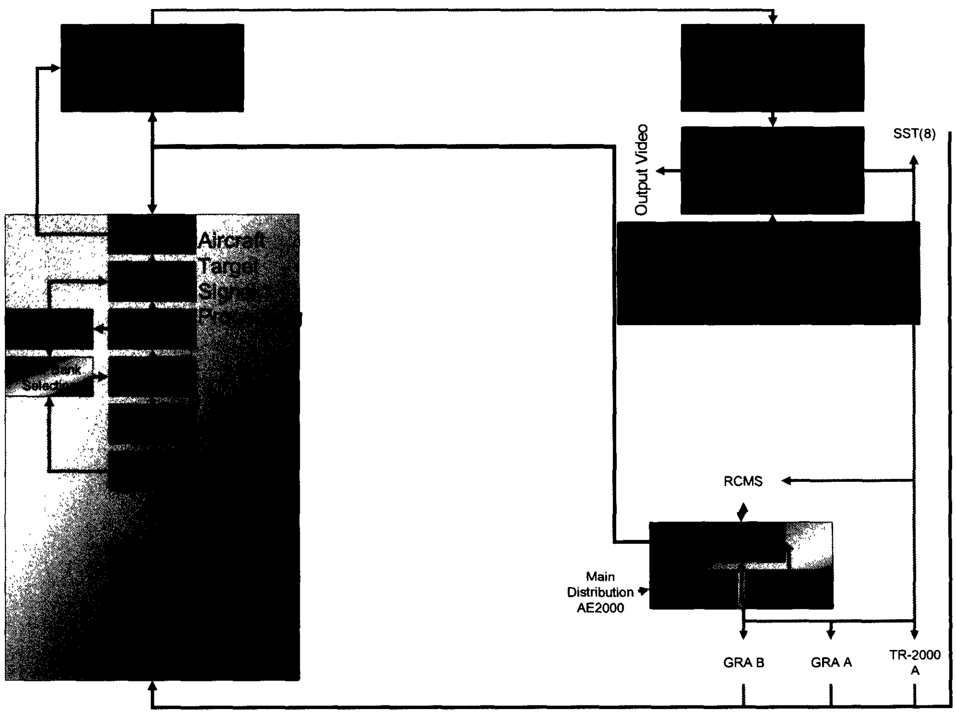


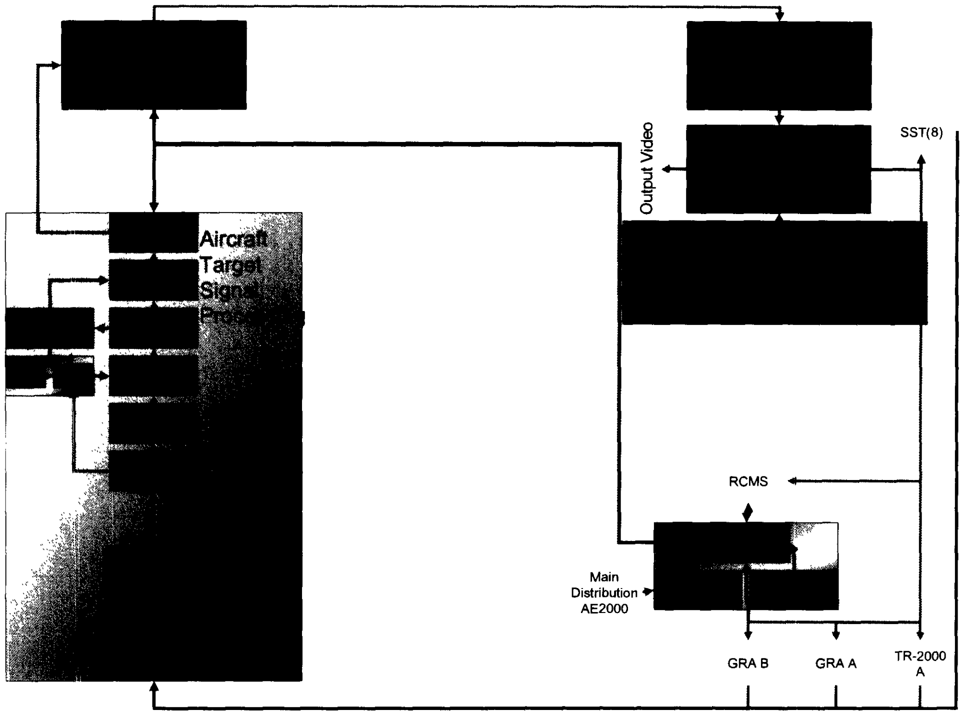
The diagram used to compute TR-2000 A Complexity is:



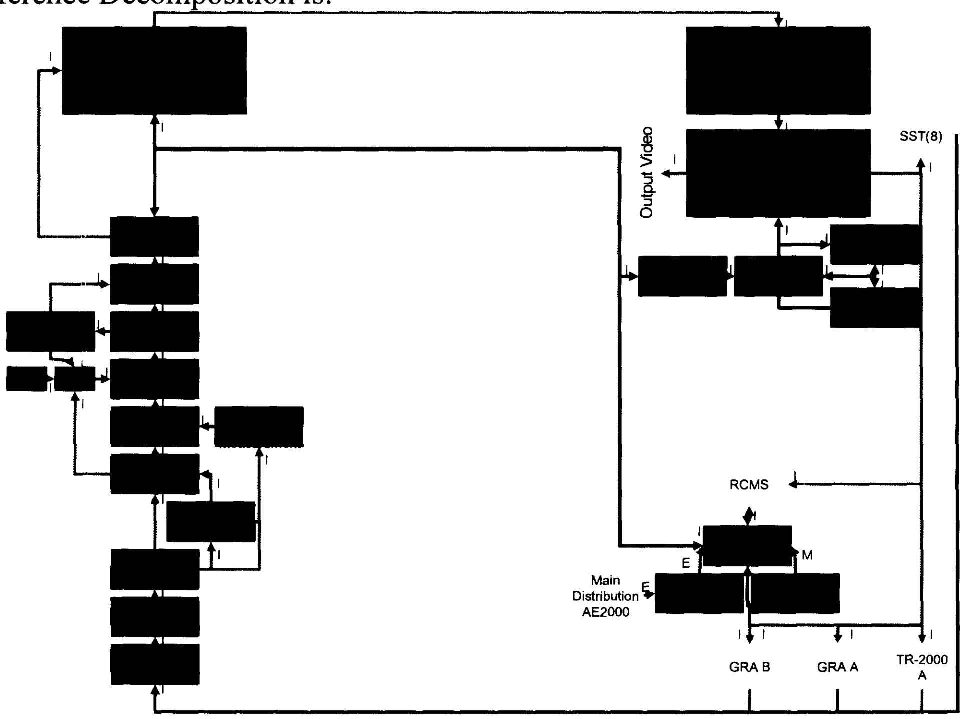
Appendix D: Complexity Diagrams of the TR-2000 B







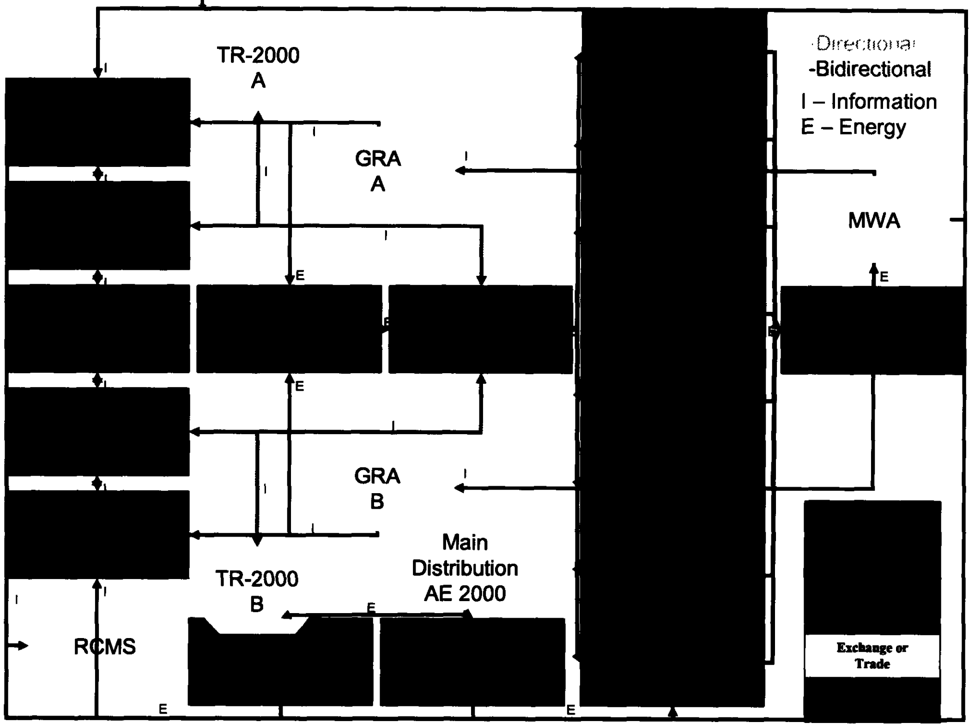
The Reference Decomposition is:



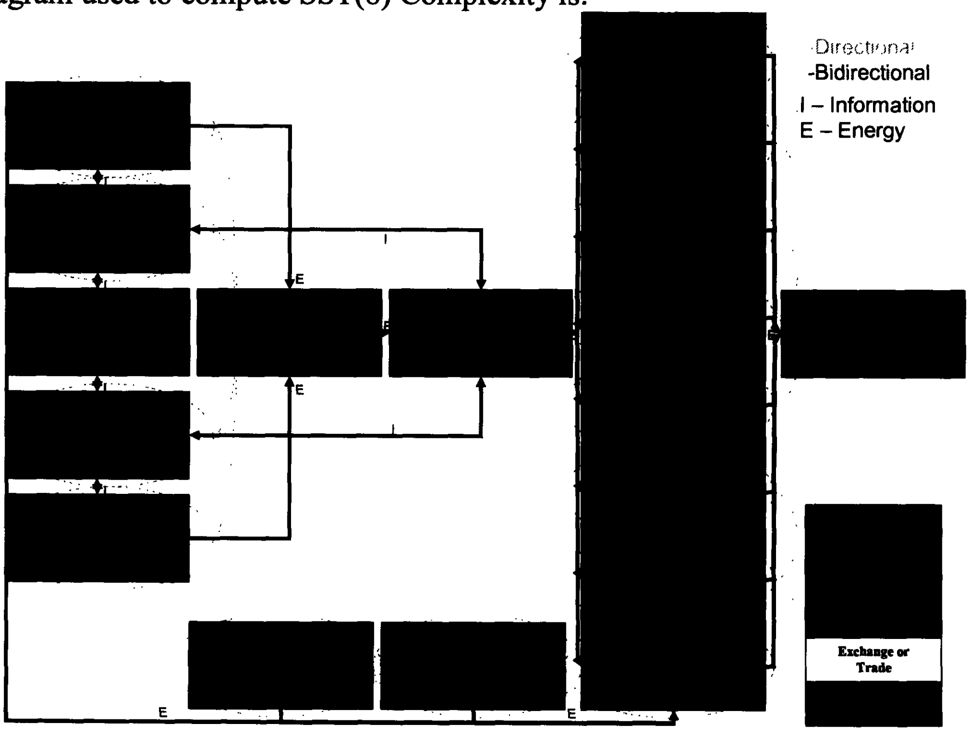
The diagram used to compute TR-2000 B Complexity is:

Appendix E: Complexity Diagrams of the SST(8)

The Reference Decomposition is:



The diagram used to compute SST(8) Complexity is:



Appendix F: Data for the Internal Complexity of the STAR 2000 subsystems

ALL THE ELEMENTS / NON REDUNDANT / NON IDENTICAL

Aerial System

E	L	V	H
14	29	4	3

Level	E_i	V_i	H_i	Level	E_i^{NR}	V_i	H_i	Level	E_i^{NI}	V_i	H_i
2	1	1	2	2	1	1	2	2	1	1	2
3	13	4	3	3	10	4	3	3	10	4	3
4	0	0	0	4	0	0	0	4	0	0	0
5	0	0	0	5	0	0	0	5	0	0	0

GRA

E	L	V	H
34	40	4	5

Level	E_i	V_i	H_i	Level	E_i^{NR}	V_i	H_i	Level	E_i^{NI}	V_i	H_i
2	1	1	2	2	1	1	2	2	1	1	2
3	12	2	3	3	12	2	3	3	8	2	3
4	11	3	4	4	11	3	4	4	11	3	4
5	10	3	5	5	6	3	5	5	6	3	5

MWA

E	L	V	H
28	32	2	2

Level	E_i	V_i	H_i	Level	E_i^{NR}	V_i	H_i	Level	E_i^{NI}	V_i	H_i
2	28	2	2	2	21	2	2	2	15	2	2
3	0	0	0	3	0	0	0	3	0	0	0
4	0	0	0	4	0	0	0	4	0	0	0
5	0	0	0	5	0	0	0	5	0	0	0

TR-2000 B

E	L	V	H
24	35	4	4

Level	E_i	V_i	H_i	Level	E_i^{NR}	V_i	H_i	Level	E_i^{NI}	V_i	H_i
2	3	2	2	2	3	2	2	2	3	2	2
3	19	4	3	3	19	4	3	3	19	4	3
4	2	2	4	4	2	2	4	4	2	2	4
5	0	0	0	5	0	0	0	5	0	0	0

TR-2000 A

E	L	V	H
43	63	4	4

Level	E_i	V_i	H_i	Level	E_i^{NR}	V_i	H_i	Level	E_i^{NI}	V_i	H_i
2	3	2	2	2	3	2	2	2	3	2	2
3	36	4	3	3	36	4	3	3	23	4	3
4	4	3	4	4	4	3	4	4	4	3	4
5	0	0	0	5	0	0	0	5	0	0	0

SST(8)

E	L	V	H
18	52	3	2

Level	E_i	V_i	H_i	Level	E_i^{NR}	V_i	H_i	Level	E_i^{NI}	V_i	H_i
2	18	3	2	2	8	3	2	2	8	3	2
3	0	0	0	3	0	0	0	3	0	0	0
4	0	0	0	4	0	0	0	4	0	0	0
5	0	0	0	5	0	0	0	5	0	0	0

SST(16)

E	L	V	H
26	84	3	2

Level	E_i	V_i	H_i	Level	E_i^{NR}	V_i	H_i	Level	E_i^{NI}	V_i	H_i
2	26	3	2	2	8	3	2	2	8	3	2
3	0	0	0	3	0	0	0	3	0	0	0
4	0	0	0	4	0	0	0	4	0	0	0
5	0	0	0	5	0	0	0	5	0	0	0

Appendix G: Recapitulation of Internal Complexity components and metrics for the STAR 2000 subsystems

	Link Complexity				Scale Complexity - MACRO									Scale Complexity - micro								
	(C_L)				(C_s)	$C=(C_s^2+C_L^2)^{1/2}$				$C=C_s \times C_L$				(C_s)	$C=(C_s^2+C_L^2)^{1/2}$				$C=C_s \times C_L$			
	1	2	Cty	Cty'	C_s	1	2	Cty	Cty'	1	2	Cty	Cty'	C_s	1	2	Cty	Cty'	1	2	Cty	Cty'
Aerial	0.159	2.07	74	10.6	30.3	30.3	30.3	79.9	32	4.82	62.7	2239	320	29.4	29.4	29.5	79.6	31.2	4.68	60.9	2174	311
TR-2000 A	0.035	1.47	187	4.25	92.9	92.9	92.9	209	93	3.24	136	17376	395	85.4	85.4	85.4	206	85.5	2.98	125	15970	363
GRA	0.036	1.18	139	4.09	83.4	83.4	83.4	162	83.5	2.97	98.1	11595	341	69.5	69.5	69.5	155	69.6	2.48	81.8	9660	284
SST(8)	0.17	2.89	192	9.6	32.3	32.3	32.4	195	33.7	5.48	93.2	6195	310	32.3	32.3	32.4	195	33.7	5.48	93.2	6195	310
SST(16)	0.129	3.23	352	12.6	46.6	46.6	46.7	355	48.3	6.02	151	16405	586	46.6	46.6	46.7	355	48.3	6.02	151	16405	586
MWA	0.042	1.14	55	1.96	44.4	44.4	44.4	70.7	44.4	1.88	50.7	2441	87.2	44.4	44.4	44.4	70.7	44.4	1.88	50.7	2441	87.2
TR-2000 B	0.063	1.46	103	4.29	55.7	55.7	55.7	117	55.9	3.53	81.3	5740	239	49.7	49.7	49.7	114	49.9	3.15	72.5	5121	213
Aerial NR	0.159	2.07	74	10.6	23.8	23.8	23.9	77.7	26	3.79	49.2	1759	251	22.9	22.9	23	77.5	25.2	3.65	47.4	1695	242
TR-A NR	0.035	1.47	187	4.25	92.9	92.9	92.9	209	93	3.24	136	17376	395	85.4	85.4	85.4	206	85.5	2.98	125	15970	363
GRA NR	0.036	1.18	139	4.09	73.6	73.6	73.6	157	73.7	2.62	86.6	10231	301	60.3	60.3	60.3	152	60.5	2.15	71	8386	247
SST(8) NR	0.17	2.89	192	9.6	14.3	14.3	14.6	193	17.3	2.44	41.4	2753	138	14.3	14.3	14.6	193	17.3	2.44	41.4	2753	138
SST(16) NR	0.129	3.23	352	12.6	14.3	14.3	14.7	352	19.1	1.85	46.3	5048	180	14.3	14.3	14.7	352	19.1	1.85	46.3	5048	180
MWA NR	0.042	1.14	55	1.96	33.3	33.3	33.3	64.3	33.3	1.41	38	1831	65.4	33.3	33.3	33.3	64.3	33.3	1.41	38	1831	65.4
TR-B NR	0.063	1.46	103	4.29	55.7	55.7	55.7	117	55.9	3.53	81.3	5740	239	49.7	49.7	49.7	114	49.9	3.15	72.5	5121	213
Aerial NI	0.159	2.07	74	10.6	23.8	23.8	23.9	77.7	26	3.79	49.2	1759	251	22.9	22.9	23	77.5	25.2	3.65	47.4	1695	242
TR-A NI	0.035	1.47	187	4.25	64.8	64.8	64.8	198	65	2.26	95	12123	276	59.4	59.4	59.4	196	59.6	2.07	87	11108	252
GRA NI	0.036	1.18	139	4.09	63.8	63.8	63.8	153	63.9	2.27	75	8867	261	53.2	53.2	53.2	149	53.3	1.9	62.5	7389	217
SST(8) NI	0.17	2.89	192	9.6	14.3	14.3	14.6	193	17.3	2.44	41.4	2753	138	14.3	14.3	14.6	193	17.3	2.44	41.4	2753	138
SST(16) NI	0.129	3.23	352	12.6	14.3	14.3	14.7	352	19.1	1.85	46.3	5048	180	14.3	14.3	14.7	352	19.1	1.85	46.3	5048	180
MWA NI	0.042	1.14	55	1.96	23.8	23.8	23.8	59.9	23.9	1.01	27.2	1308	46.7	23.8	23.8	23.8	59.9	23.9	1.01	27.2	1308	46.7
TR-B NI	0.063	1.46	103	4.29	55.7	55.7	55.7	117	55.9	3.53	81.3	5740	239	49.7	49.7	49.7	114	49.9	3.15	72.5	5121	213

Appendix H: Typical values of C_{EXT}

For C'_{EXT}

Example 1: Constant risk

$p_i = \frac{k}{i}$, $i \in \aleph^*$

$$m = \frac{\sum_{i=1}^{N} ip_i}{\sum_{i=1}^{N} p_i} = \frac{\sum_{i=1}^{N} k}{\sum_{i=1}^{N} \frac{k}{i}} \underset{+\infty}{\approx} \frac{kN}{k(\ln(N)+\gamma+o(\frac{1}{N}))} \underset{+\infty}{\approx} \frac{N}{\ln(N)+\gamma+o(\frac{1}{N})}$$

$$C'_{EXT} = \frac{\sum_{i=[m]}^{N} ip_i}{\sum_{i=1}^{[m]} ip_i} = \frac{\sum_{i=[m]}^{N} \frac{k}{i}}{\sum_{i=1}^{[m]} \frac{k}{i}} = \frac{\sum_{i=1}^{N} \frac{k}{i} - \sum_{i=1}^{[m]-1} \frac{k}{i}}{\sum_{i=1}^{[m]} \frac{k}{i}} \underset{+\infty}{\approx} \frac{\ln(N)+\gamma+o(\frac{1}{N}) - (\ln([m]-1)+\gamma+o(\frac{1}{[m]-1}))}{\ln([m])+\gamma+o(\frac{1}{[m]})}$$

$$\underset{+\infty}{\approx} \frac{\ln(\frac{N}{[m]-1})+o(\frac{1}{[m]})}{\ln([m])+\gamma+o(\frac{1}{[m]})}$$

Now, $[m] \underset{+\infty}{\approx} \left[\frac{N}{\ln(N)+\gamma+o(\frac{1}{N})} \right] \underset{+\infty}{\approx} \frac{N}{\ln(N)}$ and $\frac{N}{[m]-1} \underset{+\infty}{\approx} \ln(N)$

So, $C'_{EXT} \underset{N\to+\infty}{\to} 0$

Example 2: Linearly growing risk

The risk linearly increases with magnitude: the annual frequency of events is thus constant for every magnitude i.

$p_i = k$

$$m = \frac{\sum_{i=1}^{N} ip_i}{\sum_{i=1}^{N} p_i} = \frac{\sum_{i=1}^{N} ki}{\sum_{i=1}^{N} k} = \frac{k\left(\frac{N(N+1)}{2}\right)}{kN} = \frac{N+1}{2}$$

$$C'_{EXT} = \frac{\sum_{i=[m]}^{N} ip_i}{\sum_{i=1}^{[m]} ip_i} = \frac{\sum_{i=[m]}^{N} ki}{\sum_{i=1}^{[m]} ki} = \frac{\sum_{i=1}^{N} ki - \sum_{i=1}^{[m]-1} ki}{\sum_{i=1}^{[m]} ki}$$

$$= \frac{k\left(\frac{N(N+1)}{2}\right) - k\left(\frac{[m]([m]-1)}{2}\right)}{k\left(\frac{[m]([m]+1)}{2}\right)} = \frac{N(N+1) - [m]([m]-1)}{[m]([m]+1)}$$

Now, $[m] \underset{+\infty}{\approx} \frac{N}{2}$

$$C^1_{EXT} \underset{+\infty}{\approx} \frac{N(N+1) - \frac{N}{2}(\frac{N}{2} - 1)}{\frac{N}{2}(\frac{N}{2} + 1)} \underset{+\infty}{\approx} \frac{\frac{3N^2}{4}}{\frac{N^2}{4}} \underset{+\infty}{\approx} 3$$

So, $C_{EXT} \underset{N \to +\infty}{\to} 3$

For C^2_{EXT}

<u>Example 1: Constant risk</u>

$p_i = \frac{k}{i}$, $i \in \aleph^*$

$$C^2_{EXT} = \frac{\sum_{i=[0.9N]}^{N} ip_i}{\sum_{i=1}^{[0.1N]} ip_i} = \frac{\sum_{i=[0.9N]}^{N} k}{\sum_{i=1}^{[0.1N]} k} = \frac{k(N - [0.9N] + 1)}{k[0.1N]} = \frac{N - [0.9N] + 1}{[0.1N]}$$

So, $C^2_{EXT} \underset{N \to +\infty}{\to} 1$

<u>Example 2: Linearly growing risk</u>

$p_i = k$

$$C^2_{EXT} = \frac{\sum_{i=[0.9N]}^{N} ip_i}{\sum_{i=1}^{[0.1N]} ip_i} = \frac{\sum_{i=1}^{N} ki - \sum_{i=1}^{[0.9N]-1} ki}{\sum_{i=1}^{[0.1N]} ki}$$

$$= \frac{k\left(\frac{N(N+1)}{2}\right) - k\left(\frac{[0.9N]([0.9N]-1)}{2}\right)}{k\left(\frac{[0.1N]([0.1N]+1)}{2}\right)} = \frac{N(N+1) - ([0.9N]([0.9N]-1))}{[0.1N]([0.1N]+1)}$$

$$\underset{+\infty}{\approx} \frac{N(N+1) - (0.9N(0.9N-1))}{0.1N(0.1N+1)} \underset{+\infty}{\approx} \frac{(1-0.9^2)N^2}{0.1^2 N^2}$$

So, $C^2_{EXT} \underset{N \to +\infty}{\to} 19$

Appendix I: Data to draw Figure 20 – France: characterization of the complexity of transportation systems

Air

Magnitude	Frequency
1	2
2	1
4	1
5	1
6	1
10	1
14	1
20	1
113	1

Source: [23]

Land

Magnitude	Frequency
1	70205
2	5902
3	1018
4	287
5	80
6	16
7	4
8	3
10	2
12	2
15	1
22	1
28	1

Source: [24]

Water

Magnitude	Frequency
1	4
4	1
8	1

Source: [25]

Appendix J: Data to draw Figure 21 – United States: characterization of the complexity of transportation systems

Air

Magnitude	Frequency
1	7
2	3
4	1
7	2
9	1
10	1
12	1
19	1
20	2
21	1
25	2
27	1
37	2
38	1
56	1
58	1
64	2
70	1
81	1
83	1
105	1
110	1
126	1
127	1
137	2
148	1
152	1
212	1
243	1
248	1
251	1

Source:[26]

Land

Magnitude	Frequency
1	963322
2	82379
3	14646
4	3876
5	1104
6	360
7	150
8	57
9	14
10	12
11	9
12	2
13	2
14	2
20	1
21	2
22	1

Source: [27]

Water

Magnitude	Frequency
1	6
2	5
3	1
4	1
5	1
6	1
15	1

Source: [25]

Appendix K: Data to draw Figure 22 – World: characterization of the complexity of transportation systems

Magnitude	Frequency
1	143
2	85
3	45
4	31
5	41
6	21
7	27
8	32
9	19
10	28
11	8
12	7
13	7
14	13
15	5
16	13
17	6
18	6
19	7
20	3
21	6
22	3
23	6
24	8
25	1
26	2
27	4
28	2
29	1
30	1
31	1
33	2
34	2
35	3
36	1
37	2
38	1
40	1
41	1
42	2
43	1
44	3
45	2
46	1
49	2
50	2
51	2
53	2
54	1
55	2
57	2

Magnitude	Frequency
60	1
61	1
63	1
64	1
65	3
66	1
68	1
69	1
70	3
71	1
74	2
75	4
78	1
80	1
83	1
85	1
88	1
92	1
93	1
96	1
98	1
101	1
102	1
104	2
109	1
110	2
112	1
116	1
117	1
123	1
125	2
129	1
131	1
132	1
140	2
141	1
143	2
145	1
148	1
160	2
169	1
197	1
217	1
225	1
228	1
229	1
230	1
234	1
260	1
264	1
312	1

Magnitude	Frequency
1	205
2	74

Water >

Source:[25]

< Air

Source:[28]

Magnitude	Frequency
3	50
4	42
5	28
6	27
7	15
8	17
9	10
10	13
11	4
12	7
13	6
14	8
15	2
16	4
17	5
18	5
19	3
20	4
21	3
22	3
23	2
24	5
25	4
26	6
27	6
28	3
29	4
30	2
31	1
32	1
33	2
35	1
36	2
37	1
38	1
40	1
41	1
43	1
46	1
51	1
55	1
63	1
83	1
121	1
141	1
145	1
150	1
216	1
464	1
543	1
852	1

Room 14-0551
77 Massachusetts Avenue
Cambridge, MA 02139
Ph: 617.253.5668 Fax: 617.253.1690
Email: docs@mit.edu
http://libraries.mit.edu/docs