Influence Strategies for Systems of Systems

by

Nirav Bharat Shah

B.S. in Aeronautics and Astronautics – Massachusetts Institute of Technology, 2001
M.S. in Aeronautics and Astronautics – Massachusetts Institute of Technology, 2004

Submitted to the Department of Aeronautics and Astronautics
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Aeronautics and Astronautics
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
February 2013

© Massachusetts Institute of Technology 2013. All rights reserved.

Author . . . .
Department of Aeronautics and Astronautics
October 26, 2012

Certified by . . . .
Joseph M. Sussman
JR East Professor of Civil and Environmental Engineering and Engineering Systems
Thesis Supervisor

Certified by . . . .
Daniel E. Hastings
Cecil and Ida Green Education Professor of Aeronautics and Astronautics and
Engineering Systems
Thesis Supervisor

Certified by . . . .
R. John Hansman, Jr.
Professor of Aeronautics and Astronautics and Engineering Systems
Thesis Committee Member

Certified by . . . .
Donna H. Rhodes
Senior Lecturer, Engineering Systems
Thesis Committee Member

Accepted by . . . .
Eytan Modiano
Professor of Aeronautics and Astronautics
Chair, Committee on Graduate Students

Influence Strategies for Systems of Systems

by

Nirav Bharat Shah

Submitted to the Department of Aeronautics and Astronautics
on October 26, 2012, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Aeronautics and Astronautics

Abstract

Distributed decision making has been identified as a source of managerial complexity for leaders of systems of systems (SoS). A new framework, AIR (Anticipation-Influence-Reaction), is proposed to capture the feedback relationships between the decisions made by constituents and those made by the managers of the SoS. AIR is then used to develop a five-member set of basic influences that can bring about changes in constituent behavior thus modifying the SoS. These influences, the 5 Is, are Incentives, Information, Infrastructure, Integration, and Institutions. AIR and the influences are demonstrated through qualitative application to real-world SoSs and quantitatively through simulation of an inter-modal freight transport network. It is found that cooperation between competing constituents, i.e., rail and truck carriers, can be quite fragile and sensitive to the SoS context. Careful, dynamic planning of influence strategies is needed to maintain SoS behavior in the face of constituents who are driven by self-interest and a limited, local perspective of the SoS.

Keywords

Systems of systems; Influence; Management; Complexity; freight; Simulation; Transportation; Inter-modal; Game Theory; Distributed decision making; Anticipation; Reaction

Thesis Supervisor: Joseph M. Sussman
Title: JR East Professor of Civil and Environmental Engineering and Engineering Systems

Thesis Supervisor: Daniel E. Hastings
Title: Cecil and Ida Green Education Professor of Aeronautics and Astronautics and Engineering Systems

Acknowledgments

Many people have supported me in completing my dissertation. First and foremost I would like to thank my thesis committee. **Prof. Sussman**, thank you for your guidance, candor and willingness to put up with my flights of intellectual fancy. Your unwavering support kept me going through all those challenging years. **Prof. Hastings**, thank you for allowing me to find my own question and for your support when times got tough. I could not have overcome those obstacles without it. **Dr. Rhodes**, it was you who first introduced me to systems of systems. Thank you for sharing your unique perspective, and for giving me a research home at SEAri. **Prof. Hansman**, thank you for teaching me that the purpose of research is to bring forth new understanding of the world around us. As you once put it, "We make models not to understand our models better, but to understand the real world better."

All my colleagues at **SEAri** and **SSPARC**, thank you for providing my intellectual home. To all my fellow students who started as colleagues and became friends, thank you. **Adam**, we've had long research discussions, but they are always productive and I always walk away learning something new. **Matthew**, you have always been my greatest collaborator. Thank you for inspiring me to expand my horizons and explore new intellectual areas. **David** and **Chris L.**, thank you for teaching me game theory that came in quite handy in my case study. **Chris R.**, thank you for the jokes, research help and spiritual conversations. **Jennifer**, thank you for all your support—especially for teaching me when to stop one task and move on to the next. Thanks **Zoe** for reminding me that sometimes 'soft' sciences are hardest. Thanks **Greg** for your always insightful comments on my research and warm friendship. I was honored to defend on the same day as you. Thanks **John** for sharing your stories and making the lab a wonderful place to work. **Jason**, thank you for all your advice, personally and professionally. You always told me the truth I needed to hear. When it came time to bring my thesis to a close, your push gave me the momentum that carried me across the finish line.

Right from the start, the **Aero/Astro** department staff have helped me navigate the unexpected twists and turns during my time at MIT—thank you **Barbara**, **Marie** and **Beth**. Thank you **Prof. Nightingale**, **Dr. Bozdogan**, **JK** and all the students and staff at **LAI**. In completing my masters with you, I learned the real value of systems thinking. **Dean Staton**, thank you for your support and connecting me with the right people to move my work forward. **Prof. Frey**, though our conversations have become less frequent, you always had a knack to putting ideas into a larger context and making connections between disciplines. **Prof. Iyengar**, you taught to me that no obstacle, no matter how profound, can stop you in your search for truth. **Dr. Dolin** and **Dr. Treml**, thank you for teaching me that you can follow the rules and break them at the same time. **Prof. Hickey**, you introduced me to the wonder of research and solving a problem that has never been solved before. **Mr. Sachs**, thank you for planting the research seed in my heart.

Very special thanks to all my friends and students at **ESG** and **Concourse**. It was with you that I first learned to learn and later learned to teach. From indoor snowmen to Lego robots and steam engines, with you my passion for engineering and education blossomed. Thank you to all my other teachers and students.

I am always inspired by my spiritual family both here in Boston at **JCGB** and around the world via **YJA** and **YJUK**. Our faith gave me strength when challenges arose. **Pavak, Chintan, Shilpa, Dipti, Abhishek, Jyoti, Finale, Naman** and **Yogendra bhai**, all of you gave me advice when I needed it, made me laugh when I needed that and showed me how to live out the true virtues of our faith.

Not one day went by when my friends were not ready and willing to help. Thank you **Cappy**, it was your advice many years ago that started me on the journey to being an aerospace engineer—for that and for everything else I will always be grateful. Thank you **Phil** for helping me through the toughest times. When I didn't know who to turn to, I could count on you. Thank you **Neepa** for always being willing to listen and for sharing your love of food, dancing and stories. Thank you **Anna** for checking in with me at just the right times, saying just the right things and giving me confidence that I too would finish. Thank you **Shardule** and **Ami** for giving me hope that things will get better. Thank you **Priyanka** for your constant encouragement, laughter filled yet fascinating conversations and reminding me to never stop running even when my goal seems out of reach. Thank you **Kacy** and **Mike** for being my colleagues, friends or family as the situation required.

Thank you **Joyce**. Thank you for inspiring me to make systems my career. Thank you for giving me the chance to do research that spanned boundaries and pushed limits. You left us too soon, but your memory lives on in the hearts of all the students whose lives you changed.

Thank you to my family. I could have never done this without you. To all my aunts, uncles and cousins who never lost hope, thank you. To **Manan bhai**, you've always been like a big brother to me. Thank you for your wisdom.

Finally, thank you **mom** and **dad**. I owe this and everything else I have to you.

For mom and dad

In memory of Joyce Warmkessel, my first advisor, who always encouraged me to go beyond the answer and find the next question.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Since the mid 90's there has been a growing interest in how systems come together to form systems of systems (SoS). These coalitions of independently operated and independently managed systems can meet unforeseen needs in a timely and cost effective fashion. Traditional systems engineering theories and approaches do not fully address the technical and managerial challenges of SoS. This thesis focuses on developing better strategies for coping with the managerial complexity caused by the dynamic interactions between constituent systems within a SoS. By understanding these interactions, systems engineers and managers will be better able to improve engineering and management strategies to influence an SoS.

1.1 Motivation

Current interest in these types of systems can be traced back to the early 1990's with the work of Eisner et al. (1991); Eisner (1994), and, later, Maier (1999, orig. published in 1996). In the latter paper, Maier defines two independence properties characteristic of SoS that have subsequently been used by many authors to define the class of systems termed SoS (a review can be found in Keating et al., 2003). These two properties, *operational independence* and *managerial independence*[1], specify that, from both a technical and a social perspective, an SoS is composed of independent yet interacting entities. This formulation has been extended and refined over time. More recently, Karcanias and Hessami (2011a) echoed Maier's claim stating:

> "A System of Systems is a "super system" comprised of other elements which themselves are independent complex operational systems and interact among themselves to achieve a common goal…"[2]

[1]See section 2.1 for a full discussion of the defining characteristics of SoS.

[2]Karcanias and Hessami's notion that SoS elements have a common, shared goal is not universal among definitions.

> "The distinguishing feature of the SoS case is that the subsystems participate in the composition as intelligent agents with a relative autonomy..." (Karcanias and Hessami, 2011a)

Karcanias and Hessami further state that:

> "The multi-agent dimension of SoS has characteristics such as [20]:
>
> **Autonomy:** the agents are at least partially autonomous
>
> **Local Views:** no agent has a full global view of the system, or the system is too complex for an agent to make practical use of such knowledge
>
> **Decentralisation (sic.):** there is no designated single controlling agent, but decision and information gathering is distributed." (Karcanias and Hessami, 2011a)

This third characteristic, distribution of decision making, is a core challenge within SoS engineering and is the focus of this thesis. Traditional systems engineering relies upon centralized coordination of decision making to ensure that decisions made at lower levels of abstraction are concordant with and produce the desired behavior at higher levels, e.g., subsystem objectives are chosen based upon system objectives. Indeed it is this ability to coordinate that underpins the V-model of systems engineering (Forsberg et al., 2005). This distribution of decision making causes several additional challenges for the systems engineer who seeks to gain value from the operation of the SoS. For example, those who manage SoS components may not share the central goal of the SoS. This was seen in the U. S. Army Task Force XXI exercise (Krygiel, 1999). The exercise attempted to integrate several independent army fighting unit into a combined force in a rapid, efficient manner. However, the participating units did not have an incentive to implement the needed standards and protocols. They had their own program level objectives to meet. When the the individual units met to combine, the systems failed to integrate. Only with concentrated effort while the systems were collocated was integration eventually achieved.

Another source of difficulty is the often asynchronous nature of decision making in an SoS. Different actors makes decisions at different times and consequently with different sets of information guiding them. In networks, for example, this can cause localized congestion even when sufficient capacity exists globally. This is observed in Internet traffic, road networks and even interbank settlement (Beyeler et al., 2006). Characterizing such a situation requires an understanding of the temporal dynamics within the SoS. Resolving it may require changes to both the systems that comprise the SoS (i.e. changes intrinsic to the SoS) as well as changes in the 'offline' interaction between those managing the constituent systems (i.e. extrinsic of the SoS).

1.2 SoS terminology

As with any emerging field, definitions and terminology can be quite fluid. While no SoS terminology has been universally adopted, the following definitions are common and are used in this thesis:

System "A combination of interacting elements organized to achieve one or more stated purposes." (INCOSE, 2006)

Subsystem A subset of the elements of a system identified as such to enable hierarchic description of a system and its elements.

System of systems "A set or arrangement of systems that results when independent and useful systems are integrated into a larger system that delivers unique capabilities. Both individual systems and SoS conform to the accepted definition of a system in that each consists of parts, relationships, and a whole that is greater than the sum of the parts; however, although an SoS is a system, not all systems are SoS." (DoD, 2008)

Constituent system An operationally and managerially independent system that is part of an SoS. (Maier, 1999; Krygiel, 1999)

Component system A system that is a subsystem of an SoS but lacks managerial independence that would make is a constituent system.

Interaction A connection between entities (constituent systems, component systems and other subsystems) in an SoS via which matter, energy, information or 'value' is transferred. (after Magee and de Weck (2004)).

Interface A purposefully created interaction. An SoS is created when several constituent and component[3] systems interact via interfaces.[4]

Context or system environment "The environment of a system is a set of elements and their relevant properties...[The] elements are not part of the system but a change in any of which can produce a change in the state of the system. Thus a system's environment consists of all variables [that] can affect its state. External elements which effect irrelevant properties of a system are not part of its environment." (Ackoff, 1971) With respect to SoS, see Shah et al. (2007a).

Decision maker One who makes decisions with respect the design and operation of a system. This role may be split among multiple parties as a system goes through its life cycle.

[3]If there are no constituent systems involved then the SoS engineering problem is essentially the same as in traditional systems engineering and the SoS challenges examined in this thesis don't occur.

[4]The case of virtual SoS is slightly different. It is possible for a virtual SoS to exist purely through unintended interactions between constituents.

SoS influencer A decision-making entity that has a preference on the structure and/or behavior of an SoS. It influences the constituents decision makers to alter their systems so as to produce the desired SoS.

The Internet is a good example for seeing how these terms are used in practice. Detailed definitions and discussion are provided in chapter 2. A simplified representation of the Internet is shown in Figure 1-1.

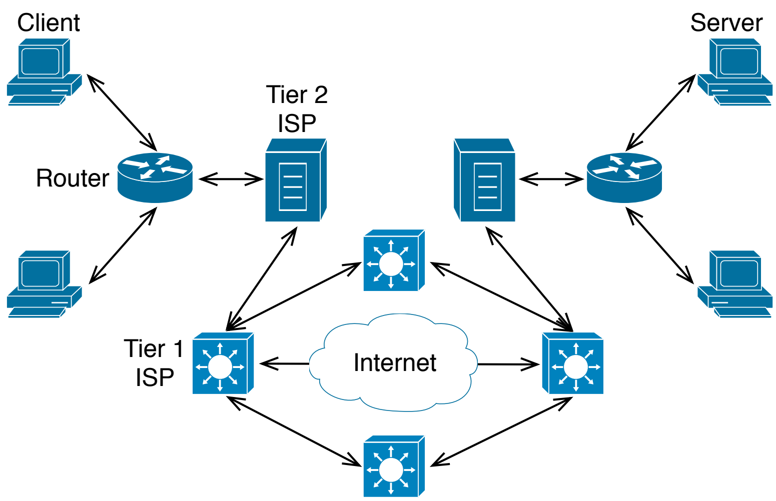


FIGURE 1-1: A simplified representation of the Internet

In this figure, traffic flow is represented as going between a server to a client through a series of intermediaries. Both the server and client part of local networks. These networks include a interface point to the outside world, i.e., a router. The router keeps traffic that is local within the local network and directs external traffic to/from computers on the local network. The router gains access to external routes via transport provided by a Tier 2 ISP or Internet Service Provider. An Tier 2 ISP is a business that provides connection to the Internet to end-users. The Tier 2 ISP in turn purchases its connection from a Tier 1 ISP, usually on a wholesale basis. The Tier 1 ISPs interconnect to form the Internet. They do so using so called 'peering agreements' providing transport for the other's traffic without exchanging fees. The use of two tiers in a hierarchy is only for demonstrations purposes. The real Internet consists of a somewhat more complicated mesh of agreements between ISPs but the essential structure of end-users having to purchase access from an external provider who in turn purchase access from larger wholesale providers is a common pattern.

First, one must remember that the specifying a group of interacting entities as constituting a 'system' is subjective categorization created to bound the set of relevant entities and context for a given problem. Looking at the same group of interacting entities, two different people may draw the system boundary differently depending upon the specific problem they are trying to solve or aspect of the system they are trying to understand. The same is true for specifying subsystems or SoS.

If an analyst was examining the effect of new peering agreements on traffic congestion, a natural classification would be to consider each Tier 1 ISP as a system interconnected with each other to form the SoS of the Internet. In setting up these agreements, the relevant decision makers are the ISPs. As they are independently owned and operated, they would be considered the constituents. While the end-user clients are also independent, the analyst might chose to consider their behavior as internal to the larger Tier 1 ISPs as end-user traffic only becomes visible to the larger network via the ISPs. In this formulation the Internet is the SoS and the tier 1 ISPs constituents. An alternative formulation would be to extend the role of constituent to tier 2 ISPs or even end-users. Both formulations are useful. The former restricts the analysis to a relatively small number of constituents allowing them to represented in greater depth, while the latter recognizes that actual traffic arises from the much larger number of 'small' decisions made by end-users that, in combination, lead to the congestion effect under study. To fully understand the problem, a combination of both approaches would be needed. There is no one universally correct abstraction for a given SoS. Given the complexity of modern systems, the choice of what to consider a constituent is often dictated as much by the problem under consideration as by the SoS.

To ensure clarity, whenever an SoS is described, its constituents will be explicitly defined. In addition, as is needed, the decision makers for each constituent will be defined as well. Generally, the decision maker and the constituent systems can be thought of as a single unit when viewed from the SoS perspective. Thus the term constituent in isolation refers to this unit. When the distinction between the decision maker(s) and the technical systems needs to be made explicit, the more specific terms will be used.

1.3 Research questions

The challenges described above, among others, are common in SoS and require the systems engineer to use new tools and strategies in managing their SoS. As is evident from the examples above, developing such tools requires an understanding of the dynamic, multi-faceted decision making process that drives SoS behavior to a much greater extent than in traditional systems engineering. Furthermore, addressing those issues may require a combination of social and technical efforts. To that end, the following research questions are proposed:

The first research question concerns the establishment of a new framework to characterize the decision making processes of an SoS. The framework should capture both the fact that there are multiple decision makers working at the constituent and SoS levels and the interaction between these decision makers. As a matter of scope, the inquiry is limited to an extant SoS with a fixed set of constituents.

> *What are the feedback relationships between the constituents and SoS influencers, and how do their influences result in changes in the constituent individually and the SoS as a whole?*

The second research question concerns the use of the additional perspective gained by the using the new framework to better produce strategies to influence the behavior of other stakeholders.

> *What approaches can be used by external SoS influencers to cause constituent decision makers to change constituent systems so as to induce a desired behavior from the SoS?*[5]

1.4 Key contributions

In regard to the first question, a descriptive framework, known as Anticipation-Influence-Reaction (AIR), for decision making in SoS is proposed. The commonly used taxonomy of collaborative, virtual, directed and acknowledged SoS (Maier, 1999; Dahmann and Baldwin, 2008) is mapped into the framework. It is then shown that all four classes differ only by how the constituents decision makers interact with the influencer(s).

Building upon this descriptive foundation, a basis set of influence strategy types are proposed by treating constituent decision making as a value maximizing process. These types, known as the five I's, are Incentives, Information, Integration, Infrastructure and Institutions.

'Incentives' is rewarding constituents for particular behavior that they would not do otherwise. 'Information' refers to providing constituents information to change the priors they use to make decisions under uncertainty. 'Integration' is the re-assignment of particular SoS components to different constituents. A common example would be combining two systems into one. 'Infrastructure' refers to introducing new technology into the SoS. An example would be a new high-speed data network to facilitate higher bandwidth inter-connection between constituents. Finally, 'institutions' refer to the rules and regulations that constituents follow.

The use of the AIR framework is demonstrated in a case study of an intermodal freight transportation network. The purpose of the case study is to demonstrate a process that an SoS influencer could use to change the performance of the SoS via changes in constituent behavior. In doing so, first the intermodal freight transportation problem is characterized using the AIR framework and transportation literature. Underutilization of intermodal rail service is identified as a key concern. To gain a better understanding of this SoS problem, an example transportation system that uses both rail and truck routes is modeled using a agent-based simulation that represent shipper and carrier decision making over a 15 year period. The model is then used to examine several intervention strategies based upon the 5 Is. Different strategies can have vastly different impact on the

[5]This is not to imply that all SoS have influencers, rather, the question concerns techniques influencers should use when they are extent.

constituents even though they produce similar behavior in the SoS. This result reinforces the need for SoS influencers to consider the effect of intervention on constituents locally, not just on the SoS as whole.

These contributions fill a significant gap in the SoS Engineering literature. While existing frameworks describing SoS identify the multi-stakeholder, multi-layer decision making structure as a key issue in SoS, they do not provide much proscriptive guidance to the systems engineer as to how to handle such a situation. The roles, interactions and processes described in AIR capture in a succinct form the key structures needed to understand SoS behavior where the constituent set and value proposition (at both the constituent and SoS levels) are fixed. With further research, the framework can be extended to allow for both these constraints to be relaxed.

1.5 Research approach

In addressing these questions, a three stage research approach was used. First, empirical grounding in real-world SoS was developed by reviewing existing case studies of SoS in the literature as well as systems, that while not using the SoS term, exhibited the independence properties Maier (1999) identified. The result was a clear indication that, in addition to technical considerations, organizational considerations were important in the development and management of SoS. The SoS literature further supported this conclusion.

Second, a review of the relevant SoS literature was conducted. This revealed an emerging consensus on the key features that differentiate SoS engineering from traditional systems engineering. Among those differences is the additional complexity with respect to decision making found in SoS arising from the independence of constituents. There is not a unitary decision maker from whom requirements can flow; rather a diverse set of decision makers are each with their own agendas. The development of techniques to better handle this challenge is identified as the focus area of the thesis.

Third, examining systems outside of the self-identified SoS domain, revealed that the same issues exist in other areas. In particular, the distributed decision making literature that is grounded in the logistics world offered key insights that were applicable to SoS. The AIR framework was developed by extending an existing framework (Schneeweiss, 2003) in distributed decision making. While that work dealt primarily with direct incentives as mechanisms for changing constituent behavior, the current research extended that to include other modes of influence. To move from the qualitative to the quantitative, a mathematical formulation of decision problems involved is proposed. This formulation turns out to have the structure of a principal-agent problem. The principal-agent approach and the related discipline of mechanism design form the inspiration for the 5 Is.

Finally, the intermodal freight transport case study revealed key lessons for putting the AIR framework and 5 Is into practice. AIR aided in organizing the modeling process for capturing the key interaction between shippers and carriers that create SoS behavior. The model used an agent-based approach with each agent attempting to make value maximizing decisions within a competitive environment. Issues such as the availability of information and different performance characteristics of transport modes were incorporated into the model. However, given the complexity of such systems in the real-world, predictive modeling is difficult. Rather, more qualitative insights such as the differential impact on carriers of influences strategies are extracted from the simulation results. The process of understanding why particular behaviors were observed was more valuable than the particular numerical results.

1.6 Impact upon systems engineering practice

The AIR framework and 5 Is can have significant impact on systems engineering practice. They provide a simple, consistent representation of the key roles decision makers take in an SoS. At the highest level, these are the constituents and the influencer. While the notion of constituent is not new, the notion of an 'influencer' is novel. More often than not (e.g. managing a communication or transportation network) system of systems engineers find themselves in this influencing role and can only indirectly effect the constituent systems within the SoS. Traditional systems engineering is predicated on the ability of the highest level stakeholder to proscribe requirement which determine decisions making at the lower levels. Such an approach would not work in SoS when there was a conflict between the needs of the system of systems engineer and that of the constituents. Rather strategies that account for the local needs of the constituents are required. The 5 Is are a first steps towards developing such strategies.

As is demonstrated in the case study, counter-intuitive results can occur when attempting to intervene in systems of such significant decision-making complexity. Therefore modeling such as the agent-based approach used in the case is crucial to gaining a sufficient understanding of the dynamics of the SoS before intervening in the real world. Examples of this are replete in case studies of real SoS (Krygiel, 1999). When trying to modernize document production in the DoD, the need of for common standards was identified. In implementing these standards problems arose given the diverse areas in which the standards needed to be applied. Furthermore, making such changes without disturbing on-going operations was quite challenging. Even though the end-state was much better than the status quo, there was a need to ensure local buy-in to make the transitions happen. AIR and the 5 Is can help the systems engineer think through such issues systematically before making changes in already operating systems.

1.7 Thesis outline

This thesis begins with an overview of the SoS literature and identifies the specific gaps being addressed (Figure 1-2). The benefits and limitations of commonly cited SoS frameworks are discussed. In addition, selected case examples are presented to highlight SoS concepts, particularly related to the dynamic interactions among constituent systems and other SoS stakeholders. It is shown through the literature and these examples, that distributed decision making is a key feature of many SoS. No one entity controls all the relevant factors needed to create SoS behavior. Rather those wishing to change the SoS must influence the other stakeholders to affect the desired change. How such an influence can be designed and implemented is the research focus of this thesis.

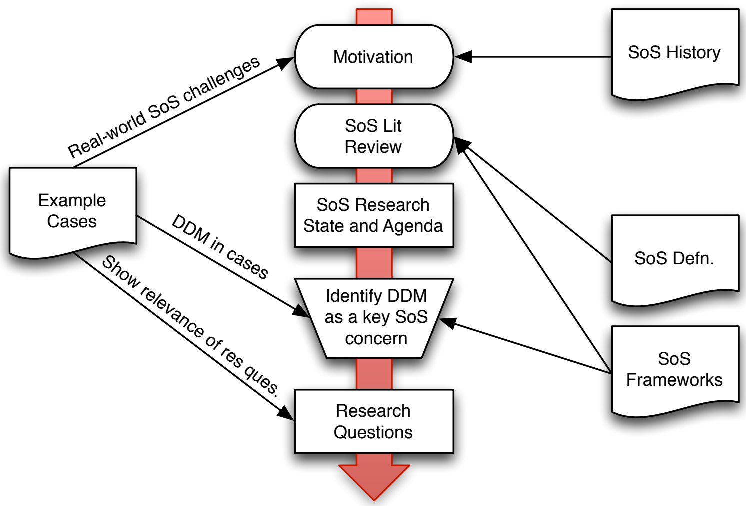


Figure 1-2: Flow of Chapter 2, Problem Formulation

Next, a new framework called AIR (Anticipation-Influence-Reaction) is proposed (Figure 1-3). The AIR framework organizes interactions between the relevant stakeholders of an SoS in a set of feed-forward and feedback loops between those who desire certain behavior from the SoS (known as *influencers*) and the constituents. The feed-forward loops, 'Anticipation', refer to the SoS influencer's attempt to estimate both current state of the SoS and model the decision making of the constituents. 'Influence' refers to the mechanism principals use to change constituent behavior. 'Reaction' describes the response of influenced constituents to the influencer after influences have been applied. This allows the influences to be modified to better drive SoS behavior. Each of the canonical SoS types, directed, acknowledged, virtual and collaborative (Maier, 1999; Dahmann and Baldwin, 2008), are described using the AIR framework.

Using the AIR framework, five strategy types that can be used by the influencer are proposed based upon a utility maximization representation of the coupled decision problems being solved by the influencer and constituents. These types, known as the 5 Is, are Incentives, Information, Integration, Infrastructure and Institutions.

These theoretical concepts are applied to the real-world example of an intermodal freight transportation system (Figure 1-4). A simplified intermodal freight transport system is represented us-

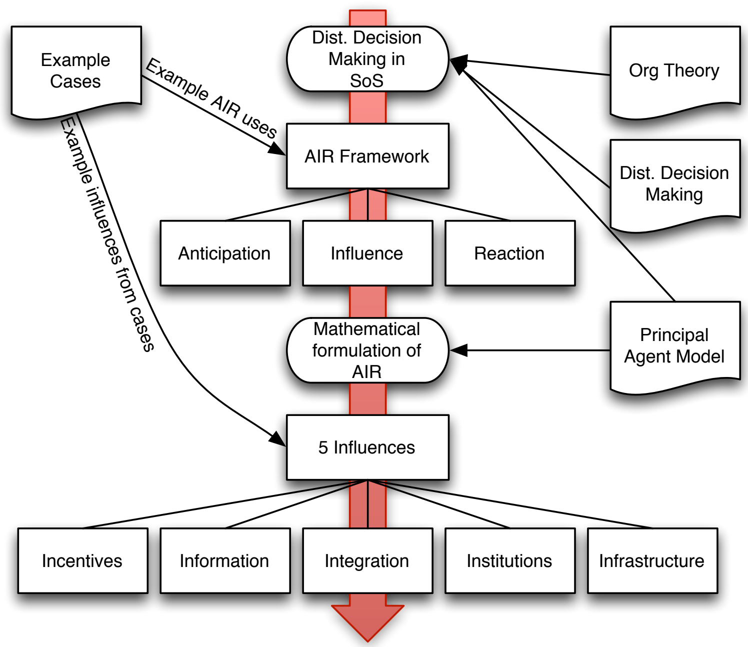


Figure 1-3: Flow of Chapters 3, 4 and 5, Distributed Decision Making, AIR and the 5 Is

ing the AIR framework. The constituent agents are the operators of the individual transportation modes. An influencer is defined in the form of a government agency who wishes to alter transport mode choices being made by users of the network. Baseline behavior is found to split traffic evenly between uni-modal and intermodal solutions. Using a simulation model, several influence mechanisms are demonstrated and their implications for the SoS as well as the constituents are discussed.

The thesis concludes with a discussion of the main contributions—the AIR framework and the 5 Is. Implementation issues are discussed with respect to availability of the information needed to formulate the problem and computational tractability as one scale to larger, more complex SoS. Challenges to the validity of the AIR representation of decision making are addressed along with possible alternatives. Finally several suggestions for extension of both theoretic and practical aspects of this view on SoS architecture are offered.

The next chapter provides a review of the field of SoS engineering, identifies the key research gaps and argues for the importance of the research questions listed earlier in addressing those issues.

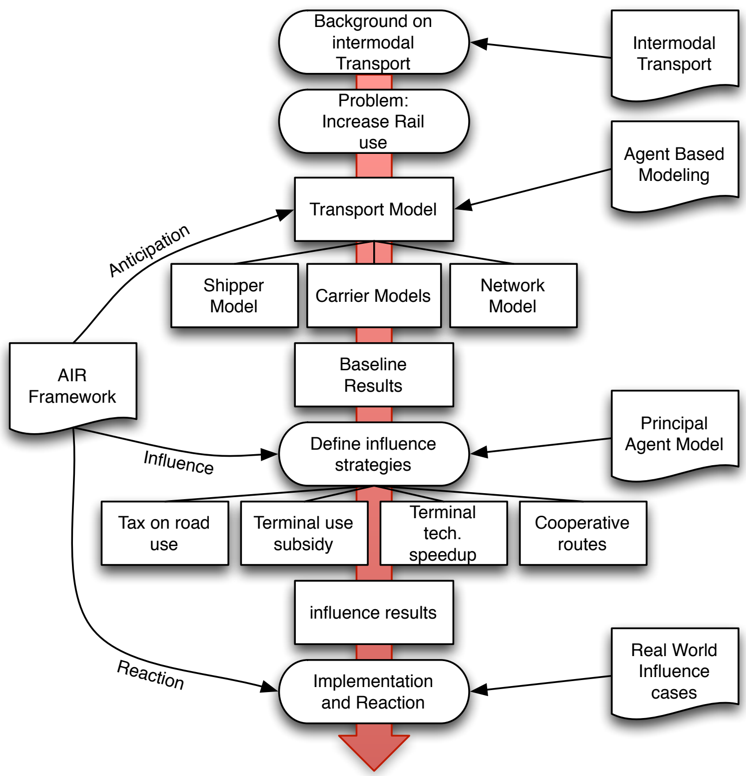


FIGURE 1-4: Flow of Chapters 6, Intermodal Freight Transport Case Study

Chapter 2

Case Examples and Literature Review

In this chapter a two-sided approach is taken in developing the research questions to be addressed. Reviewing the literature on SoS reveals a clear need for expanding the toolbox of techniques available to the systems engineer when dealing with SoS. It is argued that as a consequence of the independence of SoS constituents, decision making is distributed and, therefore, multiple, possibly conflicting value propositions must be accounted for. Further support is found by examining several real world SoS in which such stakeholder complexity has led to a unique set of challenges for those managing the SoS. Two research themes are identified. The first focuses on describing how interactions among constituents and SoS influencers determine SoS behavior. The second takes this framing and introduces a prescriptive framework to examine possible strategies that an SoS influencer could use to change this interaction.

2.1 What is a system of systems?

As SoS is an emerging discipline, establishing key definitions has been an active area of research. Before discussing the key research issues in SoS, a more detailed look at the definition of SoS and its development is provided in the next few sections expanding upon and giving context for the definitions presented in section 1.2.

Before defining system of systems, one must first define 'system'. While the word system is used in many different ways in many contexts[1], the sense intended here is that of an engineered creation. The International Council on Systems Engineering (INCOSE) defines a system as (INCOSE, 2006):

System A combination of interacting elements organized to achieve one or more stated purposes.

[1]Examples of other kinds of systems include: biological system such as blood circulation, social systems such as families, economic systems such as markets, etc. See, for example, the definition of system in Blanchard and Fabrycky (2005).

It is important to note here that defining such a combination of elements as a system is a categorical abstraction and thus depends upon how the specifier of the abstraction defines the elements

At first glance, the term system of systems (SoS) seems unambiguous enough. Taken literally, it would seem to mean simply a system that is composed of other systems:

System of Systems (Literal definition) An interacting combination of *systems* that accomplish a defined purpose.

While this definition is logically consistent, it does not capture all the significant distinctions between systems and SoS. Many systems, both big and small meet this definition. For example, the Internet is a system composed of interconnected computers that are system in their own right.

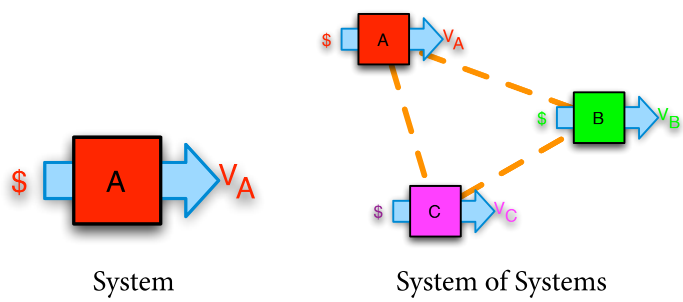


FIGURE 2-1: System vs. System of systems.

A system (Figure 2-1) is organized to convert resources into value for one or more stakeholders. A system of systems involves interaction between multiple systems to create emergent behavior. In trying to better characterize the difference between SoS and systems in general, authors produced a large variety of definitions (Over thirty are cited in Jamshidi, 2005). One approach which has some consensus in the literature are the criteria proposed by Maier (1999). He defines two characteristics that, when taken together, ensure that the systems of which the SoS is composed retain independence. The first characteristic, *Operational Independence*, requires that the SoS may be decomposed into its constituent systems, i.e., the constituent systems can be readily identified within the SoS. The second characteristic is *Managerial Independence*. This requires that the constituent systems retain the ability to make decisions regarding their internal operation while participating within the SoS. It is because of this form of independence that term *constituent system* is used to describe the systems that are joined in an SoS. There may also be other entities within the SoS that lack these independence properties. As an SoS is also a system, these components are referred to as either component systems or subsystems of the SoS. As will be seen in later examples, often these entities form the infrastructure that enables connection between the constituents and as such are a key leverage point for those who wish to modify SoS behavior. See Table 2-1 for a summary of the different types entities within an SoS.

Table 2-1: Constituent system vs. component system vs. subsystem

Element of an SoS	Operational Independence	Managerial Independence
Constituent system	**Yes**	**Yes**
Component system	**Yes**	No
Subsystem	No	No

Detailed discussion of the definition and characteristics of SoS is deferred to section 2.3.

As part of the literature review process, a series of historical SoS case examples were identified with a focus on the distributed decision-making. Background and narratives for these examples are presented in section 2.2. They will be referred to throughout the remainder on the thesis to provide empirical support for the theory being discussed.

2.2 SoS case examples

The following subsections present a select set of real-world SoS and highlight several common challenges that arise in SoS.

2.2.1 Peering among Tier 1 Internet Service Providers

In October of 2005, Level 3 Communications, a Boston based Tier 1 Internet service provider, decided to terminate its peering agreement with Cogent Communications, another Tier 1 provider (Team Register, 2005). By refusing to peer with Cogent, Level 3 cut-off direct traffic flow between their respective networks. This forced routing via third-party network increasing congestion on those links. Some customers whose only connection was via Level 3 were disconnected from those hosts whose only connection was via a Cogent network. The same was true in the other direction. After a few days, cooler heads prevailed and the peered connection was reestablished (Cowley, 2005). The underlying cause of the dispute was an imbalance in traffic flow between the two networks. Level 3 felt that Cogent was in violation of their contract when Cogent tried to make inroads into Level 3's market of selling access to Tier 2 providers. If a given Tier 2 provider, directly connected to Cogent instead of going through Level 3, this might create a traffic imbalance to Cogent's benefit.

Peering disputes continue to this day. Cogent itself has been involved in disputes with Telia, a Scandinavian provider, and Sprint (Singel, 2008). While at one level, these are purely business decisions, the operation of shared high bandwidth link is not fully automated and transparent. More human concern about how the two ISPs work together are important considerations in forming peering agreements. Recently, at a network operators conference (NANOG 47), Hurricane Electric, a smaller ISP, presented Cogent with a cake (Figure 2-2) asking for them to form a peering relationship (Miller, 2009). The numbers prefixed with "AS" identify Cogent and Hurricane Electric's networks as autonomous systems that can exchange traffic with multiple networks.

Figure 2-2: Cake presented by Hurricane Electric (Miller, 2009)

Looking at this case from an SoS perspective reveals some of pitfalls of combining systems that are managed and operated independently. The Internet is an SoS where the constituent systems are the individual ISP networks that are being interconnected. Each ISP is its own independent entity and, as demonstrated by Level 3's action. Level 3 had the ability and made the choice to terminate the peering agreement with full knowledge that this would have network-wide implications. Being both competitors as well as cooperators, the agreement between Level 3 and Cogent represented a balance between these two forces—giving them the benefits of peering while still leaving room for competition to provide service to Tier 2 ISPs. The decision making structure is distributed among ISPs and only through mutual self-interest does this SoS continue to opperate. When that mutual self-interest breaks down, as in this case, the SoS can quickly dissolve. Finding approaches to prevent such breakdowns is a fundamental challenge for SoS.

2.2.2 DSP Satellites providing early warning to Patriots

During the Gulf War, a key problem facing troops in the field was Iraqi Scud missiles. Getting early and precise warnings of Scud launches was crucial to successful use of countermeasures such as the American Patriot system and providing time for civilian populations to go to air-raid shelters. The Defense Support Program Satellites (DSP), originally launched to detect Soviet ICBM missile launches, had highly sensitive infrared telescopes. These telescopes were so sensitive that they could detect launches of individual Scuds. However, since the DSP system was designed for warning against strategic attack, the communication system for it was geared towards getting launch information back to commanders in the continental United States, not out to Patriot missile batteries in the field. Over several weeks, an automated interconnection was devised to allow DSP launch warnings to be communicated to Patriot batteries in an automated fashion gaining precious seconds of warning time. Two systems, Patriot and DSP, which were never meant to work together, were brought together to satisfy an emergent need (Cunningham, 1991; Anson and Cummings, 1991).

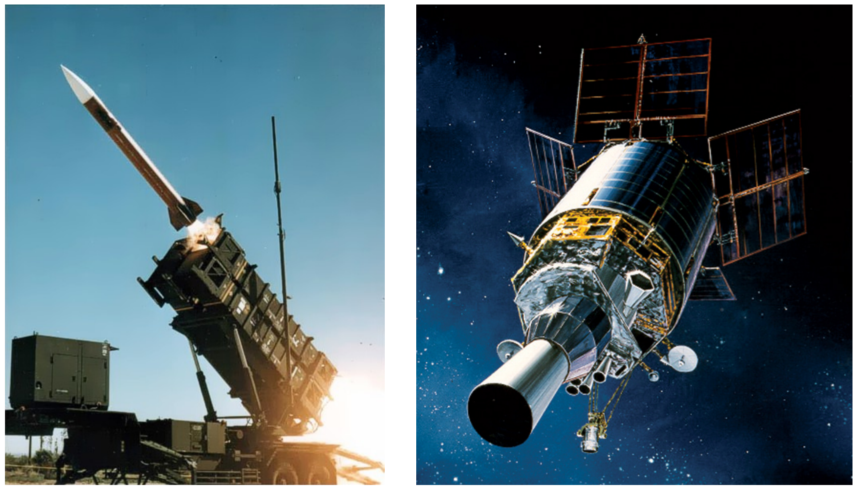


FIGURE 2-3: A Patriot missile (left) being fired and an artist's rendition of a DSP satellite (right)

Unlike the peering ISPs, the systems being combined here are not competing. Rather, they are both owned and operated by the DoD and, at some high level, share a common purpose and objective. However, when that common purpose got expressed as the specific requirements and technical forms for each of the systems, the result was two systems that performed their intended functions well, but were difficult to combine. In this particular case, the issue was that communication architecture for the DSP was engineered with that system's intended use in mind, i.e., to provide early warning of ICBM attack to strategic decision makers in DC, while the Patriot was built to best support its role as a theater-level asset. The main challenge in creating this SoS was not inter-organizational as in the case of the ISPs, but rather technical. Had communication standards been in place that foresaw the potential for inter-operation, much of the difficulty could have been avoided. A natural question then arises as why such standards were not in place and, in the future, what form should the standards take? The DoD is working through this very issue right now (AF/SAB, 2005a).

2.2.3 Housing Maps

In April of 2005, Paul Rademacher was looking for an apartment in the Bay area to match his new job at DreamWorks. Reverse engineering the JavaScript behind Google's at that time new mapping service, Google Maps, Rademacher overlaid Craigslist rental listings on Google Maps creating the first 'mash-ups' that have become at mainstay of the Web 2.0 world (Rademacher and Marks, 2012). Whereas before each website was an island of content, there is now a focus on combining content from disparate sources to form new and innovative applications. Such widespread use of content does come at a cost for content providers. In the case of Google Maps, housing maps and its imitators started to cause strain on servers. Google had been planning to release an API to allow access to map data in a manner that fit well with their overall infrastructure. HousingMaps beat them to the punch. An API or application program interface is set of software libraries and communication standards that specify how the website should interface with the Google Maps system and provides tools to help implement that specification. As a result, Google released an API shortly after HousingMaps went live and hired Rademacher (who has went on to lead the Google Earth API team).

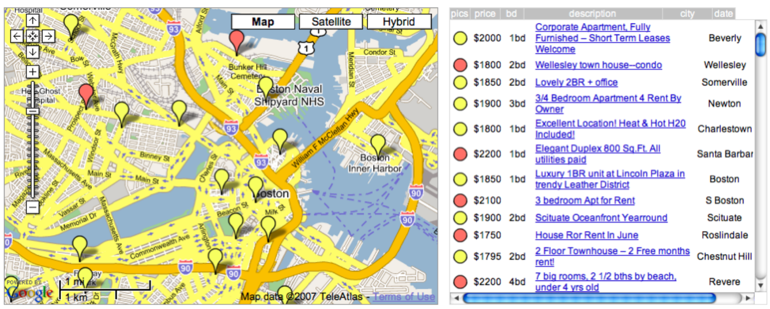


FIGURE 2-4: Screenshot of www.housingmaps.com

HousingMaps demonstrates that the emergence of an SoS may not involve active participation by all constituents. Given the opportunity to make connection between systems, users will make every effort to build SoS that fit their needs best. This has the clear advantage of shifting some of the problem of SoS design onto the user, but also requires that the constituent system managers are comfortable giving up some control to allow for user experimentation. Finding the right balance between control and user innovation (Oliveira and von Hippel, 2011) is a key challenge for those who wish to realize the 'flexibility' benefit of SoS in its full flower. Many data-sharing style SoSs such as the Global Earth Observation System of Systems (GEOSS) and DoD's Global Information Grid (GIG) face this issue.

2.2.4 Global Earth Observation System of Systems

The Global Earth Observation System of Systems (GEOSS) is an effort to combine and coordinate the collection, dissemination and exploitation of earth observation data[2]. A multinational effort, it is coordinated by the Group on Earth Observation (GEO), an inter-governmental organization with membership from 80+ countries. Each country contributes its own local data and expertise. As the assets that produce this data are all locally managed and operated, GEOSS is an SoS. One area of focus for GEO has been the establishment of data sharing standards to allow re-use of data collected by various GEOSS constituent systems (Rao et al., 2006). Establishing an information

[2]See http://www.earthobservations.org/geoss.shtml

system (of systems) for data sharing has been an early goal for GEO. Khalsa et al. (2009) describes a pilot program by which GEO is developing such a system.

A key challenge in building this SoS has been been the diversity of needs of the end-users combined with the distribution of decision making amongst globally (and, therefore, culturally) dispersed constituents (Eliot and Christian, 2005). GEO met this challenge by implementing a service-oriented architecture (SOA) for data sharing. The SOA allowed each constituent to chose which data they published and specified a common repository that served as a catalog for these data sources. The architectural choice to have a SOA allowed information sharing while also retaining autonomy of the constituents. The SOA was developed using an iterative, collaborative approach that incorporated feedback from progressively wider groups of stakeholders (Khalsa et al., 2009).

Table 2-2: Case examples summary

Example	Constituents	Observations
Cogent/Level 3 peering dispute	Cogent; Level 3	Participation in SoS can hinge upon ensuring that all parties needs are met. A local dispute between two parties can have wider impact when systems are interconnected.
DSP/Patriot interface	DSP satellite; Patriot missile battery	SoS can involve connecting systems that were never meant to be connected. This can be even more challenging when the systems have differing decision makers, objectives and heritage.
HousingMaps	HousingMaps; Google; Craigslist	SoS can arise 'from the bottom up' without permission (or even awareness) of all the involved constituents.
GEOSS	Members of GEO	Participants in collaborative SoS value their independence. SoS architecture needs to carefully balance this with need for coordination to ensure SoS value delivery.

2.2.5 Observations from the case examples

Table 2-2 provides a brief summary of the case examples. They demonstrate that when systems are designed/managed separately are brought together, the consequences can be both useful new capabilities, as well as new challenges for the system owners and those who bring the systems together. In the case of Cogent and Level 3, short term business considerations (Level 3's cashflow issues) caused an action whose impact spread far beyond Level 3's connection to Cogent and caused disruption and delay throughout the network. While, Level 3 has the right to do business with whom it chooses, this episode did reveal the need for governance structure to ensure that decisions with such far reaching implications are properly communicated so that network users can adapt and are not surprised. Establishing inter-operation between the Patriot and DSP systems revealed the difficulty posed by design decision entrenched in legacy systems when those systems are pushed beyond their original operational context. The HousingMaps mashup demonstrates that interconnection between systems need not be initiated by those who manage the systems. Rather, given the opportunity, users will form SoS to meet emergent needs. It is quickly becoming an antiquated notion that any system exists as an island with control over the demands and expectations being placed upon it. As interconnection becomes easier, system managers will face a choice of designing systems to either prevent or embrace interconnection and the emergent challenges it brings. To ignore the issue is to invite surprises. One Google employee said that HousingMaps "blew our minds right off our shoulders."(Charles, 2006)

Building upon this foundational work more recent efforts have focused on (1) developing and managing complex technical interfaces that underlie SoS and (2) viewing SoS as enterprises (Chen and Clothier, 2003) and developing management strategies that account for the many stakeholders involved (Sage and Biemer, 2007). The current research belongs to this second area. A common feature across all four examples is distributed decision making with SoS-wide consequences. The Tier 1 ISPs each make their own decisions regarding routing; design decision were made separately for the Patriot and DSP; Google Maps, Craigslist and HousingMaps are all managed separately; the members of GEO all represent independent nations. It is this characteristic more than any other that makes SoS engineering different from traditional systems engineering. Furthermore, in all these cases, stakeholders could not directly change constituents' systems other than their own. Rather, they have to use indirect means that influence constituents to change their systems. Stakeholders performing this role are known as SoS influencers. In the purest sense, SoS influencers have no authority over any constituent and are left with purely indirect access to the SoS components. An example of such indirect access can be seen in HousingMaps where Google could not change Radamacher's code to behave differently; they could only modify their own system (by technical means such as an API and legal means such as modifying the terms of service for maps) to encourage Radamacher (and other users) to make changes.

The result is a distributed decision making structure in which the controls (to use a cybernetic term) for the SoS are in the hands of constituents, while goals may be set by a different entity that must influence them to affect change. Therefore, the focus of this research is the impact on distributed decision making on SoS design, operation and management.

2.3 Systems of systems engineering

Among the earliest references to a 'system of systems' in the sense used in section 2.1 is a 1964 paper by Berry (1964) entitled "Cities as systems within systems of cities". Starting well before Berry and continuing through the next several decades, the discipline of systems engineering developed. Early (formal) studies of engineered systems developed notions of decomposition and hierarchy (Simon, 1996).

By the early nineties this notion had been extended to systems composed of other systems. In a 1991 paper, Eisner et al. express a need for extending the systems engineering paradigm to include 'systems of systems'. Table 2-3 gives Eisner's seven criteria for using SoS methods and his contrast for each in the traditional SE.

Table 2-3: Systems engineering vs. SoS engineering (Eisner et al., 1991)

	SoS Engineering	Traditional SE
1	There are several independently acquired systems, each under a nominal systems engineering process.	Subsystems are acquired under centralized control.
2	Overall management control over the autonomously managed systems is viewed as mandatory.	The program manager has almost complete autonomy.
3	The time phasing between systems is arbitrary and not contractually related.	Subsystem timing is planned and controlled.
4	The system couplings can be considered neither totally dependent nor independent, but rather interdependent.	Subsystems are coupled and inter-operating.
5	The individual systems tend to be uni-functional and the systems of systems multi-functional.	The system is rather uni-functional.
6	The optimization of each system does not guarantee the optimization of the overall system of systems.	Trade-offs are formally carried out in an attempt to achieve optimal performance.
7	The combined operation of the systems constitutes and represents the satisfaction of an overall coherent mission.	The system largely satisfies a single mission.

These differences define SoS engineering as a unique class of problems related to, but different from, systems engineering. Eisner highlights several issues that are fundamental to understanding the difference (taking each criteria in turn):

1. Acquisition of constituent systems
2. Degree of centralization of control of constituent systems
3. Evolution of the constituent systems and the SoS over time
4. Degree of interconnection between constituent systems
5. Functional specification of SoS
6. Performance of constituent systems vs. performance of SoS– What does optimality mean in a multi-function context?[3]
7. Multiple citizenship – Constituent systems may participate in multiple SoS whose combined action produces the desired end.

[3]This is similar to the notion of *evaluative complexity* described by Sussman (2002).

These same themes appear in later definitions and form the basis of the various taxonomic systems that have been proposed.

In the middle of the decade, Maier (1999, first published as a conference paper in 1996) proposed a definition for SoS that has been cited frequently in subsequent literature. Observing increasing prevalence of systems that are composed of other, distinct in their own right, systems, Maier defined three classes of such systems as distinguished by their management structure:

Directed: Directed SoS are built to meet a specific need or purpose that is promulgated by central, directing, authority. While the constituent system maintain the ability to operate independently, they give higher priority to the instructions from the central authority than delivering on their local value proposition. Rather, the central authority dictates constituent action so as to accomplish the centrally specified purpose. A joint military unit under a single unified command is a example of this type of SoS.

Collaborative: Collaborative SoS differ from directed SoS in that there is no central authority with coercive power over the constituents. The constituents' actions are governed by both needs of the SoS as well as local considerations specific to each constituent. SoS objectives emerge from the collective agreement to pursue an agenda by the constituents. The Internet as governed by the Internet Engineering Taskforce (IETF) is an example of a collaborative SoS.

Virtual: Virtual SoS lack any central authority. Their behavior emerges from the unplanned, not fully coordinated interactions of constituents.

In addition, Dahmann and Baldwin (2008) later identified a fourth category that combines aspects of Maier's Collaborative and Directed:

Acknowledged: Acknowledged SoS have a central authority like directed SoS, but that authority lacks coercive power over the constituents. The constituents retain their own budgets, decision making and objectives. SoS behavior is brought about by collaboration between the constituents and the central authority. The Army's Future Combat System (FCS) is an example of an acknowledged SoS. While there is an overall objective to create modern, lighter, more networked integrated force and a program office charged with making that happen, many of the individual components of the FCS are acquired through independent acquisition effort and/or are derived from legacy systems with extant program management infrastructures.

Applying these categories of SoS to actual cases can be difficult as the particular category that best describes an SoS can change over time as its structure and objectives change to reflect new circumstances. To help illustrate this, consider the SoS examples discussed in sections 2.2.1 – 2.2.3.

Initially, the precursor networks (e.g., ARPANET, Bitnet, USENET) that would combine to form the Internet could be best described as a directed SoS. Leiner et al. (2003) recounts the story of the development of the initial Internet protocols as a result of an NSF (National Science Foundation) contract to interconnect computers at several leading research institutions to ease exchange of scientific information and improve collaboration within and among these groups. The NSF contract provided specific direction as to which independent systems were to be interconnected and the objective to be fulfilled by the resulting network. A key clause in this contract required that the network be designed such that uses (and users) beyond those specified in the contract could be added at a later date. The result of this clause was that the protocol stack developed could be ported to other applications beyond scientific collaboration and to other networks. The protocols and associated technologies were offered royalty free and were quickly adopted by others including commercial networks. As the other networks interconnected, in particular those that were outside the group, the SoS became more collaborative in character. An example of this is the peering structure used to allow traffic flow between Tier 1 ISPs (see autorefsect:InternetPeering). Standards bodies such as IETF were also formed by granting authority and independence to the working groups formed by the government in managing the precursor networks (Leiner et al., 2003). These groups established community norms amongst the now collaborating set of systems.

Over time, these technical standards and social norms become firmly entrenched. They can be treated as the infrastructure and social institutions upon which other systems rely upon. This stability enables the creation of higher-level applications that combines Internet services in novel ways never intended by those who formed the SoS. HousingMaps, introduced in subsection 2.2.3, is a recent example of this. It began as a virtual SoS in which the constituent systems, Google Maps and Craigslist, were not aware that their respective services were being combined by a third party. Once Google became aware of Radamacher's website and instituted policy to control how their service were accessed, the SoS became collaborative as Google was both acting as a constituent and using their influence to effect the behavior of other stakeholders.

This variation in SoS type over the lifetime of systems indicates that the Maier/Dahmann taxonomy is not best suited as an absolute classification of SoS but rather as descriptors of different constituent/authority relational structures that an SoS may exhibit at a given time in a given context. Viewed from this perspective, one can think of these classes as defining different types of SoS systems engineering challenges than different types of SoS.

Looking at these types of SoS, Maier distinguishes those SoS that are developed through constituent interaction[4] as fundamentally different from directed SoS. In the directed case, since control is centralized, the design challenge is, principally, a problem of managing the interoperation of systems that may have not been designed to do so. In contrast, the multi-decision maker cases of acknowl-

[4]Either via direction from an external authority in the case of acknowledged SoS or as a result of collective decision making in collaborative SoS.

edged and collaborative SoS have an additional social challenge of coordinating the multiple, possibly conflicting, agendas of the decision makers.

In addition to fitting the literal SoS definition, those SoS which arise from collaboration demonstrate the two additional properties defined by Maier, operational independence and managerial independence (see section 2.1). Systems that exhibit these properties tend to fall on the SoS side of the seven Eisner (see Table 2-3) distinctions between SE and SoSE. Operational independence maps to distinctions 4 and 5—the constituents are not tightly coupled and can function independently. Managerial independence maps to distinctions 2 and 7—the constituents are controlled in a decentralized manner (though possibly governed by centralized rules) and the constituents can simultaneously participate in multiple SoS as well produce value on their own. Maier requires only these two properties. He does mention three properties that are not required but do frequently appear in collaborative SoS. They are:

Emergent behavior A SoS's behavior emerges from the interaction of the constituents. This maps to distinction 6 as the SoS is 'more than the sum of its parts'.

Geographic distribution SoS tend to be geographically distributed.

Evolutionary development The SoS does not change in a planned or directed fashion; rather it evolves organically responding to changing needs and pressure levied by the constituents. This maps to distinction 3.

This formulation of key properties that define SoS has proved quite robust surviving into more current publications such as (Bjelkemyr et al., 2007; Sage and Biemer, 2007). Other authors have used different terms for these properties as well as adding additional considerations to the list. Boardman and Sauser (2006) asked the question: *What does the "of" in system of systems mean?* Their answer identifies elements of difference between systems and systems of systems:[5]

> "**Autonomy** System components[6] cede their decision making to the system; SoS constituents retain autonomy (or managerial independence as per Maier).
> **Belonging** System components do not choose to be members of the system; SoS constituents join by choice given a belief in the overall SoS purpose.
> **Connectivity** Interfaces between system components are designed and instantiated to enable system-level behavior; SoS interfaces are dynamically supplied by the constituents and may be driven by local constituent-constituent interaction or by SoS needs.

[5] Boardman and Sauser speak mostly to collaborative SoS. The applicability of their distinctions to the other types of SoS is less clear.

[6] The term *components* as used here by Boardman and Sauser refers to elements of a system that is not an SoS. This is distinct from *component system* as defined in section 1.2 which refers to elements of an SoS that lack managerial independence.

Diversity Uniformity is a desired characteristic of system components to ease project management issues and reduce sustainment costs; Diversity is fostered among SoS constituent by allowing them autonomy with respect to their own evolution and encouraging open connectivity with other constituents.

Emergence System behavior (both beneficial and problematic) is foreseen and planned for during system and component design; SoS behavior emerges from constituent interaction. External (or internal) agents may try to influence this interaction, but it is not centrally controlled and may not even be understood." (Boardman and Sauser, 2006)

Many of these same themes are seen in Eisner's original list of distinctions. As more and more SoS came into being, the systems engineering community quickly realized, as Eisner and Maier had years earlier, that engineering of SoS was a fundamentally different problem than traditional systems engineering. This realization led to flurry of activity in the late 1990's and early 2000's. However, since so many well-intentioned researchers from so many different disciplines tried to approach the SoS problem simultaneously, there was bound to be subtle but significant differences in their characterization of the problem. In their review paper, Keating et al. (2003) list six representative definitions from the literature along with the primary focus and application area in which they were devised:

"**Manthorpe, Jr. (1996)** Primary focus: Information superiority. Application: Military. *In relation to joint warfighting, system of systems is concerned with interoperability and synergism of Command, Control, Computers, Communications, and Information (C4I) and Intelligence, Surveillance, and Reconnaissance (ISR) Systems.*

Kotov (1997) Primary focus: Information systems. Application: Private Enterprise. *Systems of systems are large scale concurrent and distributed systems that are comprised of complex systems.*

Lukasik (1998) Primary focus: Education of engineers to appreciate systems and interaction of systems. Application: Education. *SoSE involves the integration of systems into systems of systems that ultimately contribute to evolution of the social infrastructure.*

Pei (2000) Primary focus: Information intensive systems integration. Application: Military. *System of Systems Integration is a method to pursue development, integration, interoperability, and optimization of systems to enhance performance in future battlefield scenarios.*

Carlock and Fenton (2001) Primary focus: Information intensive systems. Application: Private Enterprise. *Enterprise Systems of Systems Engineering is focused on coupling traditional systems engineering activities with enterprise activities of strategic planning and investment analysis.*

Sage and Cuppan (2001) Primary focus: Evolutionary acquisition of complex adaptive systems. Application: Military

> *Systems of systems exist when there is a presence of a majority of the following five characteristics: operational and managerial independence, geographic distribution, emergent behavior, and evolutionary development.*" (Keating et al., 2003)

Reviewing these definitions, it is clear that they all accept the literal definition of an SoS. Other similarities include a sense of 'large-scale' and a sense that the SoS can do more/behave in different ways than its constituents. There are differences however. Each community defines the constituent systems in very different ways. What one considers a system, others consider a component. For example, a single computer may be the system in a distributed computing environment (the SoS) for an enterprise. The same distributed computing environment may be a system within the larger enterprise SoS that also includes manufacturing plants and support facilities. The particular level of abstraction at which one applies the term SoS (or systems or subsystem) is very problem and application context dependent. One person's system is another person's component and yet another's SoS. Nightingale (2003) and, later, Sgouridis (2007) takes this concept even further viewing the commercial aviation industry as an *Enterprise of Enterprises*. The system boundary is also different from one application context to the next. Some may only consider the technical aspects as part of the systems while others will bring social and economic relationships.

The authors cited in Keating et al. (2003) also view the SoS as existing to serve different purposes. Those coming from the military domain tend to view the SoS as something that is constructed to achieve specific performance improvement or create specific new capabilities, e.g., improved situational awareness, faster fusion of disparate data sources or better utilization of scarce assets. On the other hand, the IT community treats corporate enterprise IT system as SoS. Unlike the military case, these systems support a wide variety of business processes. The diversity of these processes leads to non-process specific metrics and a focus on developing an effective platform as opposed to specific capabilities. In some sense, both communities are really talking about the same thing. The military needs to think about multi-use platforms to make effective use of these new capabilities in a changing battle-space, while the IT community needs to demonstrate performance in order to justify investment. However, the difference of emphasis is reflective of the norms within these two communities. This fragmentation has led to slow cross-adoption of lessons and techniques between fields.

In addition to the term system of systems, two other terms are found in the literature to refer to systems that are composed of other systems. Both terms refer to a sub-class of SoS that exhibits a governance structure that emerges from the constituents as opposed to some central authority. Building upon the work of Krygiel (1999); Sage and Cuppan (2001) define a federation of systems (FoS) as an SoS (in the Maier sense) that is governed by a coalition of the constituents and adheres to the maxim "only make common that which needs to be common; leave the rest to the states." Such federations exist to serve the needs of constituents and further their individual objectives. As such their relationship is more akin to a federation or a coalition of convenience rather than supporting an overall vision or objective. The lack of such an overall objective in federations leads to different dynamics in the economic and policy domains and thus FoS is useful a classification distinct from purpose driven SoS. A family of systems (also abbreviated FoS) is not a SoS. Rather this term refers to a group of systems collected into a 'family' because of commonalities in, for example, interfaces or operational use. An example of a family of systems is a group of systems created upon a common platform, e.g., car models that use a common chassis. Members of a family of systems need not interface with each other or produce group behavior beyond that of their individual action as is the case for an SoS.

Following this period of many different definitions,[7] there seemed to be greater recognition that there does exist a class of systems called 'systems-of-systems' that differ from other systems in that they are composed of several autonomous components who may not have been designed for the SoS. Such SoS could be further categorized by decision making structure. This was captured in the 2008 Guide To System of Systems Engineering published by the Department of Defense (DoD). In producing this guide, the DoD spent several years synthesizing the growing body of literature supporting the existence of this class of systems and proposed a definition similar to the literal definition stated at the beginning of the chapter:

> An SoS is defined as a set or arrangement of systems that results when independent and useful systems are integrated into a larger system that delivers unique capabilities. Both individual systems and SoS conform to the accepted definition of a system in that each consists of parts, relationships, and a whole that is greater than the sum of the parts; however, although an SoS is a system, not all systems are SoS. (DoD, 2008)

Given the importance of the DoD within the SoS community their definition for SoS will be used as a baseline construct for scoping those systems that exhibit SoS character. DoD also reiterated the four types of SoS distinguished by decision making structure, i.e., directed, collaborative, virtual and acknowledged. In doing so, they recognized importance of differences in decision making structure as central to understanding SoS.

One should not take this definition to mean that SoS are composed of *only* independent constituent systems, there may be other elements present that serve as interfaces or enablers between these constituents.[8] These additional elements are referred to as *component systems* if they are operationally independent or *subsystems* if not (see definitions in section 1.2). They do not exhibit the independence criteria from Maier. In practice there is often disagreement as to the extent to which a particular system is an SoS. Finding a universal bright line criteria to separate systems and SoSs

[7]For a extensive review of different definitions from the literature, see (Jamshidi, 2005)
[8]For example, gateway routers between networks on the Internet.

would require making arbitrary taxonomic choices. Many real-world system exhibit, to some extent, the characteristics ascribed in the literature to SoS. They interact with other systems to produce value and, at a minimum, exist within a changing context where those changes are driven by other, autonomous, systems.

Qualitatively, one can imagine a continuum between isolated systems and SoS. At one end are systems that are developed and operated in isolation and with fixed contexts. At the other end of the continuum are systems like air traffic control, the Internet and joint military operations that all exhibit collaboration as a fundamental characteristic. Between these two extremes lie most systems. They have some components that were designed specifically for the system in question. They also have components that are imported from the outside. Finally, they interact with other systems within a shared context. Very few systems of interest to the modern systems engineering community exist at the 'isolation' end of this spectrum. As described by Maier, the power to gain new capabilities lies at the interfaces between systems.

The challenge to the system engineering community is not one of taxonomy, but rather developing tools and techniques to design and manage systems with these characteristics whether called SoS or not. The descriptive focus, therefore, of this thesis is not on settling the debate of system vs. SoS, but rather providing a conceptual framework for understanding key aspects of complexity within an SoS. More specifically, this thesis examines the decision making and influence relationship between the various stakeholders involved and the system components they control. To that end, attention is now shifted to further characterizing these classes and developing frameworks to describe those aspects of SoS most salient to system design and management.

2.4 Frameworks for engineering of SoS

The design and management of SoS is a problem of coordinating the parallel development and operations of the SoS with its constituents. In order to manage the inherent complexity within an SoS, several frameworks have been proposed for the design and management of SoS. The first and earliest was developed by DeLaurentis and Callaway (2004). They examine four domains in framing the SoS problem: *Resources, Operations, Economics and Policies*. *Resources* are the physical entities within the SoS. *Operations* refers to the policies and procedures that direct the activities of those physical entities. *Economics* refers to the sentient entities that provide the SoS with capacity to change and respond within a multi-stakeholder market economy. *Policies* are the external forcing functions that impact the physical and non-physical entities.

In the framework, the SoS is then described through four hierarchies, one for each category—a physical hierarchy of resources, a functional hierarchy of operations, a decision/authority hierarchy of economics and a rules/influence hierarchy of policies. Given the nested nature of a system of systems implied by literal definition, the issue of systems *composed of* systems of systems arises.

DeLaurentis and Callaway address this by arranging SoSs and their component systems into a hierarchy whose base level, i.e. the set of component systems that are not decomposed is decided by the system architect so as to best scope the analysis at hand. Entities (either physical or non-physical) at this base level, labeled α, are organized into groups based upon interfaces to form entities at the β level. Entities at the β level are in turn organized into grouping based upon interfaces to form γ level entities. This process continues recursively until all entities are aggregated into a single top-level entity that encompasses the SoS under consideration.

As an example, DeLaurentis and Callaway present the national transportation system with four levels of aggregation as depicted in Figure 2-5. Entities at each level are then described with respective to each of the four domains to arrive at an integrated understanding of the system (see Table 2-4).

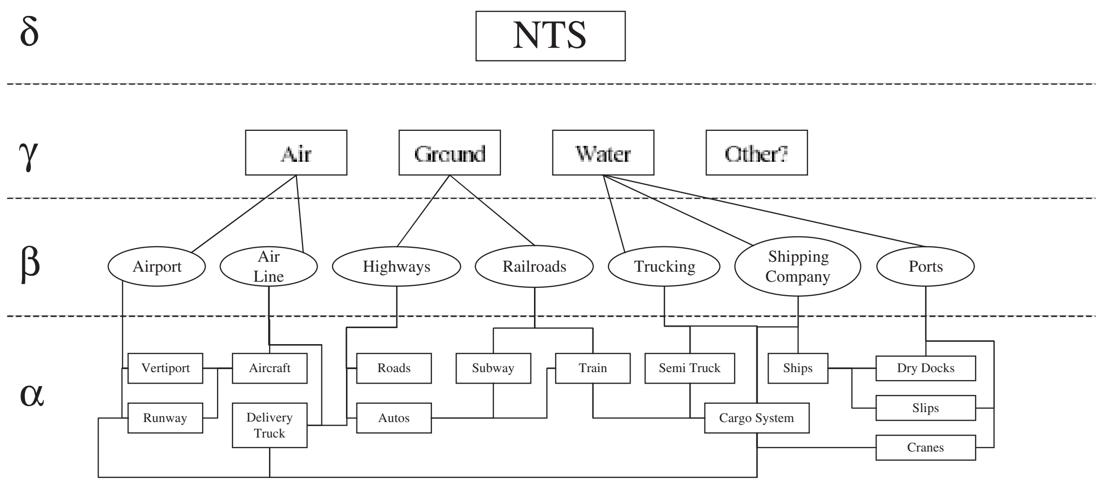


Figure 2-5: Decomposition of the resource domain of the national transportation system of systems from DeLaurentis and Callaway (2004). Links within a level indicate interfaces. Links between levels indicate aggregation.

Both problems of traditional systems engineering and SoS engineering can be represented using the DeLaurentis-Callaway framework. The two regimes can be distinguished by considering the relationship between domains/levels in the framework. For example, if all decision-making authority can be traced back to high (say γ or δ) levels in the hierarchy, then centralized control exists, and traditional SE tools are applicable. If most of the operational decisions are made within a level and there is little imposition of rules from above or below, then entities at that level will act independently and so require a different approach than when controls can be applied from higher levels.

While an excellent first step beyond the definitions, improvements can be made to the framework. One major missing element is the temporal dimension. As Eisner pointed out, a feature of SoS behavior is independent evolution of constituent systems. Conversely, as the systems interact and

Table 2-4: Definition of entities at five levels of decomposition across domains for the national transportation system of systems (DeLaurentis and Callaway, 2004).

	Resources	Operations	Economics	Policy
α	Vehicles & infrastructure (e.g., aircraft, truck, runway)	Operating a resource (aircraft, truck, etc.)	Economics of building/operating/-buying/selling/-leasing a single resource	Policies relating to single-resource use (i.e., no. of attendants per passenger for vehicle type)
β	Collection of resources for a common function (an airport, etc.)	Operating resource networks for common function (e.g., airline)	Economics of operating/buying/-selling/leasing resource networks	Policies relating to multiple vehicle use (i.e., local airport noise policies)
γ	Resources in a transport sector (e.g., air transportation)	Operating collection of resource networks (e.g.,commercial air Ops)	Economics of a business sector (e.g., airline industry)	Policies relating to sectors using multiple vehicles (FAA certification, safety, etc.)
δ	Multiple, interwoven sectors (resources for a national transportation system)	Operations of Multiple Business Sectors (i.e.,operators of total national transportation system)	Economics of total national transportation system (All Transportation Companies)	Policies relating national transportation policy
ϵ	Global transportation system	Global operations in the world transportation system	Global economics of the world transportation system	Global policies relating to the world transportation system

begin to rely on one another for certain function, their evolution may become interdependent. One can infer some of these system lifecycle issues from the data in the framework; however explicitly representing the lifecycles at each level of decomposition will reveal the dynamics of change within the SoS and show the evolutionary influence felt and produced by each resource. This would allow, for example, analysis of the co-evolution of airplane designs with the size and locations of the airports they serve. A further limitation is that the interaction between the domains are not explicitly addressed. How does economics effect the availability of resource? How does policy constrain reactive decision making during operation?

A different approach is taken by Norman and Kuras (2004). They develop the notion of a *complex system* (CS) as a system "Whose structure and behavior is not deducible, nor may it be inferred, from the structure and behavior of its component parts; Whose elements can change in response to imposed pressures from neighboring elements (note the reciprocal and transitive implications of this); Which has a large number of useful potential arrangements of its elements; That continually increases its own complexity given a steady influx of energy (raw resources); Characterized by the presence of independent change agents." (Norman and Kuras, 2004) They then argue that traditional systems engineering is ill equipped to handle systems with these properties and that expansion of system engineering is needed, i.e., a *complex systems engineering* (CSE). In doing so, they claim that CS, being composed of independent agents, cannot be managed with a "find a solution for a given set of requirements" mind-set, rather CSE involves invoking an appropriate set of pressures upon the interacting entities to cause them to act in the desired manner.

Taking a more prescriptive viewpoint, Sage has developed an SoS engineering process over several papers (Sage and Cuppan, 2001; Morganwalp and Sage, 2003; Sage and Biemer, 2007). In the first paper, Sage and Cuppan (2001), a connection is made between complex adaptive systems and system possessing the characteristics outlined by Maier in defining collaborative SoS. The connection to adaptive system brings a focus of dynamics and self-change that was missing from DeLaurentis and Callaway's framework. By acknowledging that the SoS's structure can be changed by the constituent decision makers, Sage and Cuppan bring focus on the relationship between those who control the constituent systems, i.e. those who take adaptive actions, in addition to the interfaces between the constituent systems.

In looking at the relationships among the stakeholders who control the constituents of the SoS Sage and Cuppan invoke the metaphor of a federal government citing Krygiel's federation of systems construct. This view emphasizes the collaborative nature of the relationship seeking mutual benefit. In contrast, DeLaurentis and Callaway invoked economies as the dominant metaphor for the constituent relationship. In reality most constituent communities fall somewhere in between these two extremes. For example Internet service providers need to cooperate through peering agreement to share traffic in order to provide customers access beyond their own networks. At the same time, they compete for connection from those very same customers. In transportation, rail and truck carriers normally compete, but can also cooperate to form intermodal routes (see chapter 6).

Morganwalp and Sage (2003) extend the work in Sage and Cuppan and propose an architecture development process for SoS. They view the stakeholder community that generates SoS as an enterprise that manages groups of systems that are combined and used in different ways over time to form specific SoSs as needs change. The process is very much a top-down, enterprise need driven model for how SoS should be developed. Mirroring DeLaurentis and Callaway's multi-dimensional representation of SoS, Morganwalp and Sage retain a hierarchic decomposition. However, instead of using generic labels, they propose a fixed hierarchy starting at the enterprise which manages families of systems (that are used to instantiate SoS). Families of systems are composed of systems that are composed of subsystems which in turn are composed of parts. At each of these levels of decomposition they explore the perspective of several stakeholders by asking who, what, when, where, how and why questions resulting in three dimensional characterization enterprise (level of decomposition, stakeholder, and question; see Figure 2-6). The architecture development process proposed based upon this framework performs top-down decomposition of the architecture with higher level defining needs at lower levels. This reflects heritage from the long tradition of decompositional system design methodologies prevalent in systems engineering most commonly instantiated in the V-model.

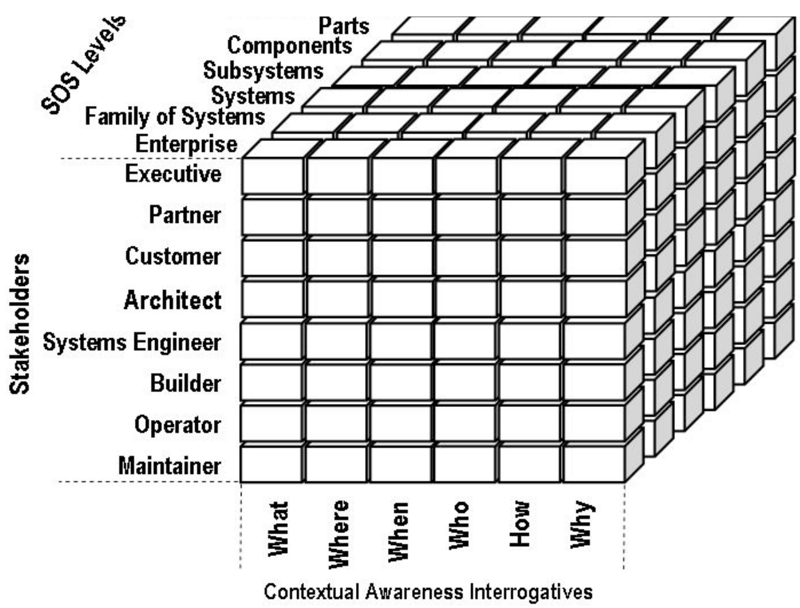


FIGURE 2-6: Enterprise Architecture Framework (Morganwalp and Sage, 2003)

Building upon several existing systems engineering government and industry standards and practices, Sage and Biemer proposes an SoS engineering process. Using the Morganwalp-Sage Enterprise Architecture Framework, the design and operation of an SoS is envisioned as the interaction between an enterprise that is providing the constituents community with resources and guiding (but not controlling) their development through scenarios with a SoS integration and operation effort that uses constituents to satisfy specific enterprise specified requirements (see Figure 2-7).

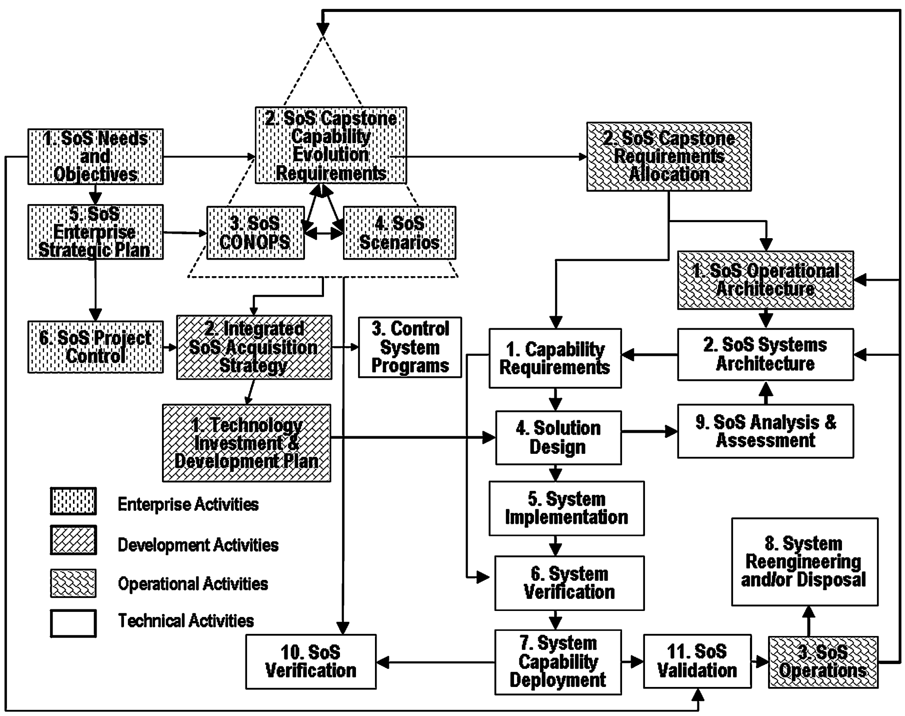


FIGURE 2-7: Process diagram for SoS engineering (Sage and Biemer, 2007).

Once again, the separation between the design and evolution of the constituents and the SoS is cited as a rationale for the need for a new systems engineering process. The implication being that traditional systems engineering, given its decompositional view of the design problem, is ill suited for situations in which design of components is done without such decomposition. While the authors discussed thus far have used this criteria as a key discriminator between systems and SoS, Walden (2007) takes a more nuanced approach.

In framing the SoS problem, Walden recognizes that simply having components that are not designed with the system's requirements in mind is insufficient to separate systems from SoS. He uses the example commercial-off-the-shelf (COTS) acquisition as a situation in which components that are not designed for the systems are nonetheless integrated into the system. In the case of COTS development, though the components are developed independently, once integrated into the system, they lose this independent identity and are tailored to meet system needs. This is not true of constituents of SoS. The independence properties first defined by Maier imply that some control is retained by the constituents, the degree and nature of that control delineates the type of SoS (collaborative, directed, etc.). Given the limited control that SoS designers have over their constituents, Walden uses the term "conducting" as the dominant metaphor. Just as the musicians in an orchestra retain independence while joining the group, so do constituents.[9] The SoS designer must orchestrate these elements to create SoS behavior while being mindful of the needs of the constituents. Compared to the other metaphor discussed earlier, DeLaurentis and Callaway's economy and Morganwalp and Sage's enterprise, this metaphor seems the most broadly applicable. The musicians exhibit the characteristics proposed by Boardman and Sauser and by Maier. The analogy fails to work however when one tries to apply the suggested systems engineering processes such as in Sage. The process by which an orchestra prepares a piece is more one of successive refinement of a baseline than decomposition followed by synthesis.

Boehm and Lane (2006) propose just such a baseline followed by refinement approach to SoS. Building upon foundational work from software engineering (Boehm, 2000), they propose a spiral development model of SoS. Refinement and alteration to the SoS occur at both planned points and opportunistically as the SoS designer understanding of and relationship with the constituents changes. Within each spiral, a decomposition-synthesis approach could be used. When the constituents have greater autonomy however, such an approach may be difficult to implement since each constituent will interpret the SoS as to best meet their local concerns. In examining the learning process by which the SoS designer discover the capabilities possible with a set of constituents, AF/SAB (2005b) suggested that venues be created within which constituents and SoS designers could experiment with potential SoS configuration prior to fielding.

In more recent times, efforts have made to further formalize SoS from both descriptive and process perspectives. One such effort is being carried out at Sandia National Laboratories. Ames et al. (2011) has proposed a more detailed framework than those above called the Complex Adaptive System of Systems (CASoS) Engineering Framework. They attempt to combine the systems engineering process constructs as seen by Sage and Biemer (2007) with the complex systems ideas used by Norman and Kuras (2004). From a process perspective they build upon a traditional SE model moving from conceptualization of the desired system, to developing and testing of its components to finally fielding and monitoring. Recognizing independence however, they frame each of the these process areas as occurring concurrently with feedback loops connecting them. At the highest level goals are specified broadly and are termed 'aspirations'. The constituents are given broad authority and are presented with opportunities to collaborate during modeling and testing. The in-

[9]There are conductorless orchestras which could be compared to federations of systems (Holland, 2005).

tent there is to create opportunities for novel SoS forms to emerge. This distribution of authority (but still governed by the 'aspirations') is an enabler for the adaptive character of the CASoS.

Karcanias and Hessami (2011a,b) attempts to formalize the notion that SoS change over time by introducing "system plays". A "system play" is a particular instantiation of an SoS at a particular time and place. In addition, the "play" includes the actions taken by the constituents (and other relevant actors) under a specified scenario. Over time a given group of constituents (termed a composite of systems by Karcanias and Hessami (2011a)), will participate in many different "plays" in response to changing external stimuli. The notion of a "play" is potentially useful to capture the transient nature of SoS.

2.5 Literature summary

To summarize the discussion of SoS definitions, frameworks and agendas; the following is the 'state of the theory of SoS'.

A *system of systems* is a system that is composed of other systems known as constituents systems (DoD, 2008). These systems retain autonomy, though not necessarily complete, over their local concerns. The degree of autonomy held by constituents classifies SoS into four types, directed, acknowledge, virtual and collaborative.

The design and management of SoS is a problem of coordinating the parallel development and operations of the SoS with its constituents. Such coordination can be externally imposed such as in an enterprise (Morganwalp and Sage, 2003) or arise as consequence of interaction between the constituents (Krygiel, 1999).

A variety of processes and approaches for managing SoS have been proposed (DoD, 2008; Sage and Biemer, 2007; Boehm and Lane, 2006); however, there is no consensus on which approach works best. Differences between these approaches are principally concerned with the extent to which the SoS designer has the necessary information and control to treat design problem as a top-down effort vs. requiring the cooperation and participation of the constituents in a more bottom-up design process (the spiral model). The system engineering community is currently struggling to find the appropriate tools and processes to be used for SoSs that fall in different places on this spectrum.

The design of SoS is a two-sided problem. On the one hand, it is a technical problem of the determination of the appropriate interfaces (Maier, 1999) between constituent systems in order to accomplish SoS objectives. On the other hand, it is a social problem of convincing constituent decision makers to actually implement such interfaces (Dahmann and Baldwin, 2008). Both challenges are recognized gaps in the theoretical SoS literature and each has been identified as a key components of the SoS community's research agenda.

2.6 Progress towards an SoS research agenda

Several community attempts have been made to establish a research agenda for SoS. Keating et al. (2003) identify four research thrusts for system of systems engineering: new system design, existing system transformation, system operation and maintenance, and evaluation and evolution. From the prior cited literature, challenges can be seen in each of these areas:

New System Design The primary focus of new system design research has been on the optimization of single systems within a specified environment. SoS exist within a changing contextual environment and, by their very nature, span multiple complex systems. New design practices must be developed to account for these realities.

Existing System Transformation SoS are often composed or/derived from existing systems. This fact places additional constraints and challenges on the SoS designer. Not only are there technical constraints imposed by legacy systems, but, since there may be an active stakeholder community invested in the legacy systems, social constraints exist as well.

System Operation and Maintenance Decentralization of decision making greatly complicate operation and maintenance of both the SoS and the constituents.

Evaluation and Evolution What are the correct objectives to drive the design of SoS and their constituents? As described by Sage and Biemer (2007), constituent and SoS needs are changing over time and may not be synchronized. In some cases the constituents' needs will be long term while the SoS needs are short-term; in others the reversed. Who, among the various stakeholders involved, decides how these various objectives are to met on appropriate timescales?

The Air Force Scientific Advisory Board (SAB) commissioned a report in 2005 (AF/SAB, 2005b) that made several recommendations to improve the state of SoS acquisition and capability development within the Air Force and its contractor community. In terms of system architecture, i.e. better understanding how to form SoS from constituent system, they emphasized the role of standards, in particular *convergence protocols*, in creating opportunities for systems to work together. Given the large-scale of the SoS the Air Force often deploys, they recommended expanding the role of experimentation in system development so the novel SoS constructs can be tested at a smaller scale prior to fielding. Experimentation would also tighten the feedback loop between system development and use. This is consistent with the spiral strategy recommended by Boehm and Lane (2006).

The SAB also recognized the social aspect of the SoS problem recommending that fundamental changes need to be made in the acquisition infrastructure in order to enable a change in mindset from acquisition of a single system to the continuous development of SoS capabilities. They discuss how such capability development occurs in the commercial sphere by discussing the development of the network technologies such as the Internet. However, they recognize that a different model will be needed for the DoD given the unique single customer generated needs context of defense acquisition. To accomplish this change they recommend the creation of a new structure to allow full participation of all stakeholders in the system design process breaking down traditional barriers between government and its contractor base.

The most recent attempt at building a research agenda for SoS architecting was undertaken by a 13-member industry-academia committee sponsored by USC's Center for Systems and Software Engineering (CSSE) (Valerdi et al., 2008). First, speaking specifically to the reliance on a distributed network, they recognized that SoS create a new set of vulnerabilities as well as opportunities. They then identified several areas which should be the focus of future research in SoS architecture. In terms of architectural fundamentals, they reiterated the need for the establishment of a clear, well-grounded, set of constructs to understand the additional properties that SoS exhibit vs. systems developed via traditional systems engineering. As part of that effort, multi-view architectural frameworks need to be developed and put into practice so that many different domains (such as that examined by DeLaurentis and Callaway) can be understood and the links between domain-specific analyses revealed. Given the limit of human cognition, they suggest that model-based architecture will be a key tool in realizing such a multi-domain approach. Even with such models, predicting SoS behavior *a priori* will be difficult. The committee recommended therefore that a guided emergence model in which the SoS designer was guiding rather than explicitly controlling the emergent process of SoS development. Finally they, like the prior attempts at defining a research agenda, recognized that SoS often do not have a single owner, but are rather under the influence of many different stakeholder each with their own agenda.

Progress toward addressing the issues raised in SoS research agendas has been focused much more on the technical issues of system architecture that arise from combining constituent system than on the social issues that arise from interaction between stakeholder groups. On the former point, examples include investigation on the use of multi-disciplinary optimization Yan and Haimes (2007); Crossley and Nusawardhana (2004) and the development modeling and simulation techniques to evaluate SoS (Sloane et al., 2007b,a; Biltgen and Mavris, 2007; Dagli and Kilicay, 2007). Less progress has been made in addressing the multi-decision maker, or, as Valerdi et al. (2008) put it, the multi-owner aspect of SoS. While this issue is evident in discussion going back to Eisner et al., existing framework do not get much beyond creating a hierarchy of stakeholders,[10] just scratching the surface regarding their individual needs and interactions. The next chapter begins to fill this theoretical gap by first examining the distributed decision-making processes of several various real-world SoSs and laying the theoretical foundations for a generalizable model describing constituent system interactions.

[10] Boehm and Lane (2006) move well beyond a list by their use of the Boehm's win-win spiral model for multi-stakeholder development.

2.7 Descriptive, normative and prescriptive research

In formulating research questions, both the theoretical considerations identified in the works cited earlier in this chapter along with the practical challenges that arose in the case examples were taken into account. The focus of the thesis is on descriptive representation of decision making in SoS and prescriptive guidance for influencers on using influences to alter constituent decision making.

Consider the SoS influencer as going through three stages in transitioning the SoS from a current state to a desired state (see Figure 2-8).

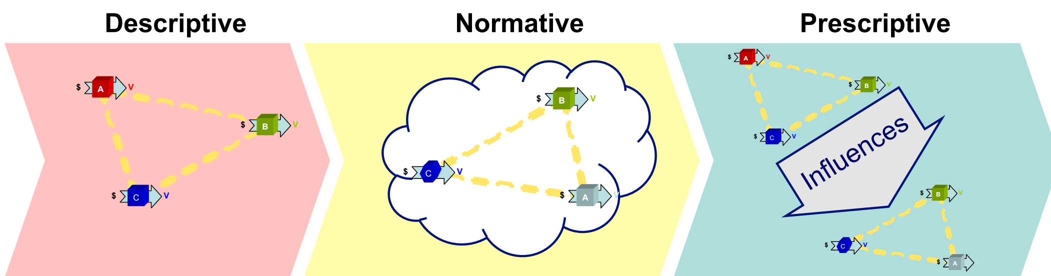


Figure 2-8: Three stages of SoS transformation

They must first be *descriptive* and characterize the current state of the SoS. This is the topic of the first research question and the intent of the AIR framework.

The influencer must then decide what the future state of the SoS should be (from their perspective). This is commonly referred to as finding a *normative* or desired state. This middle step is the purview of SoS architecting, i.e., attempting to define what constituents should do so as to create the desired SoS behavior. Several researchers are focused on this area and the reader is encouraged to consult the references cited at the end of section 2.6 that address the SoS architecture and modeling. Further discussion of the normative step is outside of this research's scope. It will be assumed that the influencer has determined the desired SoS state. This expression of desired state need not be a completely detailed list of instructions for the constituents, it could simply be guidelines, or 'aspirations' that are understood by constituents and knowledge which the constituents can incorporate into their decision making should the right influences be in place to induce them to do so. In fact, such an arrangement would take advantage of the distributed authority enabling emergence (Yang et al., 2010).

Finally, in the *prescriptive* step, they must find a way to transition from current to desired state. As was argued above, since the constituents are managerially independent, the influencer can use various influence mechanisms to cause the decision makers for each constituent system to make changes that collectively result in the desired SoS behavior.

2.8 Research questions

Given the above framing, two research questions, originally stated in section 1.3, are considered in this thesis. Focusing on the descriptive stage identified in the previous section, the first research question concerns the establishment of a new framework to characterize the decision making processes of an SoS. The framework should capture both the fact that there are multiple decision makers working at the constituent and SoS levels and the interaction between these decision makers.

> *What are the feedback relationships between the constituents and SoS influencers, and how do their influences result in changes in the constituents individually and the SoS as a whole?*

The second research question concerns the prescriptive stage. Working within the framework developed in the first question, a basic set of strategies to influence the behavior of other stakeholders is defined.

> *What approaches can be used by external SoS influencers to cause constituent decision makers to change constituent systems so as to induce a desired behavior from the SoS?*[11]

In addressing these questions, the scope of inquiry is limited to SoS with a fixed set of constituents who have fixed local value propositions.[12] Extensions to the variable constituent case are left to future research.

The next chapter lays a foundation for addressing these questions by examining decision making in SoS in more detail.

[11] This is not to imply that all SoS have influencers, rather, the question concerns techniques influencers should use when they are extent.

[12] The influencer may try to change these local value propositions, but they do not change otherwise.

Chapter 3

Distributed Decision Making in SoS

The study of decision making has a long history. Formal study can be traced back to von Neumann and Morgenstern (1953) who established the basic axioms of single attribute utility theory. This work was extended to multiple attributes by Keeney and Raiffa (1993). Extending this work to decision with multiple parties was the foundation of game theory. Within specific areas such as supply chain coordination (Schneeweiss and Zimmer, 2004) this work has been extended to the case when the multiple parties are distributed in time and/space. Such is the situation in SoS and so will serve as a starting point for examining decision making in SoS.

3.1 Perspectives on decision making in SoS

A closer examination of some of the frameworks introduced in section 2.4 reveals several shortcomings in the existing framing of the SoS problem as it relates to decision making. Sage's framework (Sage and Biemer, 2007) splits the tasks involved in managing an SoS into four categories: (1) enterprise activities, (2) development activities, (3) operational activities, and (4) technical activities. *Enterprise activities* refer to the contextual actions that the constituent systems and their stakeholders engage in to ensure that the SoS remains together and independence is maintained. This includes, for example, information sharing and collaboration between constituents, but excludes the actual operation of the constituent systems. *Development activities* concern the change in the SoS over time. *Operational activities* are related to the actual operations of the constituent systems and the SoS as a whole. Finally, *technical activities* refer to "traditional" systems engineering functions applied within the SoS context (Sage and Biemer, 2007). The key difference between this formulation and the traditional system engineering approach is the inclusion of 'enterprise activities' within the purview of the system engineer. The other three categories are concerned with the development and operation of the system itself, while enterprise activities affect the social structure within which the system is developed and operated. SoS engineering involves both social as well as technical elements. The source of this social element is the distribution of decision making amongst the stakeholders of the participating systems. For example, Sage describes the issue as follows:

> "Collaboration among programs and participating enterprises during SoS development is essential to the success of the SoS effort. This is especially the case when dealing with FoS efforts. How to achieve collaboration among programs not under one's control and potentially which one would not even wish to have under their control is challenging! Even with some amount of control, ensuring collaboration is problematic. Therefore, collaboration guidelines need to be established early. Furthermore, program "buy-in" is essential; therefore, developing a collaboration plan in conjunction with the programs is paramount. These issues are at the heart of F/SoS engineering program management." (Sage and Biemer, 2007)

The case examples in section 2.2 give real world credence to Sage's observations. Peering among the Internet service providers is an issue of choosing with which other systems one wishes to connect, i.e., with whom to collaborate. The DSP/Patriot SoS (subsection 2.2.2) does not have a participation issue per se, as it is a directed SoS in which participation can be compelled; however, there were limits to the authority of the SoS integrator. In establishing the communication link between the DSP satellites and the Patriot batteries, the existing mission of the DSP could not be compromised. Rather, protocols needed to be developed that augmented the existing capability. HousingMaps (subsection 2.2.3) did not initially require coordination of the systems, as Radamacher, who created the initial mashup, was taking advantage of existing interfaces. However, as time went on and the website (along with its clones and competitors) became more popular, there was a need for organized, as opposed to opportunistic, collaboration. This caused Google to impose terms of service that balanced the desires of those using the Maps as a service with Google's own need to control and monetize Maps. Bjelkemyr et al. (2007) identified the same issue for SoS more generally:

> "Each system within a SoS is a self-interested node in a network. These system nodes try to maximize their own utility under the influences of and in competition with the other nodes. The global SoS behavior thus emerges as a result of the actions at the lower levels of the SoS, down to the system element level." (Bjelkemyr et al., 2007)

In the traditional model of systems engineering, social consideration (i.e. issues of self-interest, competition, etc.) occur at lower levels within the system. Top-level decision making is done by a unitary decision maker who resolves the trade-offs in the design. As top-level tasks get decomposed into lower-level tasks, there is a natural distribution of decision making amongst the lower-level actors; however, there is a common guiding set of objectives specified at the top-level. This is the operating rationale behind the *waterfall* and *V* models for product development presented in traditional systems engineering resources (Unger, 2003).

In the SoS case, as was seen above, decision making is distributed even at the top-level. As such traditional systems engineering models that rely upon a consistent and unitary decision maker at the top-level are not well suited for SoS. However, this does not mean that the traditional approaches are to be abandoned completely. Each of the constituents is still a system and can be designed and managed using the traditional approaches as long as appropriate care is taken to account for the context changes that arise from being in an SoS (Shah et al., 2007a). Rather, SoS engineering can be seen as a complement to traditional SE needed to address the additional decision complexity arising from constituent interaction. As Chen describes it:

> "In order to cope with SoS challenges, SE practice must advance from its traditional context of projects into an organization context. This change declares a need for SE practice beyond the traditional scope defined in many classic SE processes, such as Waterfall, V Model, Spiral Model, Evolutionary Acquisitions, and Synchro X Model. The concepts and processes discussed in the paper should be considered complementary to traditional models and as a basis for an organization to plan and design its SoS SE practice." (Chen and Clothier, 2003)

The essential difference between the decision structure in traditional SE vs. SoSE is one of alignment. In the traditional SE case, since lower level requirements are derived from higher level requirements, there is a more natural path to alignment of the value propositions of those attempting to satisfy lower level requirements with those responsible for higher level requirements. In the SoS case such an alignment is less likely to exist as the constituents are independent. Instead, the SoS influencer (defined in section 1.2) may need to influence the constituents to behave in a manner that is not necessarily locally optimal for them but that does serve the interest of the SoS. This relationship between the SoS influencer and the constituent decision makers is a principal-agent problem (Binmore, 2007b) with the influencer as principal and the constituent decision makers as agents.

The principal-agent problem arises from game theory and describes the situation in which one party, the principal, has a payoff that is dependent upon the actions taken by another party, the agent. Each party makes choices in their own self-interest which may or may not be aligned. A classic example of this situation is the employer–labor relationship. In that case the employer, wishes to maximize the productive output of her firm. The output is dependent upon the effort put forth by the employees. The employees wish to maximize their total wages while minimizing work hours. Each player, the employer and the employees, makes their own choices with regards to the variables they control.

This same decision structure can be seen in the the case examples. For example, HousingMaps, involves Google as the 'principal' and the websites as 'agents'. Google's benefit from its map service is dependent on their ability to (1) get Internet users to make use of the service and (2) monetize that use. While some users will make use of the service directly, many others will use the service via a third-party website. Initially, Google did not have an API (application program interface—see subsection 2.2.3) and so third-parties such HousingMaps generated traffic, but did not generate revenue to cover the cost associated with that traffic. Once Google deployed an API the third-parties became agents for Google and entered into agreements that specified the relationship between Google and the third-party websites. To be effective, the agreements needed to provide benefit to both Google and the third-party website. From the website's perspective, Google provided a robust mapping capability that could be embedded within a larger service. As was demonstrated by HousingMaps, such a capability is of great value to website operators. Even better, Google was willing to offer the service at no direct cost. However, Google does require (Google, Inc., 2009) that websites who embed maps for free also allow the display of advertising connected to the map being shown. In this way, the benefit to Google is ensured. Mediating this relationship between Google and the websites is the Google Map API. It specifies how Google wishes websites to interact with Maps to ensure that both the websites and Google benefit. By providing the API, Google creates an infrastructure by which they manage the SoS.

As was demonstrated by HousingMaps, website developers do not need to use the API; they can access the service directly. However, doing so exposes them to the possibility that Google could change the service in an unexpected way. By using the API, website authors have a reasonable guarantee that Google will not change the interface. In addition, they benefit from programmatic access Google provides to Maps feature they may find useful. This creates an incentive for the website authors to use Google Maps via the API as opposed to using another map service or bypassing the API. It also ensures that Google's ability to place ads and use data collected from map related searches that occur through websites is maintained. Prior to the introduction of the API, the interaction between Google and the websites was ad-hoc or, to use Maier's terminology, a virtual SoS was formed. By introducing the API, Google intervened in this SoS changing it into a collaborative SoS governed by the terms of the API agreement which imposes conditions on both the websites and Google ensuring mutual benefit to both parties. Google's introduction of the API is one example of an intervention by which an influencer changes the behavior of a constituent to effect the overall SoS. The focus of the remainder of this chapter and the next is more formally characterizing the relationship between influencers and constituents within an SoS. Then that framework is used to identify various intervention strategies that an influencer could use to change the behavior of its agents and thereby the SoS.

3.2 System interaction and constituent interaction

The essential features that have been observed in the examples presented thus far (and also recognized in the cited literature in chapter 2, e.g. Table 2-3) are listed below. Observations 1-4 can be directly found in the existing literature, while 5 and 6 are being introduced in this thesis based upon the earlier discussion and case examples.

1. Systems-of-systems are composed of independent systems known as constituents (Krygiel, 1999).
2. Constituent systems are independently operated and managed by entities who are often different from those operating and managing the SoS (Maier, 1999).
3. Being independent, these entities have their own objectives and decisions to make in meeting those objectives (Bjelkemyr et al., 2007).
4. The particular decision made by constituents (along with choices taken by the SoS managing entity) determine SoS behavior (DiMario et al., 2009; Karcanias and Hessami, 2011a).
5. As the SoS influencer does not control the constituents, they are in a principal-agent relationship.
6. Within the principal-agent relationship, the influencer (i.e. principal) uses influences to change how the constituents (i.e. agents) value the outcomes of their choices with the intent of enacting change in constituent system behavior to maximize value for the influencer.

In discussing the proposed framework to represent decision making in SoS, two simplifying assumptions are made to allow for clearer explanation of the framework.

1. There is only one SoS influencer. There are multiple constituents.
2. The constituents are distinct from the influencer; that is, no component of the SoS is under direct control of the influencer. This ensures that there is a pure principal-agent relationship.

To begin this decision-making centric view of systems-of-systems, the basic element of the SoS, the system, is represented using the labeled box shown in Figure 3-1. The particular symbolic idioms used in representing a system are meaningful. From the SoS's perspective the constituent system's internal operations are a black box. Thus the box represents the boundary of control between the constituent system and the SoS. Entering the box are resources represented by a dollar sign. These resources are not only financial, but also time, personnel and material. The system transforms resources so as to create value[1] for the decision-making stakeholder who controls the system[2]. This transformation is represented by the arrow exiting the box labeled 'V'. In isolation and with a fixed context, the decision problem for the system can be thought of as a value maximization problem

[1] *Value* here means a decision maker preference on the result of the system acting upon the resources provided.

[2] Here it is implicitly assumed that a single decision-making entity constructs, manages, provides resources for and derives value from the system.

in which the system decision maker is trying to find the combination of resource expenditures that maximize the value delivered to them subject to the constraints imposed by the context. They preform the optimization dynamically, adjusting resource usage as the context changes. Of course, the success of a decision maker in performing this optimization is dependent upon the underlying complexity of system and context however, the decision maker's intent is still captured by the optimization.

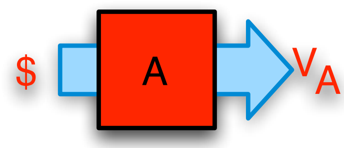


FIGURE 3-1: A system in isolation that acts upon resources producing value for its decision maker

Moving from systems engineering to system-of-systems engineering, additional systems are added to the diagram (Figure 3-2). These systems are labeled with different letters from the first showing that they are managed by different decision makers. Connecting the systems are dashed lines that represent interactions between the systems. Interaction can be either intentional or unintentional. Unintentional interactions come about, for example, when two systems use the same limited resources. A simple example of this is a shared communication channel in which the system using the channel at a given time must announce its intention to do so to prevent interfering with another user (e.g. several people using radios to communicate). Intentional interactions or interfaces arise when there is a mutual benefit realized by each systems participating in the interface. In either case, once multiple systems are interacting directly, they form an SoS.

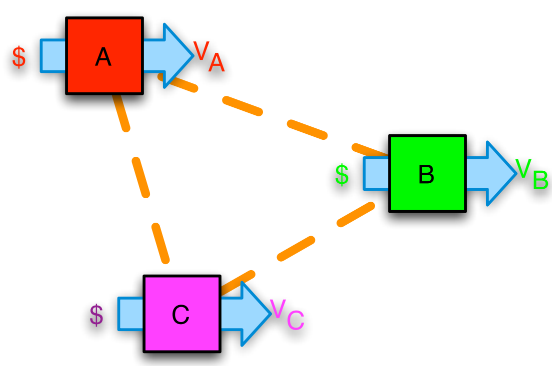


FIGURE 3-2: Three systems interacting via both intended (interfaces) and unintended (interactions) means

The particular class[3] of SoS formed depends upon the awareness of the constituents as to the existence and effect of an external coordinating actor. The simplest case is that of the virtual SoS. Here, the constituents are unaware of the existence of the SoS around them. At first glance, this seems an unlikely state. Why would any system manager be unaware of his/her interfaces with other systems? Maier (1999) uses the example of a market economy for a virtual SoS. In a pure market society firms may only be aware of the effect of other firms with which they are closely connected. The further two firms are apart in the economy (i.e. the more layers of exchange between the transactions of one firm with another), the less likely they are to take into account the other firm in their decision making. Given a sufficient number of firms with many interactions, it becomes impossible for any single firm to comprehend the whole economy at the same level of details they perceive their direct partners. Firms need to use abstractions to gain a limited understanding of the economy as a whole and as a consequence will miss details that may be important.

Even with small numbers of firms involved, all interactions may not be readily apparent. As was described in the HousingMaps example, Google and Craigslist were not initially aware of the interactions that Radamacher was creating between their systems. More precisely, they did see his use of their service via requests to their website, but did not comprehend that such use was part of a higher level tool that combined their respective service offerings. As such, Google and Craigslist were participating in a virtual SoS. The reason they did not notice is that Radamacher was taking advantage of existing interfaces as well as reverse engineering internal interfaces that were not available for public use.

SoS change over time. A source of those changes are decisions being made locally by constituents in response to changes in the context and the actions of other constituents. Constituent decision makers may also also interact with each other directly. This distinction is made to separate interaction arising from an interface between systems from interaction arising from inter-organizational contact. To represent this source of inter-organizational interaction, an additional network is added to the diagram showing constituent decision makers and their interaction with each other (Figure 3-3). This network is connected to the network of interacting systems via constituent actions that cause changes in the constituent systems. This results in time varying network of systems comprising the SoS. The constituent systems could change, new systems could be introduced/old ones removed. Interfaces may change as well. Any of these or other actions are taken by constituents are in response to changes that they observe in the constituent systems and feedback from the actions they take. The distinction between "observation" and "post-facto feedback" is defined by the particular constituent action being considered. "Observation" refers to information gained about the SoS prior to a given action being taken, while "post-facto feedback" refers to the results of the action as

[3]Class refers to the classification described in section 2.3 based upon Maier (1999); Dahmann and Baldwin (2008), i.e., directed, acknowledge, virtual and collaborative.

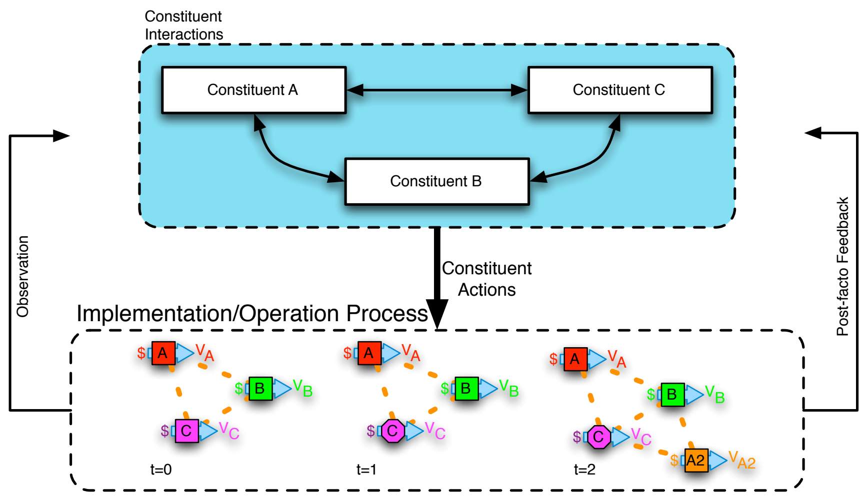


FIGURE 3-3: Systems and interactions change over time as the decision-making stakeholder (i.e. constituent decision makers) change their systems in response to each-other as well as changes in the context.

perceived by the constituents. One action's "post-facto feedback" can be another actions "observation". The two flows are separated since constituents do not necessarily act synchronously. Those constituents not involved in an action merely observe the changed caused by the action, while those who are acting look for specific feedback as a result of their actions.

An example of the interactions described above can be seen in the Internet Peering dispute introduced in subsection 2.2.1. Level 3 and Cogent interacted as constituents and signed a contract to peer, i.e., share traffic between their networks on a no-cost basis. This is an agreement between constituent decision maker and is represented by the arrows connecting constituents in the top portion of Figure 3-3. As a result of this agreement, physical connections (i.e. cables at peering centers) were made between the two networks. Making the changes in their respective networks are constituent actions that introduce an interface between the systems they operate. Level 3 then observed (the observation arrow) that the traffic wasn't flowing symmetrically between their network and Cogent's. The degree of asymmetry (which was in dispute) was enough to cause Level 3 to take an action, i.e., to effectively cut the interface cable between the two network. They then observed the effect of that change on both their and Cogent's customers and saw a need to resolve the dispute quickly[4]. Once the agreements were clarified/refined, Level 3 acted again to re-establish the peering connection between the networks (Goldman, 2005).

[4]Whether or not Level 3 knew the full ramifications of ending the peering agreement cannot be discerned from public reports and press coverage.

The structure in Figure 3-3 captures the key interaction that occur when multiple systems interconnect to form virtual SoS. There are two interaction networks, one between the systems and one between the constituent decision makers. These two networks themselves interact via a feedback structure of constituent action changing the systems based upon observed system behavior and the systems providing feedback on the effect of the constituent actions. Missing is any coordinating entity that desires particular behavior out of the SoS as a whole. Constituents decision making is driven by local concerns and when interfaces between systems are formed, they are only done so in support of these local agendas. This is the virtual SoS described by Maier. Introducing a coordinating entity, i.e., a principal in the principal-agent problem described earlier, leads to the other SoS types described in Maier (1999) and in Dahmann and Baldwin (2008).

The constructs developed in this chapter set the stage for the next. This chapter described the basic interactions that occur between constituents that give rise to SoS behavior. It was observed that these interaction are both social and technical in nature. The next chapter, building upon this foundation, adds the SoS influencer who wishes to change constituent behavior to obtain a desired SoS.

Chapter 4

Anticipation, Influence and Reaction

The previous chapter described the distributed nature of decision making in SoS. The role of the SoS Influencer (or principal) is one of inducing constituent action to generate SoS behavior that the influencer desires via influencing the constituents. This type of relationship is hardly new to the field of decision theory or organizational management. In logistics, for example, the problem is quite commonplace. A manufacturer who sources raw materials from multiple vendors and then sells to multiple retailers is managing an SoS. Each of the components in the supply chain is a constituent that is maximizing its local value, while the manufacturer, in bringing together the chain, is trying to extract value from the SoS as a whole.

Schneeweiss (2003) extends the work done in logistics to more generic distributed decision-making problems in organizations. The current work applies and extends his formulation to the SoS. While Schneeweiss's focus was on the inter-organizational relationships, the current research also includes the connection between organizations and the systems they control. As a consequence, a broader selection of influence mechanisms are considered. While Schneeweiss looks at direct incentives and information, the current research extends that to include technological and institutional mechanisms as well. To enable this extension, building upon Schneeweiss's work, a new framework, AIR, for capturing the feedback relationships between key actors in an SoS is proposed. AIR stands for the three phases that occur when an SoS influencer attempts to change an SoS—*Anticipation*, *Influence* and *Reaction*. Note that AIR only focuses on the distributed decision making aspect of the SoS, i.e., interaction between constituents and an influencer. There may also be component systems and subsystems within the SoS. While these may support the operation of the SoS as a whole, providing infrastructure for example, they are not explicitly represented in AIR.

In this chapter AIR is introduced by considering the process by which an SoS influencer might effect change in SoS given that they do not have direct control of the (entire) SoS. Then, to demonstrate the use of the AIR framework, it is applied to each of the four canonical SoS types, directed, virtual, acknowledged and collaborative (Dahmann and Baldwin, 2008; Maier, 1999). With each type, real-world SoS are used as examples. These examples are then used to argue that the major difference between the four canonical types are the feedback relationship between constituents and influencers as captured in AIR.

4.1 The AIR framework

Three terms, constituent, constituent system and decision maker play key role in the discussion to follow (see section 1.2 for definitions). Whether or not to consider the decision maker[1] who control the constituent system as within or outside the boundary of the constituent systems depends on one's view of the boundary of the constituent system. For the purposes of introducing AIR it is useful to treat the decision maker as being outside the constituent system, i.e., treat the system and its decision maker as separate entities. This is because AIR attempts to capture the various feedback mechanisms that lead to changes in an SoS via changes in its constituent systems. By separating the decision maker from the constituent system, the change agent (the decision maker) and the object being changed (the constituent system) can be explicitly identified. This partition is more intuitive in cases where the constituents are pieces of technology such as a computer or a vehicle. As one considers constituent systems in which the decision maker plays a more active operational role it is helpful to think instead in terms of the state of the constituent system. The decision maker specifies a certain set of variables that define the state of the constituent. As long as those variables remain fixed, it is still the 'same' constituent system. It may change in other ways during its operation (including changes made by the decision maker), but, from the decision maker's perspective, the constituent has stayed the same. Sometimes, for the sake of brevity, a constituent is said to 'change'. This should be taken to mean that a constituent decision maker has changed its constituent system as it is only the constituent decision maker who control and can therefore change the parameters of the constituent systems.

To generate these influences (see Figure 4-1 for a visual treatment of the processes being described), the influencer first observes current SoS behavior. This observation is used by the influencer to capture the current state of the SoS and evaluate direct changes they could make to SoS entities under their direct control (i.e. SoS component systems and SoS subsystems).

Second, they anticipate constituent decision-making and interactions. The word 'anticipate' is used instead of 'observe' since, unlike system behavior, constituent decision-making process is not generally visible to the influencer. As independent agents, constituents can make choices without

[1]Decision maker is taken to be a singular entity here. In reality, it may be an organization that has many individuals within it. That additional complexity is not addressed here to limit scope.

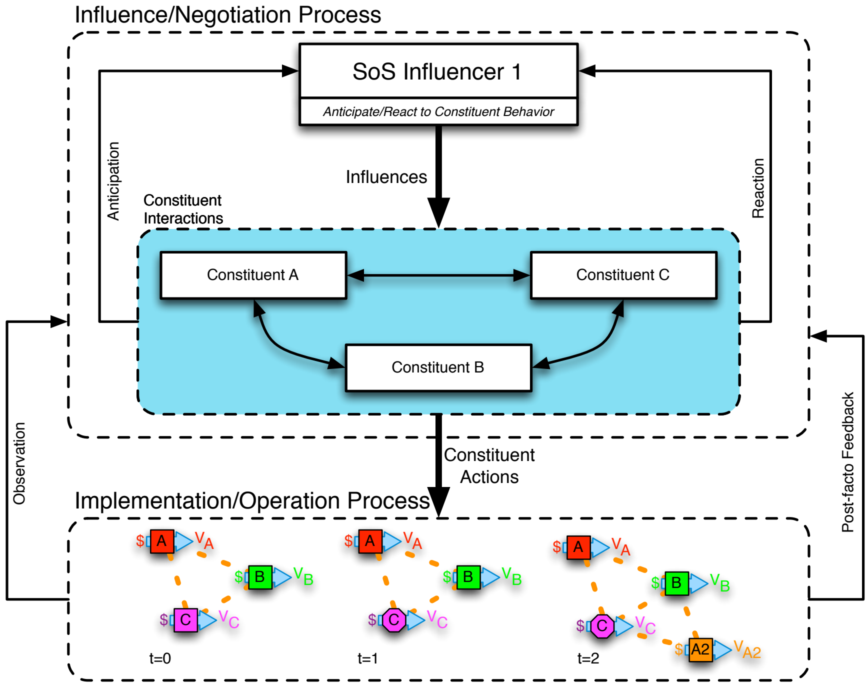


FIGURE 4-1: An SoS influencer influences constituents to take action that modify SoS behavior such that the influencer's objective are met.

transparency only revealing them through their actions. Therefore the influencer must use their best estimate of constituent action in assessing influence strategies. Third, based upon observation of the systems and anticipation of constituent decision-making, influences are brought to bear upon the constituent with the aim of modifying their behavior.

Constituents respond to these influences in two ways. First, they take actions to modify the systems they control in response to the changes in their decision problem created by the influencer. If the influences were well-formed and the influencer's understanding of the system and constituents accurate, then the effect of those changes in the systems will modify SoS behavior in an manner of strategic value to the influencer. The influencer will observe the extent to which this has occurred (post-facto feedback). The constituents may also react directly to the influences, signaling their (dis-)satisfaction.

These three interactions, *anticipation, influence and reaction*, form the core social feedback mechanism between SoS influencers and their constituents, and give the AIR framework its name.

The next several sections will represent common SoS types using the AIR framework as well as map real-world examples of each type.

4.2 SoS as represented using the AIR framework

AIR highlights the major differences between the SoS types defined in Maier (1999); Dahmann and Baldwin (2008). In the next several sections it is shown that the key difference between these types is the decision making relationship between the constituents and the SoS influencer (should it exist).

4.2.1 Directed systems of systems

Directed SoS as defined in Maier (1999) are SoS in which managerial control is ceded to a central authority. Since this central authority has coercive power over the constituents there is no direct feedback between constituents and the SoS influencer. Furthermore, the influencer can instruct (as opposed to merely influence) constituent decision making. The resulting modified AIR diagram is shown in Figure 4-2.

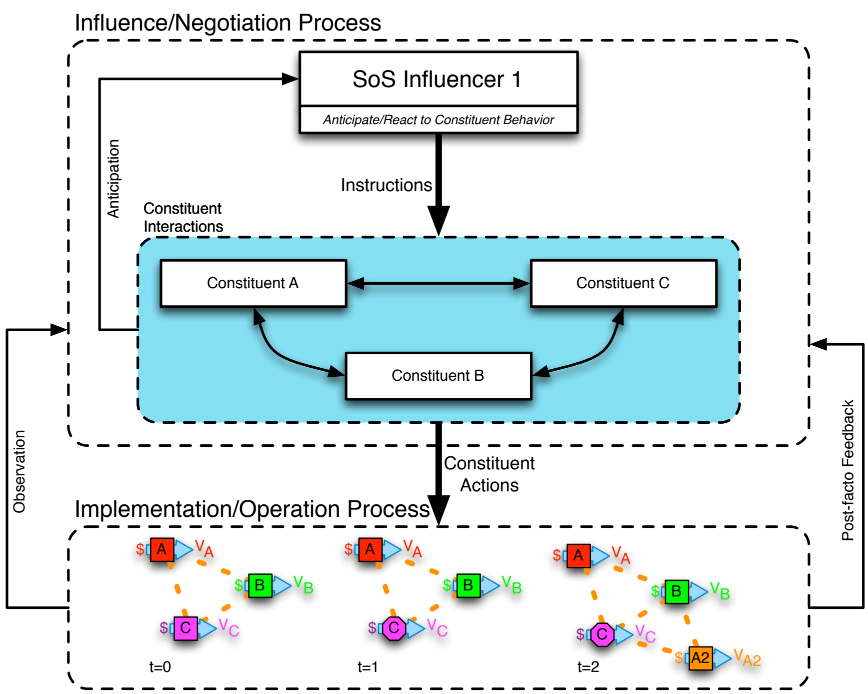


FIGURE 4-2: In a directed SoS the influencer has coercive authority over the constituents. Constituents however, retain identity and direct control over their respective systems. When compared to Figure 4-1, the reaction arrow has been removed and influences become instructions.

Such control over the constituents does not imply that the constituents are no longer independent. On the contrary, operational and managerial independence still hold; the constituent systems retain identity and are under the control of their respective constituent decision makers. Maier (1999) uses the example of an integrated air defense system in defining a directed SoS. In that example, the various constituent systems such as aircraft, radar installations, and maritime surveillance assets continue to exist as systems in their own right while being directed by the central authority. It is important to not confuse direction by the SoS influencer with the constituent systems becoming subsystems of the SoS. While superficially it may appear that a subsystem of a directed SoS is the same as a constituent system, they are in fact quite different. The decision of how much authority is retained by the influencer vs. being granted to the constituents is a key design consideration in managing directed SoS. In addition, operational independence implies that the constituent systems are separable from the SoS, while subsystems are not.

Consider Maier's example of integrated air defense. The individual components, while performing missions as directed by the SoS influencer, still retain autonomy with regard to certain local functions such as self-defense. For the influencer, with authority often comes responsibility. For example, an SoS such as an integrated air defense system will require a communication network to allow coordination of the constituent systems. Depending on the nature of the SoS it may be worthwhile to treat such a network as a constituent system rather than a subsystem of the SoS. If, for example, the SoS is short-lived relative to the constituents, piggy-backing on an existing commercial network could have significant cost savings. Of course, such savings would need to be traded against the more limited authority the influencer would have on the co-opted network when compared to network designated for its sole use. Finally, constituents may participate in multiple SoS simultaneously. In the air defense case, for example, the same radar may be used for monitoring both civilian and military aircraft. Such a situation can be represented in AIR by introducing multiple SoS influencers.

4.2.2 Acknowledged systems of systems

In examining a variety of the DoD SoS, Dahmann and Baldwin (2008) realized that many did not fit the definition of either a collaborative or a directed SoS. These SoS, referred to as acknowledged, had, as in a directed SoS, an externally specified objective with an influencer who wished to see that objective satisfied. However, the distributed nature of funding and programmatic authority in defense acquisitions implied that the SoS influencer did not have coercive authority as would be required for a directed SoS. Rather, the management of the individual programs needed to collaborate in order to accomplish the mutually acknowledged external objective, all the while maintaining their normal operations as managerially independent systems. In Dahmann and Baldwin's words:

> "Acknowledged SoS have recognized objectives, a designated manager and resources for the SoS, however, the constituent systems retain their independent ownership, objectives, funding and development and sustainment approaches. Changes in

> the systems are based on collaboration between the SoS and the system." (Dahmann and Baldwin, 2008)

To represent such SoS using the AIR framework, several changes must be made when compared to the directed case. First, the lack of coercive authority means that the influencer can only influence as opposed to instruct the constituents. Furthermore, since the constituents retain and act upon both locally specified and SoS influencer specified objectives, they can react to those influences either directly or through modifying their local system. Constituents may chose to reject or otherwise hinder the accomplishment of the SoS objectives.

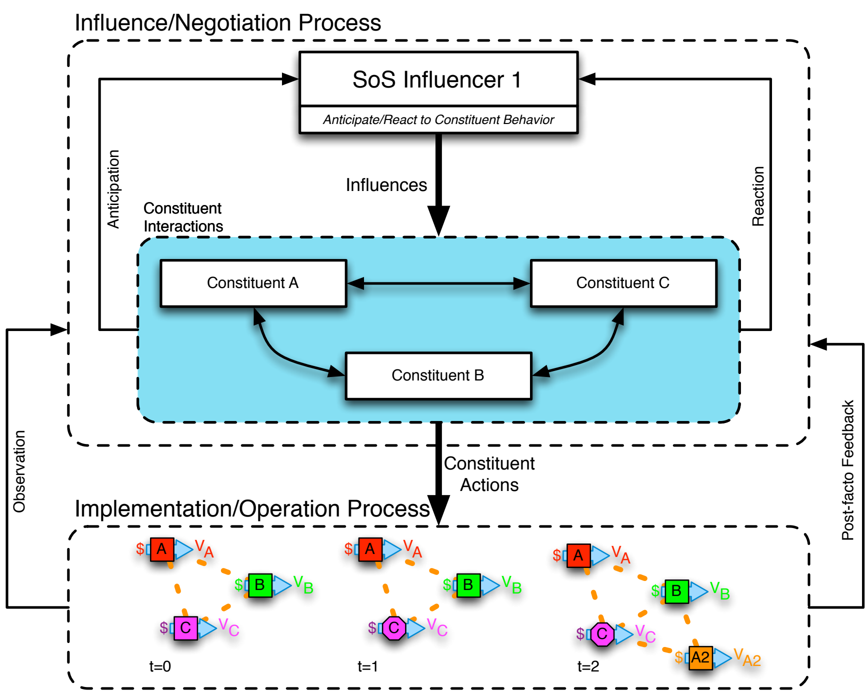


FIGURE 4-3: In an acknowledged SoS, the influencer has a limited ability to influence the constituents. Constituents retain autonomy and can react to influences imposed by the influencer.

Krygiel's (1999) case study of the development of the Defense Mapping Agency's (DMA) Digital Production System (DPS) is a good example of the issues that can arise when integrating an acknowledged SoS. The DMA is part of the Department of Defense and is responsible for producing timely and accurate Geo-Spatial representations (i.e. physical and electronic maps and associated documents) of areas of interest to customers within the DoD. Traditionally, it carried out this mission via several largely independent systems. Having its in roots in film and physical map production, the need for change in the production process was recognized as more digital technologies became available in the 1980s. In 1982 the DMA was directed to re-architect and integrate its production process into a digital workflow with the aim of substantially improving the timeliness and quality of imagery and mapping products. However, organizationally, DMA's production was not a single, unified program. Legacy production processes had been developed around the formats of the various source materials (film, maps, etc.) and was thus spread out over several acquisition program offices. Each of these programs had their objectives and would need to keep production going as the transition to the integrated digital system occurred. After 10 years of effort, the transition to a digital production system was complete.

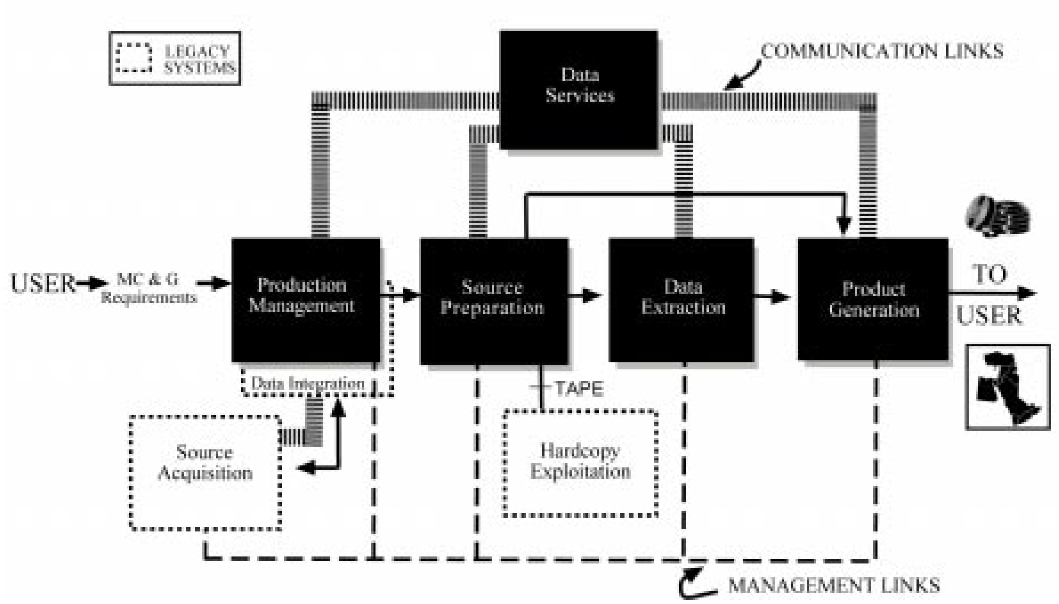


FIGURE 4-4: The constituent systems and interfaces of the DMA's digital production system. User needs enter the flow at the left and products exit at the right. (Figure 3.1 in Krygiel, 1999, p. 61)

Krygiel identified several key success factors that were needed to make the transition. First, organizationally, getting all the players to think in terms of the SoS and not their home programs was a challenge for the DPS integration team. To create this change mindset among the constituents, an Activation Control Team was established at each of production facilities to manage the transition. Each constituent had membership on this team which was led by a senior officer and a senior engineer who spoke for the end-user and the SoS as a whole. By placing these teams at the production site (i.e. at the actual point of integration of the SoS) the SoS vision was kept at the forefront in the day-to-day decision making during integration. Note that these teams did not have coercive authority, i.e., making the DPS a directed SoS. However, they did have high visibility in the constituent organizations and the explicit support of higher level leadership which significantly increased their ability to influence the decisions made by the still managerially independent constituents.

Second, despite several years of planning and the development of a prototype system in 1985, creation and proper management of the new digital interfaces proved challenging. An on-site engineering review board was established to re-architect and allocate functions to the various constituent systems. The board was comprised of engineers independent of the constituent programs. They focused solely on understanding the SoS as a whole. This position of objectivity (with the support of the DMA leadership) empowered the engineering board to request needed changes in the constituent systems.

Third, the constituent systems did not have common reporting/management systems and engineering procedures. Beyond the practical benefits, unifying the reporting/management systems helped re-orient the identity of the DPS as an SoS as opposed to a collection of individual systems with arms length interfaces.

The integration of the DPS reinforces the need for the SoS influencer to be fully cognizant of their constituents. Simply specifying a common objective was insufficient to foster SoS integration. The constituent needed to be involved directly in the integration process in a manner that respected their independence and expertise in the particular services they provided as part of the production flow.

The story of the DPS transition into an SoS can be understood through the AIR framework. Initial attempts by the SoS influencer, i.e., the DMA, had poor anticipation of the complexity/ability of the constituents to adopt the new integrated, SoS architecture. This was revealed to the DMA via reaction in the form of poor adoption of the new system and post-facto feedback as seen in the non-standard implementation when adoption did occur. Observing this reaction, the DMA changed the mechanism by which they were influencing the constituents introducing the on-site integration teams. These teams had much better visibility of the transition challenges of constituents, i.e., better anticipation, and were thus able to successfully transition to the new architecture.

4.2.3 Virtual systems of systems

Maier (1999) defines virtual SoS as having no central coordinating authority at all—neither externally imposed nor emergent from constituent interaction. The SoS influencer is eliminated entirely. Constituents interact and form interfaces, but no one is fully aware of the whole. In terms of the AIR framework, Figure 4-5 represents a virtual SoS (repeat of Figure 3-3). Note that there is no influencer; yet constituents still form interfaces opportunistically.

As discussed at the beginning of chapter 3, HousingMaps prior to Google's introduction of the Maps API is an example of a virtual SoS. HousingMaps was simply taking advantage of undocumented features in both Craigslist and Google's service offerings to create a new application for their data. Once Google introduced the API, they imposed a set of restriction upon the users of

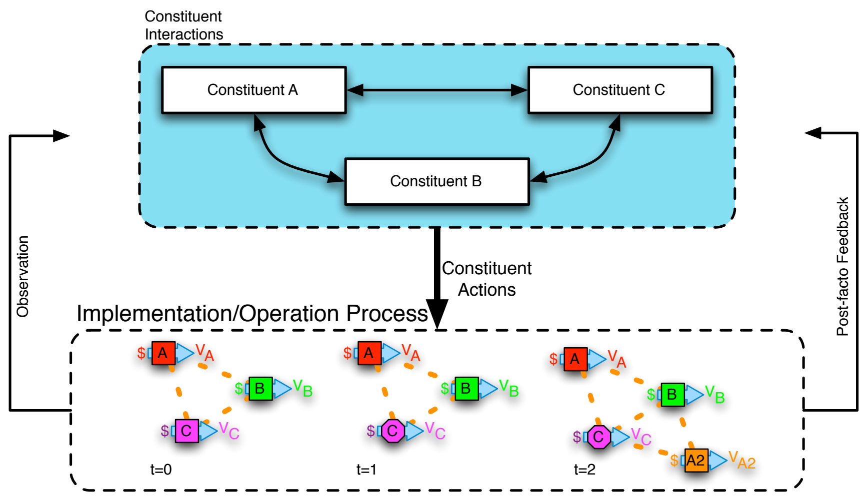


FIGURE 4-5: A virtual SoS in which there in no SoS influencer. (Repeat of Figure 3-3)

the API to influence them to use the Maps service in a manner beneficial to both Google and the API users. In doing so, Google acted as an SoS influencer changing the system from a virtual to a collaborative SoS.

4.2.4 Collaborative systems of systems

The last class of SoS defined by Maier (1999) is collaborative SoS. In a collaborative SoS, the SoS objectives emerge from a multi-party agreement of the constituents.[2] The objectives are specified and known to the constituents who then work together to meet both their local objectives and collective, SoS objectives. In this role, the constituents are now also influencers, using their influence on the other constituents to steer SoS development. As shown using the AIR framework in Figure 4-6, there may be several such constituent/influencers (not all constituents necessarily take on an influencer role). Note that there may also be external princpals as in acknowledged SoS; the difference here in that constituents are actively participating in the establishment of SoS objectives.

Continuing the HousingMaps example (see subsection 2.2.3), once Google instituted the Maps API they brought out into the open the virtual SoS.[3] They established a collective agreement on SoS objectives by codifying each party's responsibilities in their terms of service. However, to ensure that the new web services that incorporated Google Maps functionality were viable in their own right, Google had to work with fellow constituents such as HousingMaps to develop an API that worked well for both parties. To ensure that the API was successfully adopted, Google engaged

[2]It need not be all constituents. Those not party to the agreement can still be influenced as in an acknowledged SoS.

[3]To be precise, the website operators making use of maps as prior to the introduction of the API did know what they were doing, i.e., they weren't *unaware* of the larger virtual SoS in which they were participating; however, they did not take on an influencing role and so the SoS was classified as virtual since it lacked at influencer.

their developer community in beta testing and requesting feedback on new feature. This dialog also extends beyond the technical to include the changes in the terms and conditions of use for the API which are announced and discussed on the company blog. The websites have their own agendas with regards to the collaboration and do try to influence Google to act in their best interests by open sourcing key software code and supporting competitors such Yahoo! Maps, Bing Maps and the OpenLayers project.

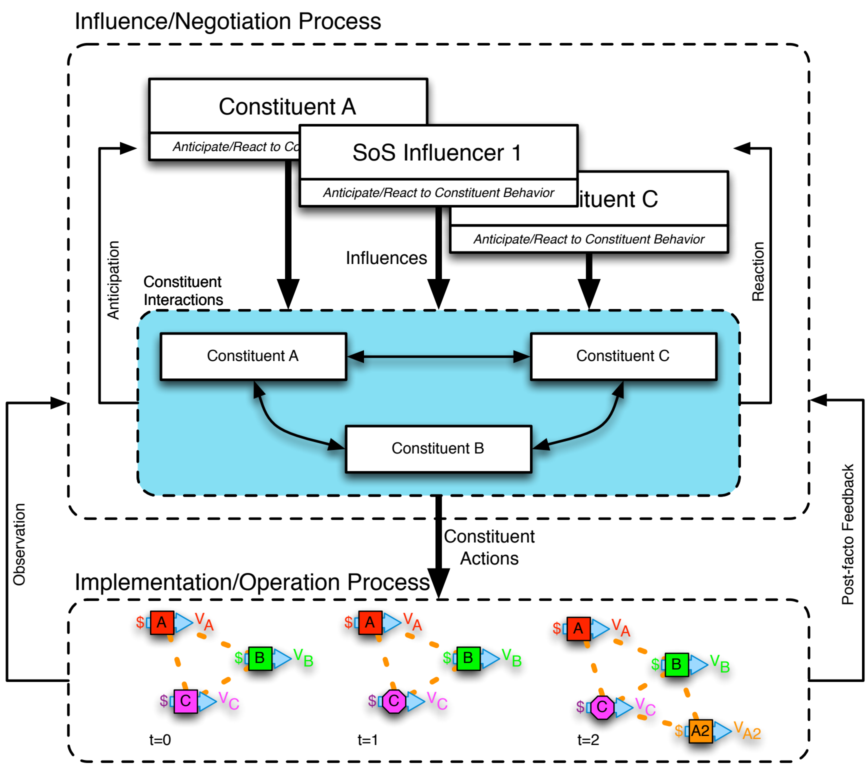


FIGURE 4-6: A collaborative SoS in which the constituents participate in setting SoS objectives and attempt to influence each other to changes their respective systems to create the desired SoS behavior.

A more extensive example is GEOSS, the Global Earth Observation System of Systems introduced in subsection 2.2.4. In trying to resolve the challenges of data interoperability in their SoS, the GEO (the Group on Earth Observation) developed a service-oriented architecture in a phased manner that was cognizant of the independence of the data providers. In this SoS, the constituents are the data providers. These same providers formed a working group that serves as the SoS influencer.

As described in Khalsa et al. (2009), the data interoperability pilot program proceeded in phases. These phases can be described in terms of observation, anticipation, influence and reaction in the AIR framework. In the first phase, recognizing, that in many cases, constituents were already exchanging data, an effort was made to document the *de facto* standards under which these exchanges took place. This is labeled *observation* on the far right side of Figure 4-6. In the next phase, communities of potential users were formed to examine what new data exchange/normalization requirements needed to be developed to harmonize the *de facto* standards from the first phase. This is the *anticipation* step within which the influencer attempts to anticipate how the constituents will respond to various influences and thereby find the influences that best induce the desired behavior of the SoS as a whole. In this case, the working group identified the new standards and protocols needed to enable the desired use-cases of the GEOSS members. The *influence* in this case is the offering of these new standards for adoption by the constituents. In the reaction phase, they will implement a demonstration version of the new data exchange service repository, thereby creating a opportunity for the users to try the new approach before broader deployment. This is a form of *reaction*, wherein private information to the constituents, i.e., the effect of the new standards upon them, is revealed to the influencer through the constituents' participation in the demonstration.

4.2.5 Unification of SoS classes

The previous four sections have demonstrated that the classes of SoS identified in the literature, directed, virtual, acknowledged and collaborative can all be represented as variations of the AIR framework. Note how each of the Figures 4-2, 4-3, 4-5 and 4-6 are variations of the basic AIR framework shown in Figure 4-1 formed by removing or adding aspects of the framework to meet the definition of the given SoS class. The examples discussed with each class also demonstrate that real-world SoS may not fit nicely into the 4-part categorization found in both the academic literature and in the guidance produced by the DoD. For example, integrated air defense is cited by Maier as an example of a directed SoS. In the short term, on the operational time-scale of a fielded defense force, this is true in that authority is ceded by the constituent to the integrated command. However, taking a longer time span, one realizes that, like the constituents in the DMA's transition to digital images, the constituents within the integrated air defense system belong to different programs and suffered in the past from similar coordination challenges as were seen in the the DMA case. In developing the Theater Battle Management Core System (Collens and Krause, 2005), a software systems used to manage multi-unit joint operations, there was a great of deal of difficulty developing requirements as well as resolving technical interfaces. When viewed over this longer time horizon, the air defense system seems more like an acknowledged SoS than a directed SoS. The HousingMaps case demonstrates that the class of SoS can change over time (e.g. from virtual to collaborative).

However, across all the examples, there is a common structure of an SoS having constituents and, except for virtual SoS, influencer(s) that interact. As demonstrated by the examples in this chapter, the AIR framework provides a flexible representation of decision making in SoS to include all four classes from the literature. The feedback relationships identified in AIR provide a way to differentiate between the classes.

Considering all four classes, there seems to be a continuum of the degree of control that the influencer has on the constituents. At one extreme is directed SoS in which the influencer has total control over the constituents. Next is acknowledged, in which the constituents can react to influences and thus requires the influencer to be more cognizant of the needs of the constituents to find solutions that they are amenable to. In collaborative SoS, there are multiple influencers and so a given influencer may now be competing with others in steering the direction of the SoS. This may result in a reduction in the effectiveness of any given influencers actions. Finally, at the other extreme are virtual SoS with fully independent the constituents that act in their self-interest without coercion or influence from outside actors. It would appear that there are at least two dimensions upon which this continuum could be defined. First is the degree of control that a given influencer can exercise ranging from a weak influence to directed instruction. Some initial work on quantifying this dimension as continuous variable is documented in (Chattopadhyay et al., 2008) by considering the likelihood of participation of a constituent as a combination of managerial control and influence. The second dimension is the number of influencers ranging from none to many. Full development of this potential continuum is left to future research; however, the AIR-based descriptions in this chapter do support the exploration of such a construct.

4.3 AIR as a dynamic framework

The role of time in SoS is complex as each of the involved actors may consider their part of SoS as operating on a different time scale. Overtime, the SoS will go through many cycles of anticipation, influence and reaction. The timing of these cycles will vary depending upon the nature of the underlying contextual driver that create the need to change the SoS. This timing need not be consistent over the history of the SoS. For example, if the SoS were to be attacked this might require multiple rapid reconfigurations that influence constituent to make rapid changes to ensure SoS viability. Once the attack is over, the influencer may take lessons learned from the experience and use them to influence constituent to alter their longer-term strategic investments so as to better cope with future attacks. These investment might bear fruit on a timescale orders of magnitude longer that the attack that caused them. The timing of influence need not be the same across all the constituents. Some constituents that play a well-established infrastructural role may only require occasional adjustment to the influence strategy, while those that, for example, incorporate rapidly changing technology may require more regular attention to keep their needs met.

Given such variability in timing driven by variability in context, a robust way of representing context changes will be required. This needs to occur for both the SoS as a whole and for each of the constituents. One promising approach is Epoch-Era Analysis (Ross and Rhodes, 2008). In Epoch-

Era analysis relevant variable that define a systems context are specified and then assigned different sets of possible value that correspond to possible future states of the context. The parametrized future states over which context remains constant are termed epochs. Eras are formed when multiple epochs are considered in sequence as in a time-line. If extended to include the interaction between constituent and SoS contexts (Shah et al., 2007a), Epoch-Era analysis could serve as useful starting point for understanding the timing complexity described above. However, such an extension will be non-trivial to implement as actions by the influencer within an epoch (as seen from their perspective) may change the context as perceived by a constituent bringing the constituent into a new epoch while still keep the influencer in the same epoch. Finding appropriate ways to incorporate this complexity into AIR is an important area for future research.

4.4 Implementation challenges for AIR

Actual implementation of AIR in real SoS poses several challenges. These are described working from left to right in Figure 4-1.

As explained earlier, observation consists of characterizing the current state of the SoS. Such characterization includes identifying the constituents, interfaces and resultant behaviors. In addition, the context within which the constituents are operating and the SoS context must be defined. This requires a minimum level of visibility of the SoS. If characterization of the current state is poor then the ability of the influencer to define a future state and find influences to transition to that state will be quite limited. In such a situation, the influencer should initially focus on influences that do not require a high degree of transparency on the part of the constituents. For example, if the intent of the SoS is to facilitate information exchange, but communication interfaces are proprietary and poorly documented, the influencer might offer a standard interface to act as a common bridge between the various undocumented interfaces. This changes the problem for the constituents from one of connecting to many undocumented interfaces to only needing to connect to a standard interface that is well documented. This also removes the need for the constituents to reveal their proprietary interfaces; they only need to certify that they can work with the standard bridge[4].

Anticipation also relies upon the influencer having a reasonable characterization of the constituent. However, unlike observation, the focus is on the drivers and processes by which the constituent make decisions. The reason why decision making process needs to be captured is that the intent in anticipation is to produce a model that reflect (as best possible) the impact of influences upon the constituents. Access to such information may be quite rare especially without direct participation of the constituents in forming the SoS. Surrogates who have similar value propositions to the constituents can be used to make educated estimates of constituent decision making.

[4]A real-world example of this kind of 'bridge' is the packet structure of Ethernet that is maintained across a wide variety of physical transmission media.

A special case that warrants further discussion is that of constituents who are hostile to the intent of the SoS and take active measure to hide their actions from the influencer. In this situation, initial steps should focus on finding ways to observe the hostile constituent and establishing a strong interface boundary to mitigate the effect they can have on the larger SoS. Doing this within an SoS context can be difficult; however, work in network security (both physical and information networks) can be helpful in this case (Bodeau, 1994; Ellison et al., 1997, 1999; Kyamakya et al., 2000). Further discussion of survivability within an SoS context can be found in (Ellison et al., 2008) and (Mekdeci et al., 2011).

Influences are the focus of the next chapter; however, for completeness, some initial comments are provided here. In formulating influence strategies, the influencer must be mindful that SoS are dynamic and so the particular set of influences used may need to be changed over time. To make such changes effectively, it is essential that the influencer can observer the effect of their influences (post-facto feedback) in a manner that reveal not only the result at the SoS level, but also the changes that occurred in each constituent. Having visibility into the actual effect of influences on constituents will allow the influencer to understand the cause of the SoS behavior that they observe.

Finally, reaction is concerned with the direct response of the constituents to the influences. A key piece of reaction is communication between the constituents and influencer. Anticipation is never perfect. There will always be aspects of the constituents and interfaces that could be better characterized. No one knows the constituents better than themselves. Therefore, having effective means of communication and providing incentives to use them are crucial to catching potential mismatches between the influencer perception of the constituents and the constituents actual decision making/behavior.

Changing focus from the descriptive to the prescriptive, the next chapter examines one part of the AIR framework in greater detail: Influences.

Chapter 5

Influences

The previous chapters have focused on describing and understanding the emergent class of systems problems known as systems of systems. This chapter shifts the focus to a prescriptive lens asking the question:

> *What approaches can be used by external SoS influencers to cause constituent decision makers to change constituent systems so as to induce a desired behavior from the SoS?*

Existing SoS research cited in chapter 2 recognizes this as a challenge. Sage and Biemer (2007), for example, refers to the need to coordinate between lifecycle development of the constituent and of the SoS. In terms of how such coordination might occur, he argues that the acquisition process by which the SoS enterprise acquires capability from independent constituents can be used to steer constituent system development in support of the SoS. The report on SoS from the AF/SAB (2005b) found that "motivation mechanisms" used in defense programs such as directives and standards may prove insufficient in the SoS case. They recommended that wider set of mechanism be considered. They further recognized that the "mechanisms" employed, to use their term, may change over time and that venues for feedback between constituents and those responsible for SoS (who are, by their nature, influencers). This same sentiment was echoed by Boehm and Lane (2006) in their spiral acquisition model for SoS. Ames et al. (2011) put it succinctly "we know a lot more about how to create models of system behavior than about potential ways of changing their behaviors."

The chapter begins with a mathematical representation of AIR that builds upon the idea of a system as a value creating entity developed in chapter 3. This representation enables an expression of the decision problem being solved by the constituent decision makers and by the influencer. Note that once again the distinction between constituent decision makers and constituent systems is maintained as in chapter 4. For brevity, references to constituents changing mean constituent decision makers taking action to change constituent systems and constituent decisions are choices made by constituent decision makers.

From that representation, using key leverage points within the constituent decision maker's problem, five basic influence mechanisms are identified. Examples are then provided for how these influences might be used in some of the cases introduced in chapter 2.

5.1 A mathematical representation of AIR

As in the development of AIR in chapter 4, the mathematical formalization of AIR begins with constituent systems. Assuming a fixed[1] set of constituents, the decision problem of the constituent decision makers can be formulated as follows.

The decision makers in control of the constituent systems are said to be continuously solving a utility maximization decision problem. They are making changes to their system in response to changes in the context in which their system(s) operate. This context imposes constraints upon the choices that the constituent decision makers can make. In addition, the constituents are interacting with each other and, therefore, impacting each other's decision problem.

The following notation is introduced to represent the constituents' decision problems. Time is viewed as discrete steps wherein constituents successively refine a set of decision variables so as to maximize their desired objectives. As formulated, it is assumed the decision makers only consider making choice one time step into the future, make their decisions in private and implement their decisions simultaneously. Relaxing these assumptions to allow multi-period and asynchronous decision making would be more realistic, but would require a more careful representation of changes in the knowledge set of each constituent decision maker over time. Constituents are labeled $c = 1 \ldots C$. Each constituent decision maker has an objective function, u^c. The value of this objective function is dependent upon the behavior of the corresponding constituent system which in turn depends (in part) on the design and operational choices made by the constituent decision maker. These decisions are represented by a vector of decision variables, x^c. The context within which the constituents operate imposes constraints on their actions. These constraints, g^c, bound the space of decisions that the constituents can make. The constituents are not isolated from each other, so, both u^c and g^c are affected by the decisions of the other constituents, $x^\bullet$. More precisely, if another constituent decision maker, c_1, chooses to change their constituent system, i.e., change their x^c, that will effect any other constituent system that the c_1 interacts with. The decision makers responsible for those constituents systems must therefore take into account the actions c_1 might take in solv-

[1] The set is fixed in the sense of no new constituents entering or existing constituents departing. Constituents will change as constituent decision makers take actions to modify them.

ing their own decision problems. As constituent decision makers choices are not visible to other constituents (at least not before they have become manifest in their respective constituent systems), a given constituent must make his decisions based upon a prediction of the actions of the other constituents. This prediction is indicated by a hat, $\hat{x}^{\bullet}$. There is also uncertainty with respect to the context and so the context imposed constraints must be estimated. At each time-step, constituents attempt to solve the following optimization problem to generate the decisions to be implemented in the following time-step.

$$x_{t+1}^{c} = \begin{array}{ll} \underset{x_c}{\text{argmax}} & u^{c}\left(x^{c}, \hat{x}_t^{\bullet}\right) \\ \text{subject to} & \hat{g}^{c}\left(x^{c}, \hat{x}_t^{\bullet}\right) \leq 0 \end{array} \tag{5.1}$$

One of these problems is specified for each constituent. The SoS influencer is also solving a decision problem. They are choosing influences to impose upon the constituents. For simplicity, the formulation is described here for a single influencer who themselves is not a constituent system decision maker. Their problem is best described by Figure 5-1.

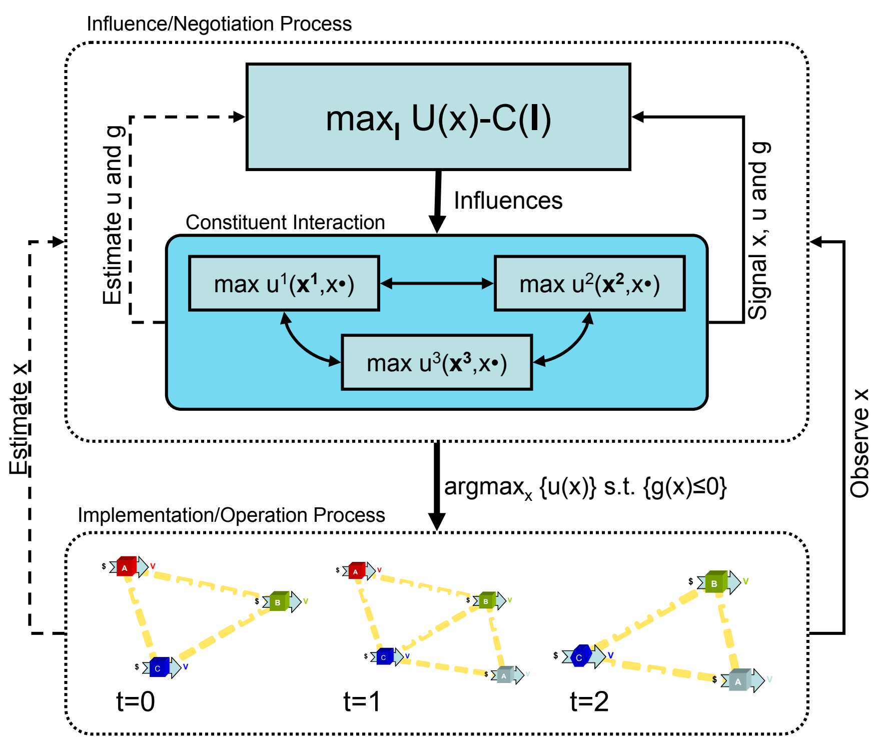


FIGURE 5-1: The SoS influencer decision problem for three constituents

The upper box is the objective of the SoS influencer. $U(x)$ is the utility perceived by the SoS influencer. This is dependent upon the decisions made by the constituents as these decisions determine (along with any interaction with the external context) the behavior of the constituents. SoS behavior then arises from that constituent behavior.[2] As the influencer has a preference on this emergent SoS behavior, his utility is a function of $x = \left\{x^1 ... x^C\right\}$. The term $C(I)$ represents the costs associated with applying the influences upon the constituents.

In Figure 5-1, the overall objective is expressed as the difference between the utility experienced from the SoS behavior and the cost incurred in inducing the constituents to produce that behavior. This, of course, require both terms to use the same units of measure. Should this not be the case, a more general function that accounts for both utility of perceived SoS behavior and cost incurred can be used. The influencer seeks to identify an influence strategy, I, that maximizes his objective. In doing so, he must estimate the constituent decision problem (anticipation) as well as estimate the current state of the SoS (observation). This allows him to capture the current state of the SoS and predict how the constituent might react to a given set of influences. Should there be a mechanism to allow the constituents to react, these signals can be used to update the influencers estimate of constituent decision making. Constituent decision makers take actions that maximize their perceived utility, $\text{argmax}_{x^c} \left\{u^c\left(x^c, \hat{x}^\bullet\right)\right\}$ s.t. $\hat{g}^c\left(x^c, \hat{x}^\bullet\right) \leq 0$. The result of those actions change the constituent systems and whose new structure, behavior and interactions are observed by both the influencer and constituent decision makers who update their estimates of x.

5.2 The principal-agent problem and mechanism design

The mathematical formulation in previous section makes plain the principal-agent relationship between influencers and constituents. Notice that in Figure 5-1, the influencer's objective function depends upon the decisions being made by the constituents. This is precisely a principal-agent relationship as was described at the end of chapter 3. The basis by which x is chosen is not to maximize the influencer's objective function, but rather to maximize the objective functions of the constituent decision makers.

The study of finding such influences in a principal-agent problem is know as mechanism design. The agents are said to be participating in a game in which their outcome are dependent upon not only their own choices, but also the choices of the other players. As they have some (but not complete) knowledge of motivations and available actions of their fellow agents, the game is strategic. Left on their own, the agents may eventually settle into some equilibrium solution to the game. This solution may not optimal from the perspective of an external principal. The goal of the principal, therefore, is design and implement a change in the rules and/or payouts of the game so that the

[2]Constituent behavior here is that which occurs when the constituent are part of the SoS, that is interacting via interfaces, not their behavior in isolation.

new equilibrium to which the agents migrate is also preferred by principal. This change is called a 'mechanism' (Binmore, 2007b).

Restated in the terminology of this thesis, the principal is the SoS influencer, and the constituents are the agents. More precisely, the constituent decision makers are the agents who make decisions with respect the constituent system and then derive value from operation of the constituent system as described in chapter 3. The coupling between the constituents' decision problems described in the prior section (the $\hat{x}^{\bullet}$ terms) create the strategic game[3] that the constituents are playing.

The key challenge is ensuring that the mechanism are *incentive compatible* with the locally expressed needs of the constituents (Sappington, 1991). Using an incentive compatible mechanism ensures that actions which are in the constituent's best interest are also in the SoS influencer's best interest, i.e., that the new equilibrium solution to the game is better aligned with the preferences of the influencer. To accomplish this requires examining the factors affecting constituent decision making and then exploiting the factors to shift the equilibrium in the constituents' game.

5.3 Five basic influences

Given this mathematical formulation, what strategies are available to the influencer? The influencer is trying to change the choices being made by the constituents. Therefore, a natural starting point for developing the influencer's strategies is the constituent decision problem. Proposed below are five distinct ways that the influencer can exert influence upon the constituent decision problem. Each influence mechanism impacts a different part of the problem being solved by the each constituent decision maker. They are outlined in Figure 5-2. Each approach is identified by a word that begin with 'I' as a mnemonic. They are summarised in Table 5-1

Constraints have been split into two groups, g_S and g_T reflecting those that arise from social and technical concerns respectively. Social constraints are those that arise from institutions and norms that exist within a given constituent's context. Examples include laws, standards, and agreements between actors in the context. Technical constraints arise from the physical limits of the technological solutions being employed by the constituents. These limits may be fundamental as in the case of laws of physics or may be a consequence of having the chose a particular technical solution given the resources available.

5.3.1 Incentives

The first influence mechanism is 'Incentives'. This refers to direct compensation for constituents acting differently than they would without the influence, i.e., different from the local utility maximizing action. In the case of a financial objective, the incentive can be viewed as an additional term in the value function of the constituent (as shown in Figure 5-2).

[3]Game as in game theory.

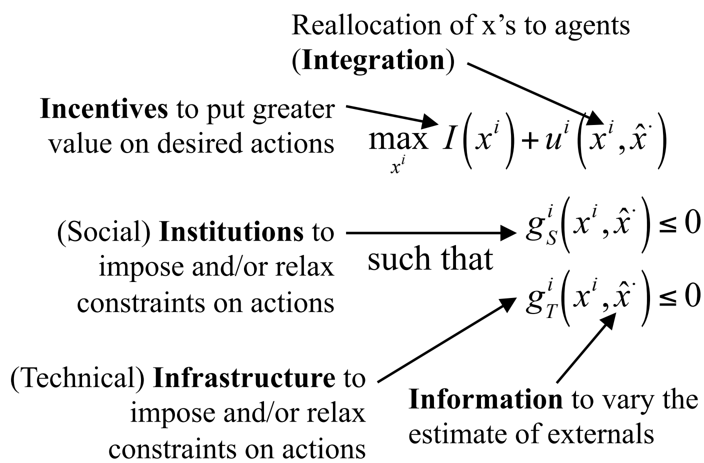


Figure 5-2: Changing the constituent decision problem via influences.

Table 5-1: Mechanisms for influencing constituents

Influence	Definition
Incentives	Compensate the constituent for taking an action that does not maximize its utility.
Institutions	Change the rules and norms under which the constituents interact to promote emergence of new, desired, behavior.
Infrastructure	Impose/remove technical constraints on constituent actions.
Information	Provide constituents with information so as to modify their estimate of the uncertain terms in their decision problem.
Integration	Change which constituents control different parts of the SoS.

Incentives can be both positive and negative (i.e. penalties—assuming the influencer has the ability to enforce it). A subsidy to offset additional costs associated with SoS participation would be a positive incentive. Contracts to perform SoS functions are also a form of incentives as they are, in essence, providing something of value (dollars) in return for taking a specific action. A related issue is that of capability maintenance. Often in SoS, not all the constituents are actively participating all the time. It may not be in the best interest of a non-participating constituent to maintain interfaces that would be used should their participation be needed in the future. Incentive payments can also be used to compensate such a constituent for having to retain the ability the participate even when not in use.

Negative incentives include taxes, penalties and fees. They differ from the positive as punish rather than reward a particular set of actions. Within an SoS context, an example is overage charges on cellphone plans for excessive use of the shared resource (i.e. the cell network).

Additional information on designing incentives can be found in Sappington (1991).

5.3.2 Information

While is hasn't been discussed in much depth thus far, uncertainty plays a key role in decision making of both constituents and influencers. This effect is represented by the need for estimation of the current value and predication of future values of terms that appear in the decision problems above. The constituent perception of his objective (payoff) function is dependent upon his prior knowledge of the world and his estimates of the other actors decisions, objectives and constraints. The influencer can provide information to change either of those factors. Mathematically, game theory provides a construct, Bayesian games, that allow such uncertainties to be represented. For a more detailed explanation of that approach the reader is encouraged to consult (Gibbons, 1992).

An example of the influencer providing information to change constituents perceptions include establishing a road map for future SoS activities to allow constituents to invest and plan ahead. In that case; however, the influencer would need to ensure that the perceived present value of the future payoff promised in the road map exceeds the present day investment or an additional incentive may be required.

The influencer can also create a mechanism in order to get constituents to reveal private information. The challenge, of course, is to ensure that the constituents are truthful.[4] For example, if different constituents have different cost structures that the influencer needs to cover when the constituents participate, the influencer may wish to find the lowest cost constituent. One way to do this would be to hold an auction for participation opportunities. In an auction designed to be incentive compatible, it would be in the constituents' best interest to submit bibs reflective of their true costs.

[4]See the revelation principle in Gibbons (1992).

Such auctions do occur in the real-world. Examples include take-off slot allocation for airlines at airports (Le et al., 2004) and transport carriers bidding on shipment offers—both cases involve bidding to participate in an SoS.

5.3.3 Infrastructure

Constraints in the constituent decision problem (Figure 5-2) are divided into two groups corresponding to those arising from physical or technological limits and those arising from organizational or institutional or even societal limits. Influencing the former, can involve making changes to component systems or subsystems within an SoS. These parts of the SoS are not explicitly represented in AIR framework as they do not exhibit the same principal-agent relationship as constituents and influencers; however, they often play a foundational role in establishing the context in which constituents interact.

As stated in chapter 1, a good example of infrastructure is a communication network used to allow constituents to share information. This could come into form of either a subsystem—if it is purpose-built for the SoS—or as a component system—if it already exists as an operationally independent entity, but is brought under the control and use of the SoS. The power of such infrastructure to enable new SoS form and behavior is quite well demonstrated by net-centric warfare. A detailed case study of that application is documented in Tisserand (2005); Cammons et al. (2005).

5.3.4 Integration

This influence mechanism may be the most difficult to implement as it involves changing mapping between constituent decision makers and constituent systems. It is, in effect, re-architecting the constituent set. As the constituents are (managerially) independent, they may have the ability to obstruct such an effort. An example is vertical integration of an intermodal transport chain as described in Van Der Horst and De Langen (2008). They report that such integration was difficult given culture and business practice clashes between chain members.

Also useful for the influencer to consider is dis-integration, i.e., breaking up a constituent into two or more operationally and managerially independent parts. A real world example of this is the privatization of rail transport. This takes what was once a unified, single decision maker system and splits it functionally into several separate systems that inter-connect (e.g. track and yard operations, passenger service and freight). While it may have been difficult for new competitors to enter into the market when faced with a fully integrated incumbent, the dis-integrated, SoS version allow for such competition by having a lower barrier to entry for a new constituent.

5.3.5 Institutions

The final influence mechanism seeks to change institutionally imposed constraints faced by the constituents. The study of institutions and how they change is a discipline in its own right and a full treatment is beyond the scope here. The political economy literature is a good place to start, in particular the papers by Gourevitch (1978); Krasner (2010).

Gourevitch (1978) proposes a typology of influences that can be used by state actors to affect the other state actors. Key to this typology is the notion that effective influence requires a clear understanding of not just the formal organs of the target state (i.e. the government), but also the underlying structure from which it derives it powers (i.e. the people) as both present opportunities. He examining four influence types. The first is “military intervention”. This is the most direct way to cause change; however, it comes with a rather high cost. In the SoS case, such destructive intervention can occur in the form of integration or highly coercive incentives. The second type is “meddling”. This is less direct, covert interference into the domestic politics of the target state. Examples include cultivating rebellion, spying, bribery and coups. Here the SoS analogy might be surreptitiously providing key information to skew a constituents view of the context. The third influence type is “international economy”. Here there is recognition that, within the international system, states are economically interdependent and so have a strong interest in taking actions that are mutually beneficial. Conversely, that dependence can also imply that negative actions towards a target state may have unintended side effects on others. This can lead to the economic interconnection effectively creating constraints of state action thus binding states to serve the “greater good”. Achieving alignment through constraints is one possible outcome of exercising the infrastructure influence mechanism. The final influence type is the “international state system”. This refers to formal institutions and diplomatic relations which states use to increase, from their perspective at least, the stability of the international system. This, of course, has a direct analogy to using an institutional mechanism to affect constituent action. These four types only scratch the surface of the scholarship within the political economy literature on this topic. However, they do provide further support for the 5 Is as fundamental, distinct mechanisms by which one party can affect another.

A concrete example of institutions is the set of technical standards that regulate communication on the Internet. They trade the benefit of commonality against the risk of sub-optimal performance in some situations. As the Internet demonstrates, standardized interface makes reconfiguration much easier. Furthermore, they provide point of common reference that can help align the objectives of constituents affected by them. Hsieh (2007) examined the history and development of the Internet standards.

5.4 The 5 Is in earlier case examples

SoS case examples from section 2.2 are now used to demonstrate the 5 Is in practice.

5.4.1 HousingMaps

In the ‘HousingMaps’ example, the SoS architecture problem of interest is how Google can manage a growing set of interactions with other services via its maps service of which HousingMaps was among the first. Each of the 5 Is provides Google a different way to impact the behavior of those integrating with Maps as part of a mashup. Google *incentivized* websites to use the API (application program interface—see subsection 2.2.3) by lower the cost of incorporating Maps into their services. Google assigned each developer a unique key by which they accessed the map service. This key allowed Google to track and collect statistics on how each developer incorporated maps functionality into their own solutions. This same analytic data is available to the developers. By sharing these analytics, i.e., new *information*, with the developers, Google created additional value for the developers, better allowing them to understand their user base. Google moved certain function (e.g. cacheing of geo-location requests) that would ordinarily be the responsibility of the client web-service to their own servers. This is an example of *integration*. Google's provided API functions that standardized the interface between the third-party websites and Maps. The API is an example of *infrastructure*. In order to use the API, developers had to agree to terms of service specified by Google. This *institutional* mechanism formalized the relationship between Google and the developers, thus influencing the developers to act in a manner that was mutually beneficial. For example, to manage traffic created by the mashups, Google's terms require that the websites be careful not to over-burden the system with unnecessary requests and provide avenues for developers to work with Google should a particular application require a larger than usual share of Google's resources. Furthermore, the terms allow Google to place advertising in the mashup (at their option) that can be used to offset the costs to Google in providing the service.

5.4.2 Internet Peering

Level 3 had several options available to it to resolve its peering dispute with Cogent (see subsection 2.2.1). Using the influence mechanism of *incentives*, Level 3 could have renegotiated the peering agreement with Cogent to require compensation from Cogent in the event that traffic is imbalanced. In terms of *integration* they could buy Cogent's business (or sell their business to Cogent). They could exchange additional *information* with Cogent, e.g., traffic data and projections to better allow planning by both parties and possibly avoid traffic imbalances. They could have modified their technical *infrastructure* to selectively reduce the quality of service for Cogent customers traversing the Level 3 network. This would encourage those customers to find alternate transport thereby reducing the traffic imbalance. Finally, they could use dispute resolution *institutions* provided for in a typical peering agreement. Each of these approaches has side-effects and one may leave Level 3 in a better position that another; however, they do demonstrate the variety of strategies available.

5.5 Limitations and extensions of the 5 Is

The formulation of the 5 Is is rooted in the coupled set of decision problems that started this chapter. For both the constituents and influencer, alternative formulations may be more applicable for some SoS. For example, instead of solving a value maximizing decision problem, an influencer might only wish to ensure that certain SoS performance parameters stay within constraints and allow the constituents to behave in a less guided manner. This is of practical importance when trying to discover emergent SoS forms and behavior that arise from constituent to constituent interaction rather than influencer direction. A simple example of this ad-hoc networking (Krag and Büettrich, 2004). By allowing the constituents freedom, novel SoS forms can emerge.

The 5 Is are presented here almost as a menu of possible choices. Real world strategies should be composed of multiple Is changing as the SoS, the constituents and the context evolve. How to develop and execute combined strategies is key consideration in making the 5 Is a practical part of the SoS influencer's toolbox.

This chapter concludes the presentation of the theoretical contributions of the thesis. The AIR framework as presented in chapter 4 provides a powerful representation of the key decision makers in an SoS, the interaction between them, and, how those interactions result in/changes SoS behavior. This chapter built upon that foundation taking a deeper look at the specific strategies available to an SoS influencer. The 5 Is are a first step in developing influence strategies. However, implementation of the framework and 5 Is can be challenging.

To demonstrate some of these benefits and expose some of these challenges, a case example of intermodal freight transport using AIR and 5 Is is presented in the next chapter. Intermodal transport is the movement of goods via multiple transportation modes, e.g., truck and rail. In this case, a hypothetical freight transport system consisting of interconnected road and rail network is used as the set of constituents. The influencer is a government authority who wishes to change relative usage of road vs. rail in the transport system in order to reduce road congestion and pollution. However, they cannot force traffic onto a particular routing and must use their influence to change the interactions between the constituents. As these interactions are dynamic and closed formed solutions unlikely, a simulation model will be used. In line with the AIR framework, the SoS influence problem will be formulated a principal-agent problem.

Chapter 6

Intermodal Freight Transport Case Study

This chapter will demonstrate the use of the AIR framework and the 5 Is when applied to an example intermodal freight transportation network. The case example begins with a brief motivation for the importance of studying intermodal freight transport. The full scope of issues surrounding the development and operation of an intermodal freight transport system is beyond the scope of this thesis; however, such a system does exhibit many of the challenges seen in the SoS examples discussed in earlier chapters and thus is a good example to demonstrate the insights that can be gained by using the AIR framework and the 5 Is. Therefore, a simplified model transport network will be studied with the intent of understanding the interactions between the different stakeholders that results in SoS behavior through a modeling exercise. The model is then used to explore potential influences to alter that behavior.

6.1 Research issues in intermodal freight transportation

As the supply chain becomes more global, there has been increasing focus on freight transportation. Driven by both increasing demand and increased concern for externalities such as environmental damage and noise, one area of focus is making more efficient use of transportation networks. In terms of overland transport, there is much interest in better understanding the unique socio-technical challenges in the design and operation of intermodal freight transport systems (Transportation Research Board, 1998; Bontekoning et al., 2004; Caris et al., 2008). Intermodal freight transport refers to transportation solutions that, from an origin point to a destination, use two or more transport modes. For example, manufactured goods arrive at a port on a container[1] ship, then are loaded onto a train for a journey inland to a logistics center, and finally are delivered to local retailers via truck. For inland transport in particular, research into rail-truck intermodal

[1]A container is a standardized metal box used to store goods during shipment. See the photographs in this chapter for examples.

has revealed potential for significant cost savings when compared to using trucks alone. This is a consequence of the greater efficiency of rail over long distances carrying large numbers of containers at a time, i.e., as a train. In addition, using modern locomotion technology, rail can generate less pollution than the same move by truck (Janic, 2007). On both these accounts, increasing the use of rail via intermodal links to the trucking system appears to be a beneficial policy objective.

Figure 6-1: A 'double-stack' of containers on flat cars in an intermodal transfer yard near Chicago

On the other hand, rail has disadvantages in terms of quality of service (Sussman, 2000). Quality of service is an umbrella term for the various factors including, but not limited to, cost that shippers look at when making transportation decisions. Other factors include pickup timeliness, overall transit time from origin to destination, variability in transit time and likelihood of damage during transit. Rail may not perform as well as trucks on these metrics. Additional handling of shipments as well as time spent waiting at terminals increases the overall trip time and introduces opportunities for damage to occur. Delays can also occur more easily on rail than trucks. Since trains are often composed of shipments from multiple origins, a shipment could be delayed as it waits in the terminal for other containers to form a train or to keep with the train's schedule. In addition, rail networks have much less routing flexibility to make adjustments for weather or other unforeseen hindrances. Finally, unlike say air transport, freight trains operate at high volumes throughout the day and night. This makes recovery from adverse events via repositioning of capacity quite difficult and causes delays to build up throughout the week only to be settled on weekends when traffic tends to be lower. Reduced quality of service makes shippers who currently use trucks wary of incorporating rail. This is especially apparent in the modern supply chain that has experienced an increase in speed and demands much tighter tolerances with respect to timely movement of goods.

Looking at a specific example, Janic (2007) models freight flows within Europe including both realized costs and externalities. Janic derives breakeven points between an intermodal solution and one that only uses the road network. At a rate of 5 trains/week, this breakeven point occurs at a door-to-door distance of 1000 km. This means that for a move of greater than 1000 km, the intermodal solution is cheaper when externalities such as air pollution are priced and included in the cost. Taking advantage of economies of scale, when 25 trains/week are used, the breakeven distance reduces to 700 km. One would expect, therefore, that an intermodal solution would capture significant share given enough train frequency. The data as reported in Janic do not show this. The European Commission (1999,2000) estimates only 10% market share (as measured by volume of containerized freight tonnage) in 1000km market and 2% share in the shorter 200-600km market.

The reasons for low utilization are many. In addition to the generic challenges for rail listed above, there are challenges specific to Europe. For example, while the high speed passenger rail network has had significant growth in recent years, their tracks and other facilities often cannot be shared by freight trains that move much slower and have different wear effects on the infrastructure. Bontekoning et al. (2004) review shipper mode choice and pricing literature to identify the factors that go into the decision to use an intermodal solution. Cost-driven decision makers tended to go with intermodal solutions (Tsamboulas and Kapros, 2000) while those who were more concerned with service quality in addition to cost tended to use intermodal solutions less. Murphy and Daley (1998) point out that non-users of intermodal services have a lower perception of its service quality than users. This difference in service quality is not just perception, but a real effect caused (in part) by delays introduced in terminal operations. As such, road operators require significant discounts in order to switch to an intermodal solution (Fowkes et al., 1991). Bontekoning concludes that while the mode choice literature reveal a strong sensitivity to cost vs. service quality, the studies tend to be situated in specific cases and the results may not be more generally applicable.

In the United States, the situation is somewhat different, as intermodal freight transport has seen wide adoption especially for high-value, long distance moves (Chatterjee and Lakshmanan, 2008). Historically, the development and offering of Intermodal routes in the US has involved coupled changes in technology, organization of the transport (esp. rail) industry and public policy. As described by Plant (2002), in the pre-WWII era, railroad tried to offer a nascent form of trailer-on-flatcar service to reduce on/off loading costs at terminals. Rail, at the time, was a regulated industry with pricing set by the ICC (Interstate Commerce Commission). The ICC ruled that while the railroads could carry the trailers, marketing and control of the service still laid with the truckers. The prices charged by the railroad under this regulatory mandate created little incentive for truckers

Figure 6-2: A container being loaded onto a truck

to not simply make the entire journey themselves and, as a consequence, keep all the revenue. Not much investment and therefore progress was made. The next several decades saw vast reorganization of the rail industry with a slow movement toward deregulation. Through a series of mergers the market became dominated by a few, larger but still regulated railroads. Deregulation finally came with the passage of the Staggers act in 1980. With greater ability to set their own prices and to negotiate their old contracts, railroads were much more willing to adapt their services to better handle inter-modal loads. Supported by technological developments in the 80's and 90's, the deregulated industry witnessed a boom in inter-modal traffic (see Figure 6-3).

Significant challenges remain in the US to sustain the adoption of intermodal solutions. DeWitt and Clinger (2000) outline four important issues at the start of the 21st century. First, globalization has resulted in an increased demand for globally integrated supply chains with many links owned by many different, geographically diverse parties. Such complex supply chains represent both a technical challenge of organizational coordination as well as a business challenge of establishing the needed agreements to form and operate the chain. These agreements involve not only the rail and truck companies, but also government as regulator or as owner of public infrastructure used by the mode operators. Second, customer requirements will further expand beyond cost to emphasize "speed, flexibility, variance elimination and relationships with other members of the supply chain" (DeWitt and Clinger, 2000). Third, they recognized the increasing and vital role that information

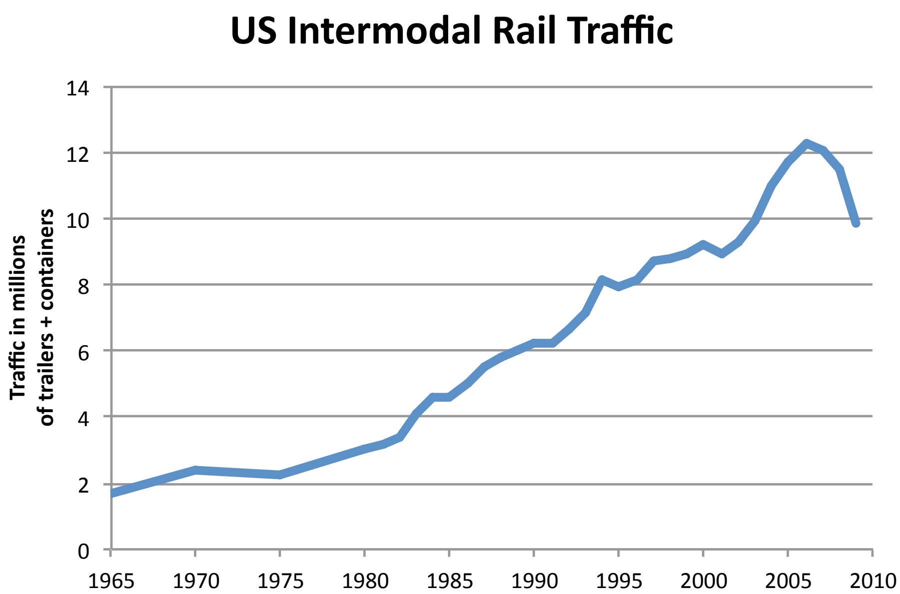


FIGURE 6-3: Intermodal traffic in the US has grown significantly in the past several decades (AAR, 2010, Table 9.9)

technology plays in operating the modern supply chain and the logistics work force will need to be re-trained to work in this new environment. Fourth, partnership between public and private entities will be needed. The transition to a more integrated supply chain involves changing both public assets (such as infrastructure) and private assets. Integration will require investment by both public and private entities. To ensure that such investment is done efficiently, partnerships will be needed to coordinate development.

Looking at the factors affecting the availability of intermodal solutions, several issues arise. First, there is an open problem of how to price intermodal solutions and how to distribute that tariff among operators of the different modes involved. Hurley and Petersen (1994) derive an equilibrium pricing solution for the overall intermodal problem, but neglect the distribution of surplus (net profit above cost) amongst the operators. Also, only cost is considered and not service quality. It is possible that the surplus could be used to compensate more reluctant shippers for loss in service quality. More recently, Fernández L. et al. (2003) approaches the problem as one of matching transport demanders with suppliers of services to form a mix of uni-modal and intermodal chains. Going beyond cost, Fernández L. et al. (2003) includes timeliness of delivery as a constraint in the mode choice decision. Neither of these authors have dealt with the issue of contract formation directly. Showing that an intermodal service offering would be profitable in the market does not necessarily indicate that such an service would be provided in practice. Rather, one needs to look at how each of the constituents involved in the intermodal service would view the agreement.

Real-world intermodal services are commonly offered via three approaches. First, a shipper (or her agent) can purchase each leg of an intermodal move separately. This is the case of a freight forwarding company. Second, a single company can own multiple modes and offer origin to destination service combining the modes as appropriate. Large logistics providers such as UPS and FedEx offer this type of service, combining, for example, air and road travel using company owned planes and trucks.[2] Third, firms may form cooperative agreements to jointly offer intermodal service. Establishing such agreements has historically been difficult with issues of equitable allocation of revenues between the parties being a challenge (Van Der Horst and De Langen, 2008).

From a technical perspective, the operation and management of intermodal terminals plays a key role in determining service quality. Interchanges between modes are an entry point for both delays and damage risk within the transport chain. Figure 6-4 shows an intermodal terminal. Roads for trucks are in the lower-left and rail lines are in the upper-right. In the center of the photograph is an orange gantry crane is visible that moves containers between rail flat cars and trucks. The large storage areas are needed to manage the complex scheduling and handling requirements such diverse traffic flows.

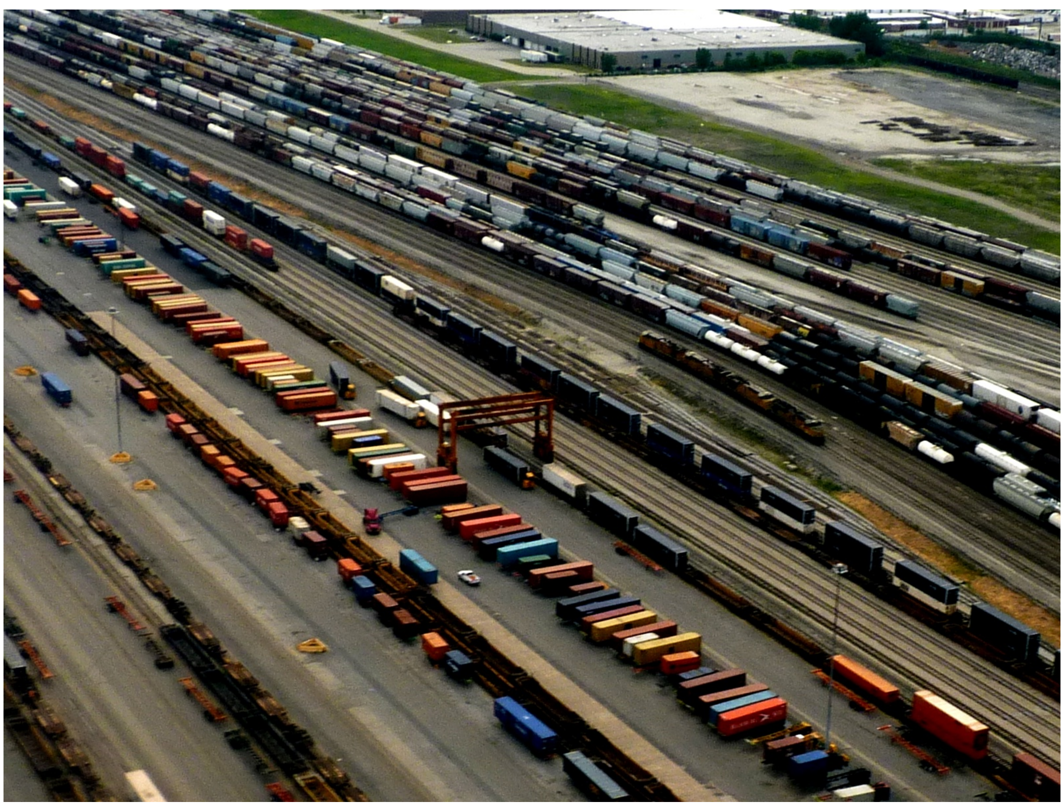


Figure 6-4: An aerial view of an intermodal terminal.

[2]They also contract with outside firms as in the third approach.

6.2 Case study objective

The objective of the case study is to apply the AIR (Anticipation-Influence-Reaction) framework to look at the above issues in the use of intermodal solutions. In framing the intermodal freight transport problem as an SoS, the constituents are taken to be the mode operators, i.e., truckers and railroads and the SoS influencer is an external entity (such as a government transportation authority) who wishes to examine various influence strategies that could increase traffic on the rail portion of their transport network. A simplified, but illustrative intermodal transport network is used as the venue for the case study. The model attempts to account for, at a high level, several of the current issues cited above in the studies of the American and European intermodal markets. For example, the shipper choice model accounts for factors beyond price in decision making.

Figure 6-5 shows the simple intermodal network used in this case study. There are two origin-destination pairs between which demand needs to be routed (O1–D1 and O2–D2). For each pair, there are multiple uni- and intermodal solutions connected through intermodal terminals A and B. For example, between O1 and D1, there are two long-haul truckers (T1 and T2) and an intermodal route (T3–R1–T7). Both O1–D2 and O2–D2 have intermodal routes that connect via R1 and therefore the usage outcome for this link is dependent upon shippers moving between both O–D pairs. Even in this simple case, intermodal freight transport can give rise to a complex competitive situation.

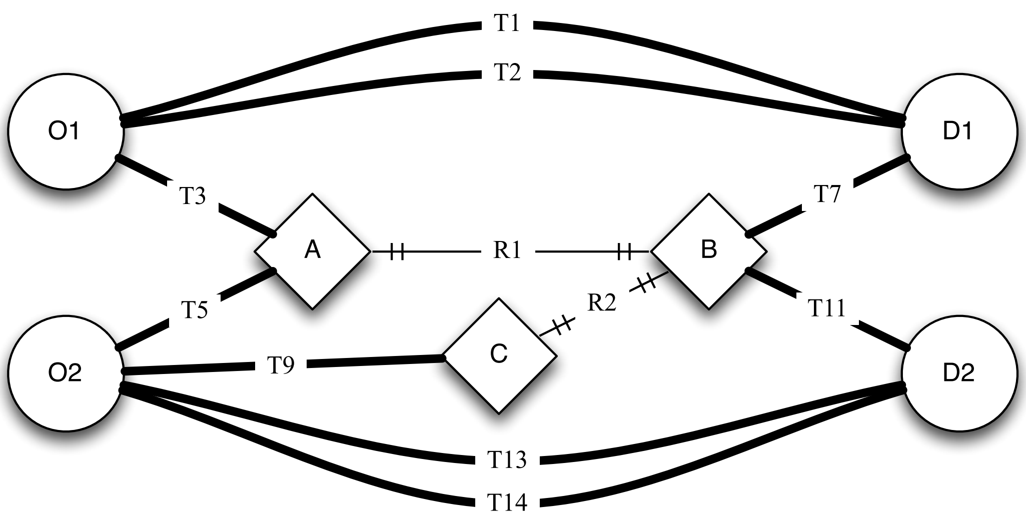


Figure 6-5: A simple intermodal network

There is an ongoing effort to improve terminal operations that points to the role of interfaces in determining SoS performance (Ballis and Golias, 2002). Rizzoli et al. (2002) modeled detailed operations of intermodal terminals and identified several areas that could be improved: (1) IT infrastructure; (2) scheduling problems for cranes and other terminal internal equipment; and (3) congestion at times of peak load. While a full treatment of the role of terminals is beyond the scope of the thesis, two influences related to terminals are examined: monetary compensation to shippers for the additional delay caused by terminal and improvement in terminal throughput via technology. Needing to travel through terminals has a cost both in terms of any handling and storage fees charged by the terminal and, indirectly, the additional logistics costs due to time spent in the terminal. To examine the effect of reducing these costs, a subsidy is provided to shipper in the second influence strategy. As detailed in the papers cited above, there is much room for improvement in processing time within terminals. The second influence strategy looks at the effect of reduction in this processing time. The reduction is sized to improvement estimates in the literature. The key difference between these influences is that the former compensates for a limitation in the terminal while the latter makes a technical change to improve the intermodal terminal.

Two different influence options are examined to encourage the formation and maintenance of intermodal routes: taxing road use and allowing cooperative routes to be negotiated. The former doesn't directly lead to intermodal solutions; rather, it increases the cost of uni-modal truck solutions (that rely much more heavily on road than intermodal) thereby making intermodal option more cost effective by comparison. Such a tax only works if intermodal routes are available. As discussed above such routes can be formed by a forwarding agent; however, this requires purchase of all links on the routes separately. If mode operators form agreements to offer such service directly, they may do so at lower total cost to the shipper. The second influence allows certain truckers and railroads to form such agreements. The Nash bargaining solution (Nash, 1950) is used to solve the bargaining problem of revenue distribution.

To apply AIR, the case moves through the three phases of Anticipation, Influence and Reaction. Anticipation (section 6.3) involves formulating a model (section 6.4) of carrier servicing offerings, shippers decision making and consequent traffic flow on the network in Figure 6-5 is developed with an emphasis on representing the issues and challenges reviewed in the prior section. For the Influence phase (section 6.5), four different influence mechanisms are tested. These are inspired by the cited above from the European and American intermodal markets. They are a tax on road use and a subsidy to recover terminal transfer costs (incentives), speeding up of terminal operations (infrastructure) and allowing cooperative agreements to be formed between railroads and truckers to form new intermodal routes (integration and institutions). The reaction phase (section 6.6) looks at how the constituent might react to these influences.

6.3 Anticipation

Recall that the anticipation phase consists of the SoS influencer attempting to understand the behavior of the constituents (and by extension the SoS) so that he may look at potential interventions. Most SoS are far too complicated to yield a closed form solution. Building a closed form predictive model is impossible.[3] Rather, the SoS influencer should seek to understand the key behaviors of and interactions between the constituents and include those in a simplified model that can be used to better understand the dynamics that emerge when all the pieces interact. This type of 'behavioral' model is much easier to produce.

Using concepts from the models proposed in Hurley (1994), Fernández L. et al. (2003) and Gambardella et al. (2002), the following local (constituent-level) decision makers (DM) are identified: (1) Shippers; (2) Road operators (truckers); (3) Rail operators (railroads); (4) Terminal operators. SoS-level decision makers are (1) Coalitions of mode/terminal operators who offer intermodal service as a door-to-door offering as perceived by the shippers and (2) external SoS influencers. For this study, the SoS design problem is framed from the perspective of an external SoS influencer who has a preference on the utilization of all different parts of the transport network. This may be, for example, a government transportation authority charged with improving the utilization of the rail infrastructure.

To complete the Anticipation phase in the AIR process for SoS engineering, a model of the SoS and its constituents must be developed. The intent of this model is to allow the SoS engineer to evaluate various intervention strategies that may be used to influence the constituents into producing behavior. To that end, the model needs to have the following characteristics:

- Capture the pre-intervention behavior of the constituents. This allows a baseline to be established from which different interventions can be tested.
  - Represent the decision making process of the constituents with respect to the decision variables in their control
  - Represent the outcome (objective) variable that the constituents use to make decisions
  - Represent interaction between constituents
  - Represent any relevant external drivers that effect constituent behavior
- Allow for the introduction of one or more interventions
  - Represent the response of the constituent to each intervention. In what way do they change behavior?
- Evaluate SoS objective function(s) such that the effects of interventions can be compared

Prior modeling efforts for intermodal freight transportation broadly fall into two categories: (1) Equilibrium models and (2) Simulation models. One of the significant difficulties in modeling freight transportation is the sheer number of decisions that are being made. There are many decision makers involved (Flodén, 2007) including carriers, shippers, and terminal operators. In

[3] This can be due to complexity, scale of the SoS, limitations on available data to characterize past behavior and, conversely, inability to use past data when consider novel SoS forms.

addition there are regulators, public interest groups and other influencers working to change the behavior of system operators.

Equilibrium models resolve this difficulty by making strong assumptions about the behavior of the system given enough time to settle to an equilibrium state. For example, Hurley and Petersen (1994) uses Wardrop's system and user equilibrium (Wardrop, 1952) to solve for the joint behavior of the carriers and shippers respectively. Equilibrium models (and, more generally, network flow models) of transportation system have been used extensively to look at strategic issues in system design (Crainic and Laporte, 1997). Examples include facility placement (Arnold et al., 2004), pricing and managing externalities such as pollution. While equilibrium models do have the advantage of being solvable even for fairly large problems, they do not readily allow exploration of the dynamics that occur when the system is not in equilibrium state. Furthermore since dynamics are not explicitly represented, time dependent effects such as congestion are represented using heuristics and/or statistical fits (Fernández L. et al., 2003). Should one wish to study such time dependent phenomena directly, a simulation model offers a richer set of behaviors to explore.

Simulation modeling in transportation has a long history. Problems such as optimal scheduling and routing of vehicles have driven much of the work in this area. Typically such operational simulation take the higher level strategic choices such as amount shipped and pricing as input and therefore define the problem as how to ship a given quantity on a given network rather than determining the structure and properties of the network itself (two examples are Flodén, 2007; Kwon et al., 1998). When looking at a uni-modal problem the assumption of a fixed network of short-run operation optimization is valid since changes to the rail network occur over long time scales. The intermodal problem, however, may involve changes in the network even in the short term as companies must agree to provide intermodal service offerings and thus requires operations-only simulation models to be extended. Recent work using agent-based models has moved in this direction. Gambardella et al. (2002) developed an agent-based model of both the operation and some higher level decision making that occurs in intermodal freight transport. The model consisted of several sub-models that describe the behavior of each of decision makers and other actors within the system. This research uses this agent-based, dynamic approach as a starting point.

6.4 Intermodal freight transportation network model

For the current case, building upon the work cited in the previous section, the following model was developed to look at the simplified intermodal transport network in Figure 6-5. The flows of interest are from two origin points to two destination points. Connecting these are a network of road and rail links. Traffic is simulated upon this network for a period of 15 years.

The model represents the interaction between two types of agents—shippers who wish to use the transport network to manage the stock of a good at the destination point and carriers (railroad operators and truckers) who provide transport service between points on the network. Each quarter, the shippers evaluate which of the available shipping options minimizes their expected total logistics cost and then contract with the one or more carriers to provide shipping service for the following quarter. The carriers meanwhile re-evaluate their service offerings on a quarterly basis as well. They adjust prices and, for rail carriers, service frequency so as to maximize expected profit. All shipments are assumed to be full twenty-foot intermodal containers and are measured in TEUs (twenty-foot equivalent units). The overall flow of the model is shown in Figure 6-6.

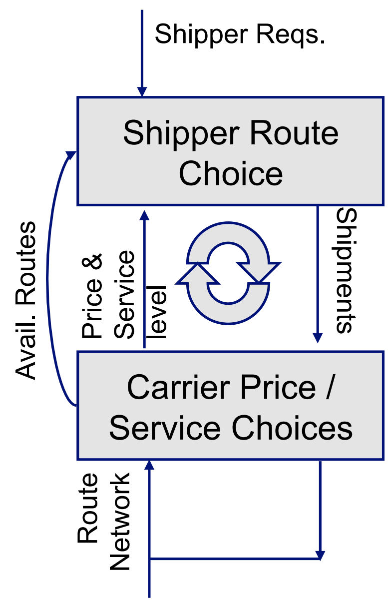


FIGURE 6-6: Overall transportation model flow

Details of the above are described in the following sections. Numbers provided are for the baseline simulation scenario which consists of the route network from Figure 6-5. The long-haul road paths are 500 miles in length. The short-haul road paths (between origin/dest and intermodal terminal) are all 50 miles except for O2–C which is 250 miles. The rail link R1 is 500 miles, while R2 is 250 miles. Remaining baseline parameters are specified as they are introduced below.

6.4.1 The carrier's problem

The model has two types of carriers, truckers and a railroad. Truck carriers are modeled as providing an on-demand service[4] that is un-capacitated, i.e., they sub-contract to however many drivers they need to meet demand. Their only decision variable is the price of the service. They use the same price (on a per container per mile basis) for all the transportation links that they operate. Rail

[4]An alternative is for the truck companies to maintain a fleet of trucks. Future development of the model should consider this possibility.

carriers, on the other hand, offer scheduled service. Their decision variables are price (per container per mile) and service frequency (trains per day).

6.4.2 Truck carriers

The parameters in Table 6-1 are used to specify the performance and costs of each truck carrier. The cost is derived from the Owner/Operator Independent Driver Association[5]. The OOIDA cost model assumes truckers cover 100,000 miles per year. In the model[6], assuming 400 trucks per truck carrier route, annual mileage is 105,000 in the baseline case which is quite close to the assumption in the OOIDA cost model.

Table 6-1: Truck carrier parameters

Parameter [units]	Definition	Sample Value
Link ID	Link(s) owned by the trucker	
Nominal Speed [mph]	Nominal speed of trucks while running	60
Duty cycle [1]	Fraction of time that trucks are operating at nominal speed; i.e., Nominal Speed * Duty cycle = Average speed of a truck along route	0.5
Price [$/TEU/mile]	Price charged to shipper on a per TEU shipped per mile basis	2
Cost [$/TEU/mile]	Cost of moving a container one mile as per the Independent Owner Operator Association cost model	1.15

Truck carrier price optimization

Truck carriers use the following heuristic to find a price that will maximize their profit in the next quarter. The heuristic attempts to use both the past experience of the carrier along with the carrier's ability to make estimates of future profit for changes in price. Carriers can only compute a limited number (11) of such point estimates thus reflecting the real-world high cost of forward looking market studies. They must also rely upon forecasting to estimate competitor behavior.

1. Forecast competitors prices forward one period using exponential forecasting (Brandimarte and Zotteri, 2007, chapter 3) on the recorded price history of the competitors.
2. Using the forecast prices for competitors, estimate profit for setting the price to each of the following: Keeping the price the same; increasing/decreasing it by 1%, 5% and 10%. This esti-

[5] http://www.ooida.com/Education%26BusinessTools/Trucking_Tools/costpermile.shtml

[6] These figures vary somewhat under different model scenarios, but don't result in unrealistically large numbers of trucks being used or large shift in annual mileage.

mate is generated by running a restricted version of the model. All traffic not passing along routes served by the carrier is removed.

3. Compute a best fit parabola for price vs. profit using the seven forecasted points generated in the previous step.
4. Find the maximum of that parabola
5. If the maxima lies within the +/- 10% price change interval estimate its profit (using the model, not the curve fit) and add the resulting point to the set of point used for the fit.
6. Repeat steps 3-5 four times. This was chosen to represent the compromise between accuracy of the heuristic generated optimization result and expenditure of resources by the carrier to gather and process the data needed for their pricing problem. Each round represents another attempt to estimate the revenue and costs for a given price. As such estimates are not free, only limited a number of rounds could be completed. Varying this number revealed that, after four rounds, there was little shift (<10%) in the estimated optimal price for the next quarter.
7. Set the price to the value that has the highest estimated profit from those generated in steps 2-6.

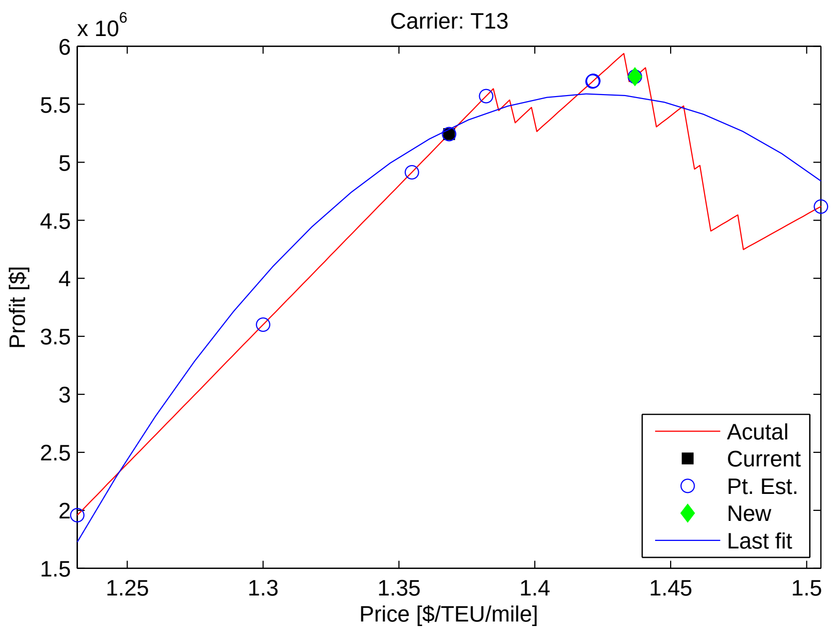


Figure 6-7: Example price finding for a trucker

Figure 6-7 shows a sample run of the truck carrier price heuristic. The black square shows the price for the previous time step. Each drop of the actual profit curve as price is increased indicates another shipper choosing a route option that does not include the trucker. More shippers choose other options as the price increases. The blue circles indicate the estimated sample points from step 2 in the heuristic. The blue line is the best fit parabola from the fourth iteration of steps 3-5. The green diamond is the final price chosen.

Limitations

This representation of trucker decision-making has several key limitations resulting from simplifications made to keep the overall modeling problem tractable. First and foremost, all the truckers on a given network link are treated as a single decision maker setting a uniform price for that link. To get a partial view of competition among truckers, some network node pairs are connected by multiple links (e.g. t1 and t2 between A and B in Figure 6-5). Relaxing this assumption would require modeling another layer of transactions between the logistics provider (i.e. a company that hires the individual trucks and the truckers themselves) and the truckers.

The truck links are assumed to have infinite capacity. Rather it is assumed that a sufficient number of trucks can be hired as needed. This guarantees that all shipment have a route even if rail links are at capacity. Having capacities on these links would require adding an additional decision to the truckers problem of determining a desired capacity and modeling the procurement of that capacity. Such a change in the scope of the trucker model was not considered but could be a fruitful area of future work.

The only pricing objective is to maximize profit in the next quarter. Alternative, more complex, pricing strategies are possible. For example, since there is a switching cost to change from one routing to another, a carrier might initially price at marginal cost to gain market share and then raise their price using the transaction cost to keep shippers from switching to other providers.

There is only a single quality of service offered. As was described at the start of this chapter, shippers have multiple objectives such as price, service reliability and timeliness. An alternative, more complex, pricing strategy might include multiple levels of service, e.g., an express service for customers who are willing to pay more for timeliness guarantees and a budget service for customers who are willing to absorb delays in return for a lower rate. This is left for future work.

6.4.3 Rail carriers

Rail carriers are modeled as providing scheduled service. Each day they run a chosen number of trains at equal intervals. Rail link capacity is modeled based upon average flows per day. This assumption eliminates the need to model the actual timing of individual shipper re-order requests during the three month contract period while insuring that, on average, the number of containers transported per day does not exceed the total number of cars that could be transported given the specified train frequency and train length. Table 6-2 shows the parameters specified for the railroad. The next section describes the cost model used for the rail carrier.

TABLE 6-2: Definition of parameters for railroad constituents

Parameter [units]	Definition	Sample Value
Link ID	Link(s) owned by the carrier	
Frequency [Trains/day]	Frequency of trains.	10
Number of engines [1/train]	Number of engines used per train	2
Train Length [cars]	Number of containers cars per train (does not include power)	100
Nominal Speed [mph]	Nominal speed of trains while running	45
Duty cycle [1]	Fraction of time that the train is operating at nominal speed; i.e., Nominal Speed * Duty cycle = Average speed of train along route	0.75
Price [$/TEU/mile]	Price charged to shipper on a per TEU shipped per mile basis	0.6
Fixed Cost Track [$/mile/day]	Fixed cost associated with upkeep of right of way per mile with straight-line depreciation.	15
Fixed Cost Moving Stock [$/car/day]	Fixed cost associated with owning sufficient rolling stock to achieve desired frequency with straight-line depreciation	50
Fixed Cost Power [$/engine/day]	Fixed cost associated with owning sufficient power to achieve desired frequency with straight-line depreciation	500
Fuel Cost Empty [$/car/mile]	Cost of moving an empty car one mile.	0.17
Marginal Fuel Cost Full [$/car/mile]	Incremental cost of moving of full car one mile.	0.07
Labor cost [$/hr]	Labor cost per hour for the train crew.	80
Backhaul time fraction [1]	Additional time required for backhaul expressed as fraction of the route travel time. Total travel time for an out and back journey is (2+Backhaul Fraction) * One-way travel time. This is used in computing the number of labor hours required to run the train.	0.3

Rail carrier cost model

Rail carrier costs are divided into three parts:

Fixed costs – These are incurred regardless of how many trains are run;

Service costs – These are dependent upon the number of trains run;

Marginal costs – These account for the actual number of filled cars on the trains that are run.

Fixed costs are assessed per mile of track per day. Based upon the 500 miles track length in the model, this results in a total fixed cost over the life of the simulation of 40.5 million dollars.[7]

Service cost are defined as follows (terms are defined in Table 6-2; the $TimePeriod$ is the contract period of 90 days):

$$\begin{aligned} serviceCost = {} & FuelCostEmpty * trackLength * trainLength * freq * TimePeriod + \ldots \\ & LaborCost * (2 + backhaulDelayFrac) * travelTime * freq * TimePeriod \end{aligned} \tag{6.1}$$

Given the figures in Table 6-2, an estimate can be made of the total cost (excluding fuel) for power and for moving stock. The baseline scenario total operating service cost is 259 million dollars over the 15 yr. simulation period. Finally, the marginal costs (i.e. fuel) for the baseline simulation case was 288 million dollars. This proportion of marginal costs vs. operating costs matches well with the OSCAR V (Operational simplified costing analysis for railways) model developed with support from the World Bank (Cripwell, 2001). in whose example case, variable costs were split nearly evenly between operating costs and fuel, and, fixed costs were about 1/5 of operating costs.

Rail carrier price and frequency setting

A rail carrier has two decision variables to choose per simulation round, train frequency and price. They proceed by finding a profit maximizing price using the same heuristic as the truck carrier for 5 different values of train frequency, their current frequency and ±1 and ±2 trains per day from their current frequency. They then choose the frequency and corresponding price that maximizes profit.

Figure 6-8 is a plot of rail carrier optimization in the baseline case. Note how as train frequency increases, so does profit up until five trains per day. After this, not enough new traffic is carried by an additional train to recoup its cost thus resulting in a decrease in profit when the sixth and subsequent trains are added.

[7] This is reasonable; however, finding references to justify it is difficult as such numbers are hard to come by and have many caveats attached. Rail projects vary so much in terms of financing that making apples-to-apples comparisons is challenging. On a per day per mile basis this number does compare well with Flodén (2007).

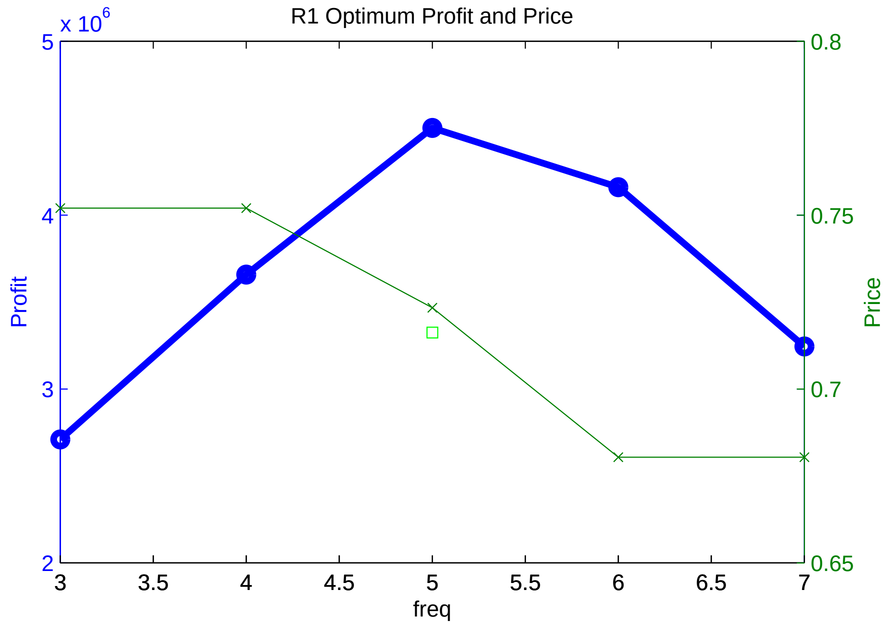


Figure 6-8: Example rail carrier price/freq optimization

Limitations

Modeling the pricing and operation decisions of a railroad is quite complex and so several simplifications were introduced. Actual movements of containers were not modeled, only the expected flow of containers per day. A side-effect of these simplifications is that when, for example, a truck arrives late and its container misses its train, it is assumed that the container is place on the next train, rather than having to wait for space on a future train. Real-world service of this type is often reserved and so the container would need to be kept in storage until a new reservation could be made. This assumption significantly reduces the expected delay in the terminal. To address this limitation and look at actual rail operations issues such as car handling and train formation, a stochastic simulation layer would need to be added to represent the actual (as opposed to expected) requests from shippers for service and track the movement of goods along the network as those request are fulfilled. Kwon et al. (1998) demonstrates the impact of varying these operational decision on overall network performance.

For costs of cars and locomotives, straight line depreciation from typical costs of such equipment is assumed. Also capital equipment acquisition costs and delays are not included; rather, the change in service frequency (and thereby car and locomotive needs) allowed per quarter is constrained to provide some limit/delay in equipment and labor acquisition.

Similar to the truck carrier model only a single class of service is offered with a simple tariff. Other authors, for example, (Hurley and Petersen, 1994), look at more complex, non-linear tariff schemes.

6.4.4 Intermodal terminals

As a matter of model scope, the representation of terminals is highly simplified. The design and operation of terminals is a key research area in the intermodal field. Often the success or failure of a particular intermodal route hinges upon the terminals along the routes. Key challenges include scheduling of operations, terminal layout for efficient handling and storage, implementing end-to-end IT to speed information flow, being able to respond to volatility of service demand while keeping high utilization and ownership structures of terminals (e.g. public vs. private vs. partnerships) (Stahlbock and Voss, 2008; Wiegmans et al., 2008; Rijsenbrij, 2008; den Hengst, 2008).

While decision making by intermodal terminal operators is not considered in the model, terminals are included as pass-through nodes along intermodal routes. Containers that pass-through a terminal are charged a processing fee and delayed by a fixed processing time (plus any additional delay waiting for their train). These parameters are listed in Table 6-3.

Table 6-3: Intermodal Terminal Parameters

Parameter [units]	Definition	Sample Value
Link ID	Virtual link that represents the terminal	
Processing Delay [hrs]	Time it takes to move one container through the terminal not including waiting	1
Price [$/TEU]	Cost charged for processing one container	10

6.4.5 The shipper's problem

Each shipper is managing supply of a good that is supplied at an origin point and demanded at a destination point. The model of total logistics cost described in Kwon (1994)[8] is used. The demand and trip time are stochastic and so the shipper maintains an inventory at the destination to reduce the likelihood of stock-out. When inventory levels fall below a specified trigger point, s, a specified quantity, Q, is ordered. After a lead time that accounts for order-processing and transportation delay, that quantity is added back into the inventory.

Kwon describes the computation of the total logistics cost, TLC, for such a resupply strategy. The TLC has the following components:

$$TLC(Q,s) = OrderCost + InventoryCost + ShortfallCost + TransportCost \tag{6.2}$$

[8] The development here follows Kwon with the exception of using a 90 day contract period (Kwon uses one year) and correcting some typographical errors in the equations.

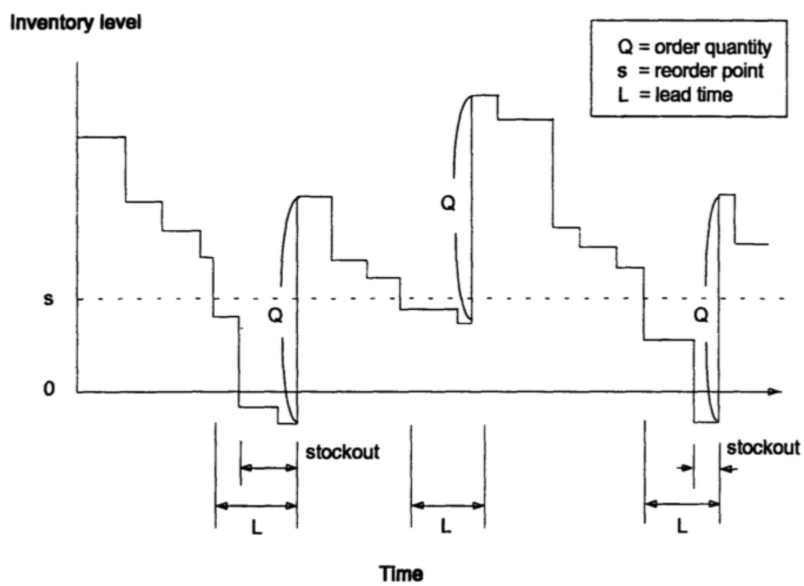


FIGURE 6-9: Trigger Inventory strategy (Kwon, 1994, figure 2.2, pg.23)

The Order Cost is the cost associated procuring each re-order and is charged once per replenishment. Inventory costs are the costs associated with holding the inventory on-hand and the in-transit inventory costs. Shortfall costs are the opportunity costs due to lost sales during periods of stock out. Finally, transport costs are the costs of moving goods from origin to destination on the specified transport route. The following notation is used to represent these costs.

Q	reorder quantity (TEUs)
s	reorder point (TEUs)
$\overline{D}$	expected demand during contract period (90 days)= $90\overline{d}$ (TEUs)
$\overline{d}$	expected daily demand (TEUs)
$\overline{L}$	average lead (transit) time (days)
$\overline{a}(s)$	expected shortfall per order cycle (TEUs)
A	order cost ($ per order)
V	per TEU value of the goods ($ per TEU)
W	inventory carrying cost (as percent of shipment value)
Y	in-transit inventory cost (as percent of shipment value)
K	per unit stock-out cost ($ per TEU)
R	per unit transport cost ($ per TEU)

Using this notation, the expression for estimated mean total logistics cost is (Kwon, 1994):

$$\overline{TLC} = \frac{A\overline{D}}{Q} + VW\left(\frac{Q}{2} + s - \overline{Ld}\right) + \frac{VY\overline{LD}}{90} + \frac{K\overline{a}(s)\overline{D}}{Q} + R\overline{D} \tag{6.3}$$

Under the assumption that daily demand and transport time are normally distributed with mean $\overline{d}$, $\overline{L}$ and variance $Var[d]$, $Var[L]$, the average stock-out is computed to be:

$$\overline{a}(s) = \sqrt{\overline{L}\mathrm{Var}[d] + \overline{d}^2\mathrm{Var}[L]}L'\left(\frac{s - \overline{Ld}}{\sqrt{\overline{L}\mathrm{Var}[d] + \overline{d}^2\mathrm{Var}[L]}}\right) \tag{6.4}$$

$L'(u)$ is the unit-normal linear loss integral, $L'(u) = f(u) - u\left(1 - \Phi(u)\right)$, where $f(u)$ is the probability density of the unit normal distribution and $\Phi(u)$ its cumulative distribution.

To verify the formulation, Figure 6-10 shows 2500 Monte Carlo trials of this reorder strategy. Fifty trials were conducted at each of 50 different levels of Q. The other parameters were as follows: s = 100, A = 500, D = 10,000, V = 5000, inventory and shortfall costs were 40% of the shipment value. Transport time was 0.8 days with a standard deviation of 0.13 days. Three months (90 days) of the using the strategy were simulated. The red-line is the analytic result from Kwon for the expected TLC and matches well with the Monte Carlo trials.

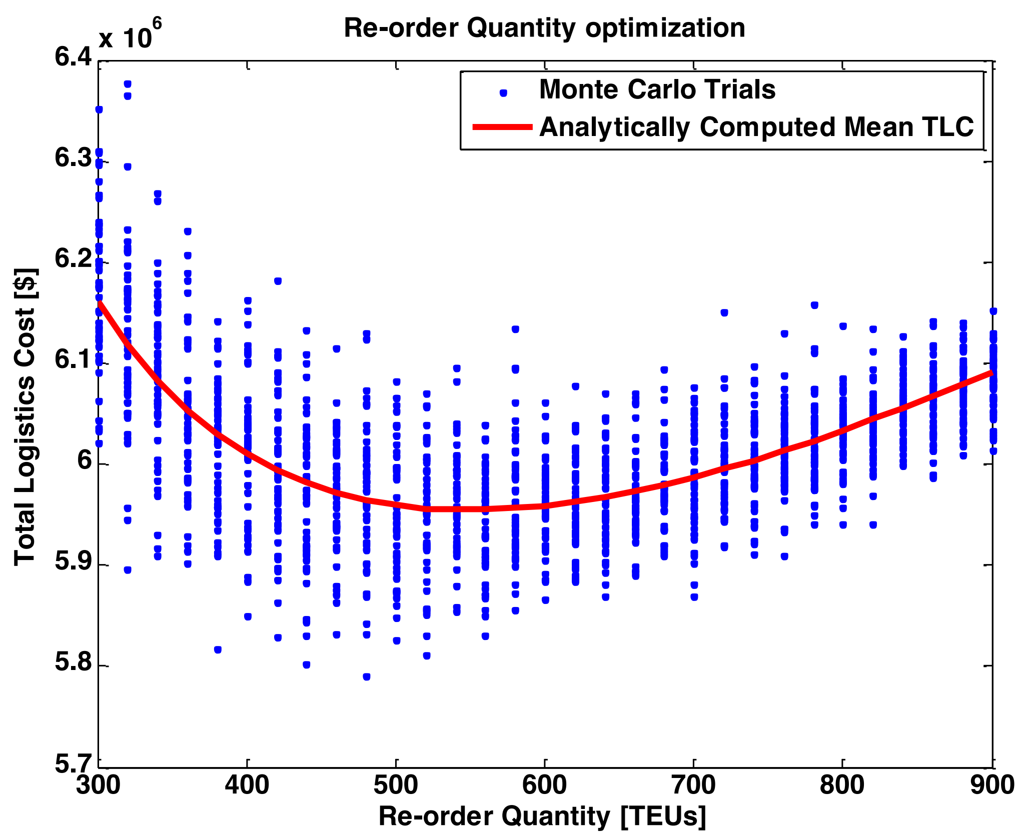


Figure 6-10: Total logistics cost

To minimize the total cost, TLC is differentiated with respect to s and Q and the derivatives set to zero. Solving for the optimal Q^* and s^* yields the simultaneous equations:

$$Q^* = \sqrt{\frac{2\overline{D}\left(A + K\overline{a}(s^*)\right)}{VW}} \tag{6.5}$$

$$1 - \Phi(s^*) = \frac{VWQ^*}{K\overline{D}} \tag{6.6}$$

An iterated procedure is specified in Kwon (1994) to solve for Q^* and s^*:

1. Solve for Q assuming $\overline{a}(s) = 0$
2. Solve $1 - \Phi(s)$ for s given the computed Q
3. Calculate $\overline{a}(s)$ for the new s under the normality assumption
4. Solve for Q' given the computed $\overline{a}(s)$ and s' given Q'
5. If $\left|Q' - Q\right| < \epsilon$ and $\left|s' - s\right| < \epsilon$, ϵ is small, then $Q^* = Q'$ and $s^* = s'$ otherwise go back to step 2.

In each simulation round, shippers proceed as follows:

1. Identify all routes available from the carriers to connect the desired origin and destination point.
2. For each route, compute a Q and s that minimizes total logistics cost using the procedure described above. If the new routing is different from the previous contract period, an additional transition cost is added in the form of an extra order transaction, A.
3. Place orders to use the transport service with the provider(s) that offer the lowest total logistics cost.
4. Should there be insufficient capacity on the lowest cost route, send the remaining demand via the next lowest cost routing. Contracts for transport are specified based upon average daily flow over the contract period. Actual flows will vary around that average.
5. Repeat the previous step until all traffic has been assigned to a route.

Baseline shipper parameters

In the baseline scenario the shipper parameters are set as follows. There are 50 shippers each of whom faces a mean demand, $\overline{D}$, of 2,000 TEU during each 90 day contract period. Demand is assumed normally distributed with standard deviation equal to 1/5 the mean, $\text{Var}[D] = \left(\overline{D}/5\right)^2$. Half of the shippers ship from O1 to D1, the other half from O2 to D2. Order costs, A, are 500 \$/order. The per-unit value, V, is chosen for each shipper randomly from \$20000 to \$100000 such that $\log V$ is uniformly distributed between $\log 20000$ and $\log 100000$. Inventory costs, W and Y, are chosen, for each shipper, from a uniform distribution between 20% and 40% of the per-unit value. Stock-out costs, K, are chosen for each shipper from a uniform distribution between 20% and 40% of the per-unit value.

Route formation, pricing and travel time

The route options available (and cost/quality of those options) to the shipper depend upon the prices and service offering provided by the carriers. Carriers only[9] sell point-to-point travel between nodes in the transport network. To form a route between non-adjacent nodes, shippers must purchase service for all the links along the routes. When there is a mode change (road to rail or rail to road), they must also pay for an intermodal transfer. This is as in the case of freight forwarding service described earlier in this chapter

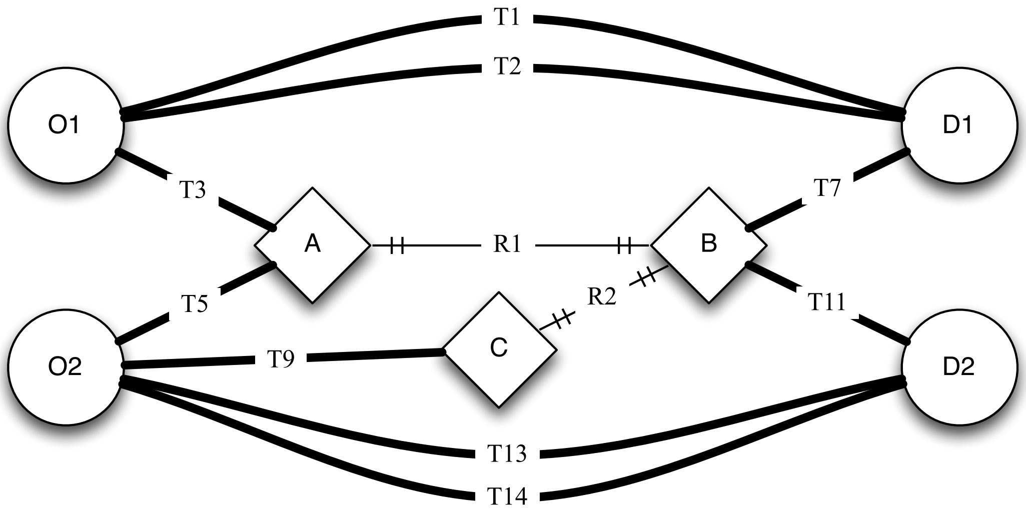


FIGURE 6-11: A simple intermodal network (Repeat of Figure 6-5)

Pricing for a route is simply the sum of the prices for each link along the route. For example, to travel from O1 to D1 via the railroad, R1, the total price would be:

$$p = p_{T3} + p_A + p_{R1} + p_B + p_{T7} \tag{6.7}$$

Travel time along the route is computed in a similar manner summing the time to traverse each link along the route. It is assumed that travel time for each link is independent of the other links. This assumption simplifies computation of travel time and variance of travel time. In the real-world, these times would not independent as it possible for disturbance that cause delays on one link may impact other links further down the chain. For example, a blizzard that delays road traffic might also impact rail movements. Under this independence assumption, the variance in travel

[9] This changes when one of the influence option is used and will be discussed when that option is introduced.

time for the route as a whole is also just the sum of the variances of travel time for each link. For truck links the standard deviation in travel time is assumed to be 2% of the mean travel time. For intermodal terminals processing time is assumed constant with zero variance. For the rail link, travel time has two components, the actual time spent traveling and the time spent waiting for the next train. They are assumed to be independent. Time spent actually completing the rail journey is assumed normally distributed with mean as per the train speed specified in Table 6-2 and standard deviation equal to 20% of that mean. Time spent waiting for the next train is assumed uniformly distributed over the headway between train where the headway is simply 24 hours divided by the daily train frequency.

Limitations

This formulation is responsive to three of the four key parameters identified by Danielis et al. (2005) as being of importance in shippers. They found via a survey of logistics managers in Europe that the most important factors in choosing a transportation option are, in order of importance, (a) price, (b) travel time, (c) reliability (i.e. variance in travel time), and (d) likelihood of damage. The TLC formulation includes terms for cost, travel time and, via the inclusion of loss due to stockout, variance in travel time. Damage/loss in route is not included. Note that there was variance observed by Danielis et al. between shippers from different sectors. The current model attempts to capture some of that sensitivity by varying shipment value and inventory costs. Shippers with higher inventory costs will be more sensitive to travel time than those that can afford to amass a large inventory and allow for long/highly variable travel times. Each shipper is only interested in a single commodity. Real shippers often need to coordinate multiple supply chains and may trade-off delays/expense in one chain for another.

6.4.6 Baseline model results

Given the above formulation, the baseline model results are as follows. Time series plots for model results are provided with a two period moving average smoothing. This removes an oscillation caused by the discrete steps taken when choosing potential prices. Using the curve fit mitigates this somewhat; however, some oscillation remains. Essentially, the exponential forecasting implies that constituents will have a delay in realizing a change in direction of the price movement of competitors. This means that they will tend to pick prices that are too low/high relative to their competitors and end-up oscillating as they overshoot and then compensate for competitor price changes. Results are shown in Figure 6-12.

The first set of plots, Figure 6-12(a), shows prices and rail frequency as a function of time.[10] The upper left chart is of rail prices in \$ per container-mile. Below that is plotted rail frequency in trains per day. To the right are truck prices. The upper-right chart shows prices of long-haul trucks. T1,

[10] Labels match Figure 6-5.

T2, T9, T13 and T14 are considered long-haul. They cover distances greater that 250 miles. The lower-right chart shows short-haul truck prices, i.e., links shorter than 250 miles. These are feeder roads in to/out of the intermodal terminals and are used by T3, T5, T7 and T11. The large number of overlapping lines on these plots may make them difficult to read; however, their main purpose is to show trends over time for each of these three groups of constituents—railroads, long-haul truckers and short-haul truckers.

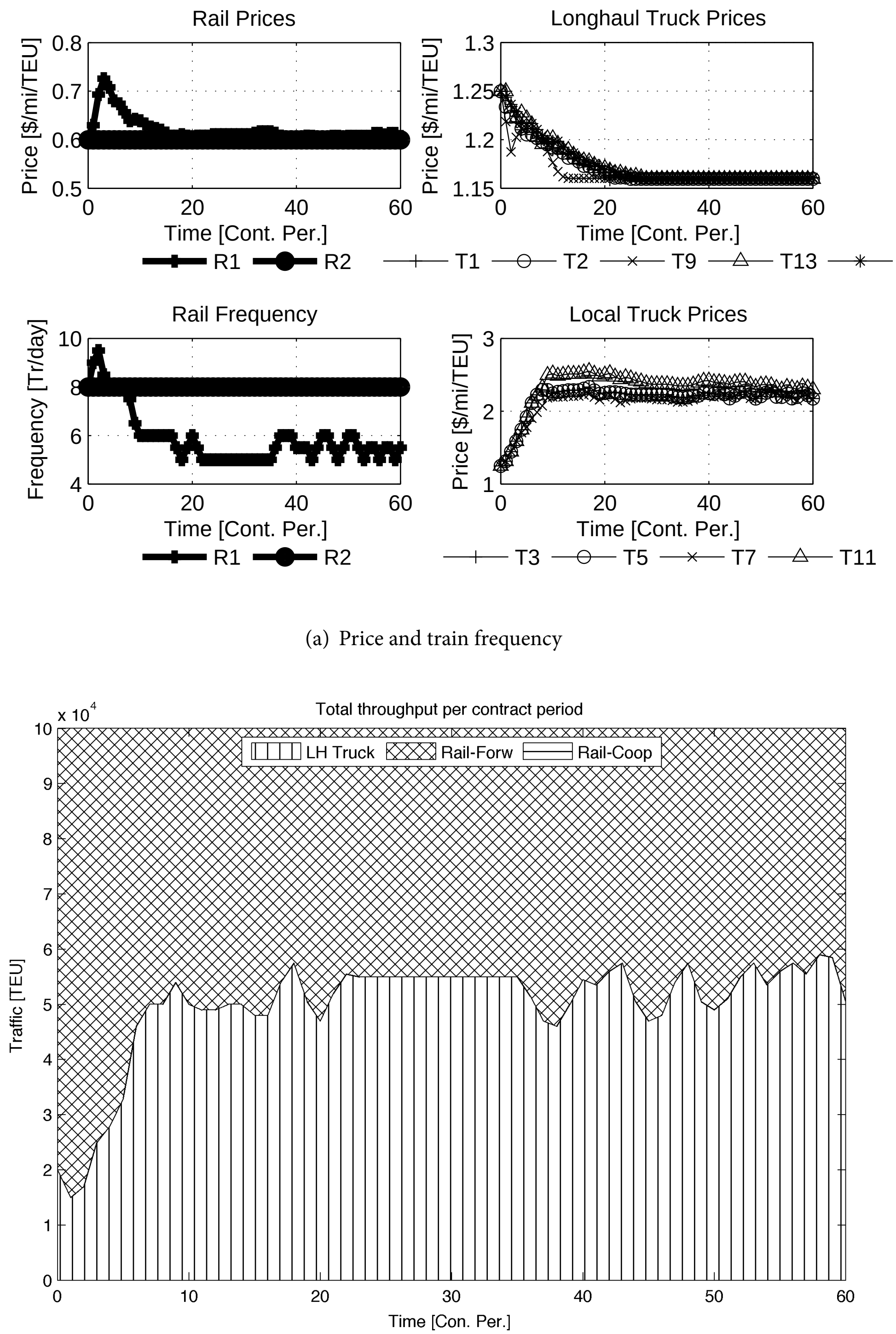
The second plot, Figure 6-12(b), shows the volume of traffic by route type for each quarter of simulation time. There are three possible route types: uni-modal long-haul truck direct from origin to destination shown with the vertical shading, truck-rail-truck intermodal routes formed via a forwarder shown with the cross-hatch shading, and truck-rail-truck intermodal formed via cooperative agreement between a long-haul trucker and a railroad shown with horizontal shading[11]. Simply put, the goal of the influencer is to make the vertically shaded portion of this graph as small as possible.

Looking at the baseline results, a few key dynamics can be observed. The A–B rail link prices appears to start low and then increase, overshooting a stable value to which they then return (upper left panel in Figure 6-12(a)). During the initial rise (t=1 to 5), there is also an increase in rail usage. Why would rail usage go up if the price of rail is also going up? The reason is that in the initial state, the cost of the alternative, i.e., longhaul truck is sufficiently larger that even if rail and the local trucks (the two carrier types needed in intermodal routes in the baseline case) raise prices as is seen here, they still gain market share. At the same time, the long-haul truckers are reducing prices eroding the pricing advantage from above. Eventually the two effects meet and, after about 10-15 periods, the division between traffic carried by long-haul truck and via truck-rail intermodal remains, with some oscillation, around a mean of ~46% rail intermodal. As this is happening, the A–B rail carrier is reducing their frequency (which reduces the number of cars and locomotives needed) to better match the realized demand. In contrast, the C–B rail carrier does not receive much traffic as they don't get the full cost benefit of rail using it for only half the move, and, is also hampered by the long and expensive O2–C road link. The C–B carrier cannot find a profit improving pricing/frequency solution within their search space and so keep a constant price and frequency.

6.5 Influence

Establishing the baseline behavior of the model ends the anticipation phase and provides a platform upon which the effect of various influences can be studied. The next several sections introduce four different influence strategies that could increase utilization of truck-rail intermodal services on the network. By comparing these strategies, insight can be gained into the effectiveness

[11] Traffic on cooperative routes does not appear here as such agreements are not allowed in the baseline case



(a) Price and train frequency

(b) Traffic volume by type

FIGURE 6-12: Baseline case

of various points of intervention by the influencer within the market. Direct incentives (taxes and subsidies to discourage/encourage certain behaviors), terminals (commonly a choke point limiting the performance and scaling of intermodal routes) and cooperative route formation are explored (an institutional change that enables integration). Using information as an influence mechanism was not considered in the case study and is left to future work.

6.5.1 Intermodal terminal

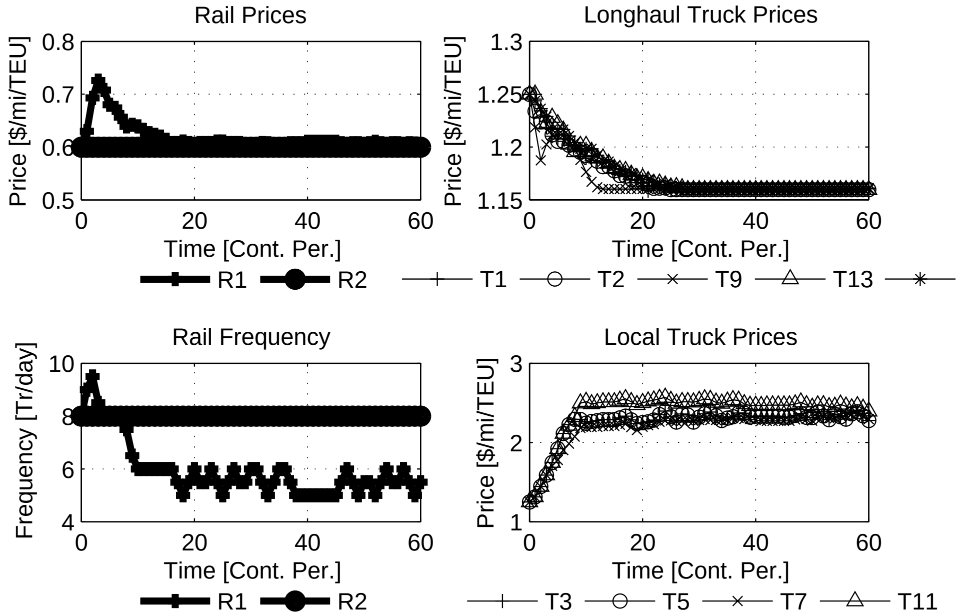
Intermodal terminals are key features of any intermodal transportation solution. They can be enablers by providing access between otherwise dis-connected modes, and also be inhibitors by becoming choke points along an intermodal chain.

Two different influences are considered with respect to intermodal terminals. The first addresses the additional cost borne by the shipper when needing to go through the terminal. This is simply a subsidy to offset a portion of that cost. This influence is an 'Incentive' in terms of the 5 Is[12]. At the start of the 20th quarter of the simulation (to allow initial settling of prices and service rates), a 50% subsidy is provided by the influencer, i.e. the government authority in this case, for all intermodal transfers. The simulation results are shown in Figure 6-13. The same plot convention as in Figure 6-12 are used.

There is little change in behavior when compared to the baseline case. A useful summary metric for comparing the behavior with the influence place vs. the baseline is the mean value of the percent of total traffic per quarter that uses rail after the influence is applied. This is simply the percent of traffic volume graph, Figure 6-13(b), that is not vertically shaded after t=20. Using this metric, rail usage increases from 46.6% in the baseline case to 47.3% with the subsidy applied. While this is technically a change in the intended direction, it is tiny. The reason for this is that, on a per container basis, the intermodal terminals only represent a very small percentage of the total shipping cost. Using the values in Table 6-1, 6-2 and 6-3 as examples, the rail portion of an intermodal move is priced at $300 per container, the truck portion is $200 per container and the intermodal transfers are only $20 per container. Reducing that $20 to $10 via subsidy does not materially impact the overall cost structure faced by shippers.

The second influence involves the travel time penalty incurred when switching modes. Current research in intermodal terminal operation has focused on improving throughput (Crainic and Kim, 2006). Estimate of up to a 50% improvement in productivity are projected with the implementation of large-scale automated container moving within the ports (Stahlbock and Voss, 2008). As this is a technological improvement, it is 'Infrastructure' in the 5 Is. To look at the effect of such an improvement, terminal travel time is decreased by 50% at time 20. The simulation results are shown in Figure 6-14.

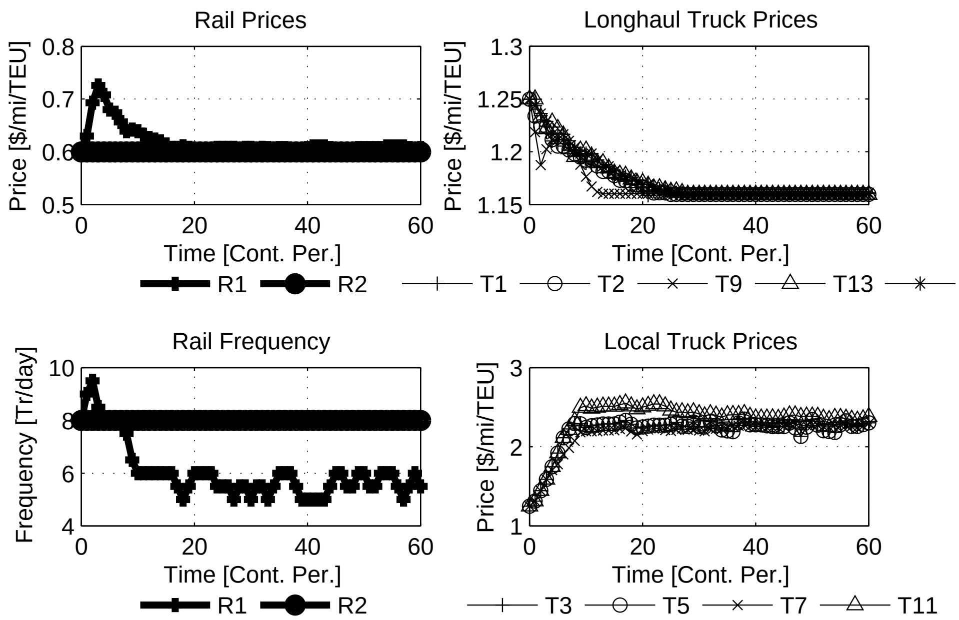
[12]The 5 Is are incentives, information, integration, infrastructure and institutions. See chapter 5



(a) Price and train frequency

(b) Traffic volume by type

Figure 6-13: A 50% subsidy on terminal costs is applied at t=20



(a) Price and train frequency

(b) Traffic volume by type

FIGURE 6-14: Time to process through a terminal decreased by 50% at t=20

Again there is little change when compared to base case. The share of traffic using rail increases by a very small amount, from 46.619% to 47.619% a change of only 1%. Once again the culprit is the relatively small impact of terminals on the overall performance of the chain. As above with cost, in terms of time, the mode-switching terminals represent only a small fraction of the total time incurred by a container on its journey. In the real-world, this time might be much longer if reserved service (rather than next available service as in the model) is used. With reserved service, a missed connection would result in a large additional wait in the terminal (compare missing a city bus to missing a flight and needing to fly stand-by). Modeling of reserved service, while not considered in this study, is recommended as a key refinement for future efforts on this problem.

A shorter transport time impacts shipper decision making via three of the costs listed in the total logistics cost equation (Equation 6.3). Inventory costs are computed based upon average inventory during the contract period. The shorter the transit time, everything else being equal, the higher the inventory costs since less product will be sold during the transit period (second term in Equation 6.3). Conversely, in-transit inventory costs are proportional to the transit time and so fall as transit time decreases (third term in Equation 6.3). Finally, shortfall costs (fourth term in Equation 6.3) decrease with decreased transit time; however, this effect is limited by s, the safety stock, i.e., the stock kept on hand to cover orders while new product is in transit. If s is much greater than the expected orders during the transit period then shortfall is very unlikely to occur thus making this third effect small. In the baseline case, shippers keep, on average, 1.5 days of safety stock, while average transit times are only 0.85 days ± 0.14 days. While stockout is not impossible (it would depend upon the variance in both travel time and demand), the data do suggest that, in the model, shippers are opting for strategies that use larger inventories thus minimizing the impact of changing the transit time. To put it another way, large inventories indicate that these shippers are able to use inventory to absorb variation in travel time and don't need 'just in time' delivery of new stock. While this is usually true for lower value goods, higher value goods tend to be more sensitive to inventory cost pressures. This effect is seen in the model as the shippers with higher value goods tend to preferentially choose a uni-modal road solution. Furthermore, the intermodal terminal transfers only represent about 10% of the total travel time and so reducing time spent in the terminal by 50% only translates into a 5% decrease in the overall time spent in transit.

6.5.2 Tax on road use

Janic (2007) argues that externalities are a significant piece of the costs associated with travel either via unimodal road or intermodal solutions. In the case of road travel, Janic claims that 20% of the total cost can be attributed to externalities. Conversely, in the intermodal case, only 6% of the cost is the from externalities. This represent a hidden cost advantage to the road carrier. Therefore, a tax on road use that accounts for this cost is explored as an influence mechanism to encourage rail use.

In terms of the 5 Is this is an 'Incentive' as it provides a penalty for an action (using roads) that the influencer does not want to have happen.

Imposing a tax does have the intended effect, reducing the miles traveled via road. Traffic volume (measured as described in the prior section on the baseline case) to rail inter-modal increases substantially from 46.6% to 61.5%. Of note is the pricing behavior of the various carriers in response to the tax. Long-haul truckers raise their price almost immediately after the tax is imposed by roughly 20% which is just enough to cover the new tax burden (sharp increase in price at t=20 in Figure 6-15(a) upper-right plot). They are still pricing just slightly above cost. As it is assumed that all shippers must ship, the effect of this change is to create pricing headroom for the other carriers. Now that long-haul truck is more expensive, railroads can, to a point, raise rates and still be the cheaper alternative. They do so and also increase service frequency allowing for both greater capacity and less travel time (Figure 6-15(a) left side plots).[13] short-haul truckers limit the ability of railroads to accomplish this price and frequency change. They also raise prices to try and capture the excess profit enabled by the long-haul truckers being forced to charge more (increase in price at t=20 in Figure 6-15(a) lower-right plot).

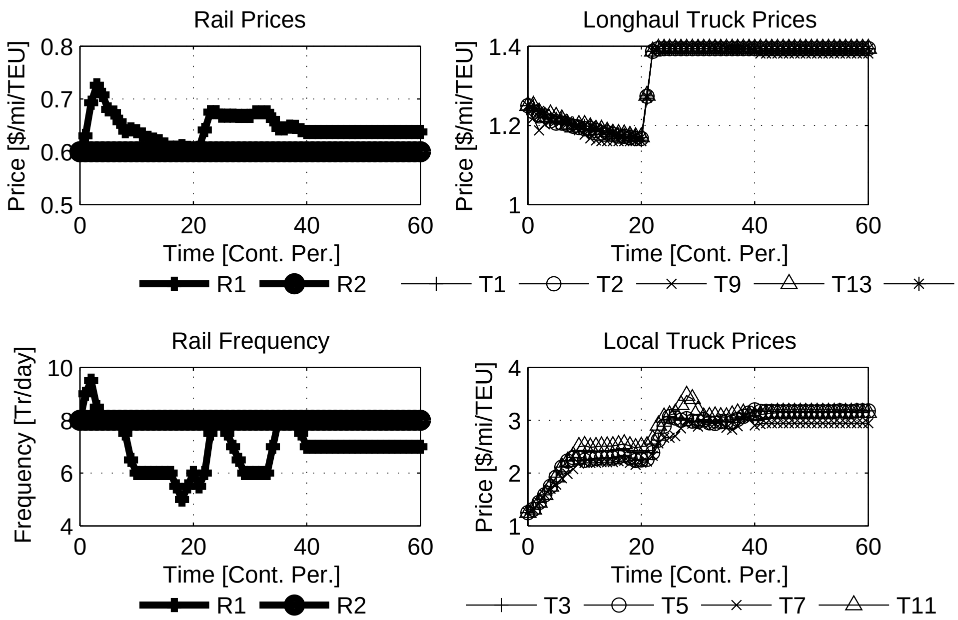
Figure 6-16 shows the sensitivity of the result to the tax rate used. The marginal effectiveness of the tax in moving traffic to rail intermodal is slowly declining as the rate increases. This is to be expected as the last shippers to continue using road despite ever higher prices caused by the tax burden are those who are at the high end of distribution of inventory costs and so are less sensitive to the transport price increase than they are to the decrease in service quality incurred when switching to rail intermodal.

6.5.3 Cooperative routes

As stated in the introduction to this chapter, intermodal routes can be formed by freight forwarders who act as middle-men between shippers and carrier or by cooperative agreement between the carriers involved. The baseline scenario only allows for forwarder formed intermodal routes and does not include the possibility of cooperative routes. The ability to form such routes is implemented as an institutional influence. Modeling the formation of such cooperative routes is an open area of research in intermodal freight transport. As explained below, the Nash bargaining solution is used to determine if cooperative routes are formed. When this influence mechanism is in place, a two step procedure is used to determine if a cooperative route will be offered and, if so, what will be its price and how will the revenue generated be divided among the participating carriers.

Formation of cooperative routes is restricted to coalitions between rail carriers and long-haul truckers as only the long-haul truckers have services at both the origin and destination points. Only

[13] Increasing train frequency reduces the delay between train thereby reducing the typical journey time overall.



(a) Price and train frequency

(b) Traffic volume by type

Figure 6-15: At t=20, a 20% tax is introduced per mile of road travel

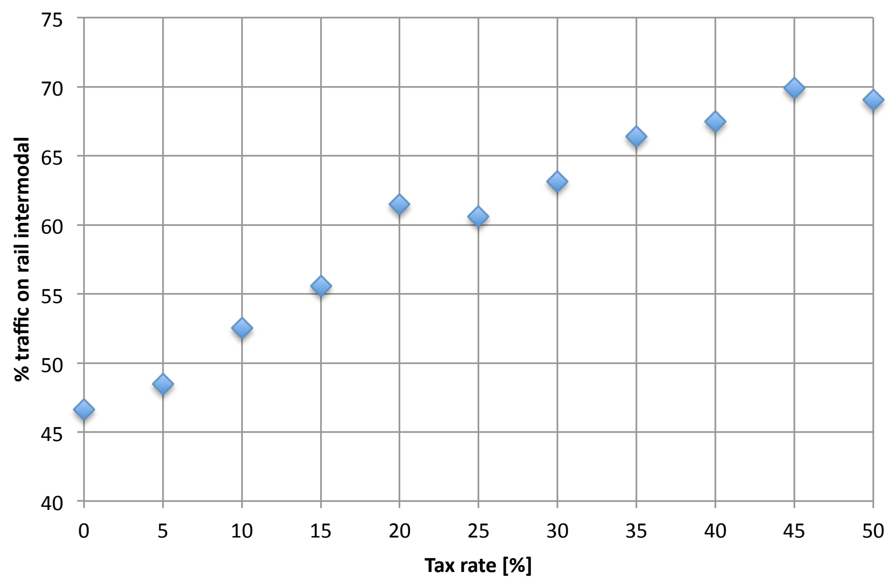


FIGURE 6-16: Sensitivity to tax rate

two-party coalitions are considered due to the computational complexity of three-plus party negotiations. As a consequence of this limitation, intermodal terminal services are purchased at market rates. Once formed, coalitions are kept in place for 4 contract periods (one year) and then renegotiated. Coalition formation occurs as follows. In doing the computation, the prices and services offering of parties that are not involved are forcasted as in the price optimization for the carriers acting alone.

1. The price of the prospective cooperatively formed route is computed using a procedure identical to that for price optimization for truckers with the exception that historical price and profit data are not used in the curve fitting step as the prospective route does not have a price history (see subsection 6.4.2). The objective is to maximize the overall profit generated by the route (total revenue less total cost incurred by all mode operators involved in the route). Fixed costs are allocated to the route as per the percentage of the flow handled by the carrier that arrives via the prospective route. For example, if for the rail carrier, total fixed costs are 10,000 and 10% of traffic comes via the cooperative route then 1,000 of fixed costs are accounted for when computing the joint profit of the route. Figure 6-17 shows a sample run of this calculation. The jagged red line represents the true profit that the cooperative route would generate for the carriers (amortizing fixed costs as described above). Its jagged shape comes from the loss of profit that occur when a shipper switches to another routing option as price is increased. The blue circles are the point estimates of profit from step 5 in subsec-

tion 6.4.2. The blue parabola is the best fit curve from step 6 in subsection 6.4.2. Finally, the green diamond is the chosen profit maximizing price.

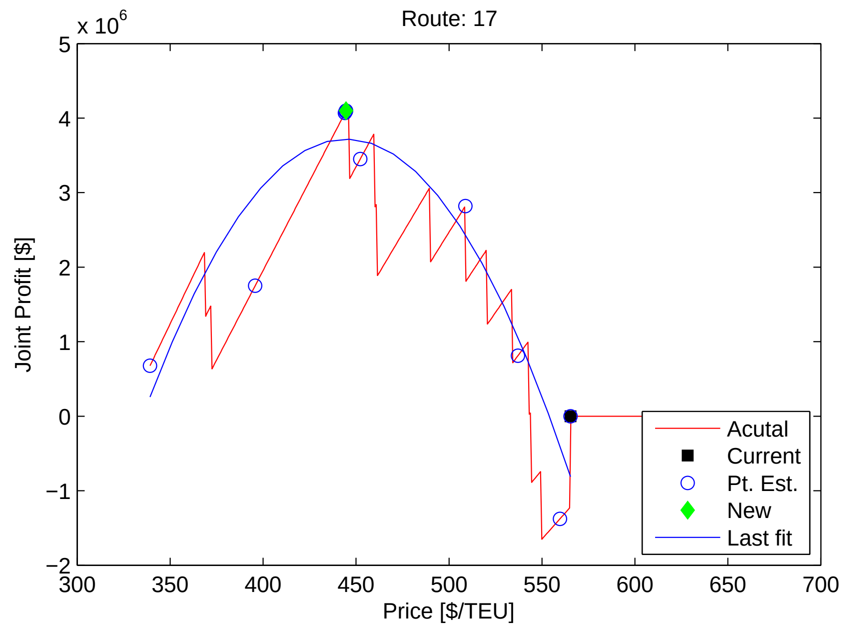


Figure 6-17: Finding a price for cooperative route #17 between T1 and R1

2. Should a price that results in a positive joint profit be found, then the division of revenue from the route is determined using the Nash bargaining solution (Nash, 1950). Nash's bargaining solution allows identification of Nash equilibria in two-party bargaining games. The solution is computed as follows:

   (a) The profit for each party (railroad and truck) in the absence of the agreement is computed. This pair (π_r^d, π_t^d) is known as the disagreement point.
   (b) The revenue generated by the prospective route (less costs paid for terminal use), R, is computed using the price arrived at above. This is the revenue available to be split. Note that R is invariant with how the revenue is split between the parties.
   (c) Let r_r and r_t be the revenue from sources other than the cooperative route received by the railroad and trucker, respectively, when the cooperative route is in place. Let k_r and k_t be the total costs incurred by the railroad and trucker, respectively, when the cooperative route is in place. Let τ be the proportion of the revenue from the cooperative route given to the trucker $(1 - \tau)$ is given to the railroad. Then the total profit for the trucker is given by $\pi_t = \tau R + r_t - k_t$ and, for the railroad, $\pi_t = (1 - \tau)R + r_r - k_r$
   (d) The Nash bargaining solution (as formulated for this situation) states that if the following optimization problem has a solution, τ^*, then that solution is a Nash equilibrium

solution to the negotiation over splitting the revenue. If one party offers τ^* as the split, then the other party's best response is to also offer τ^*.

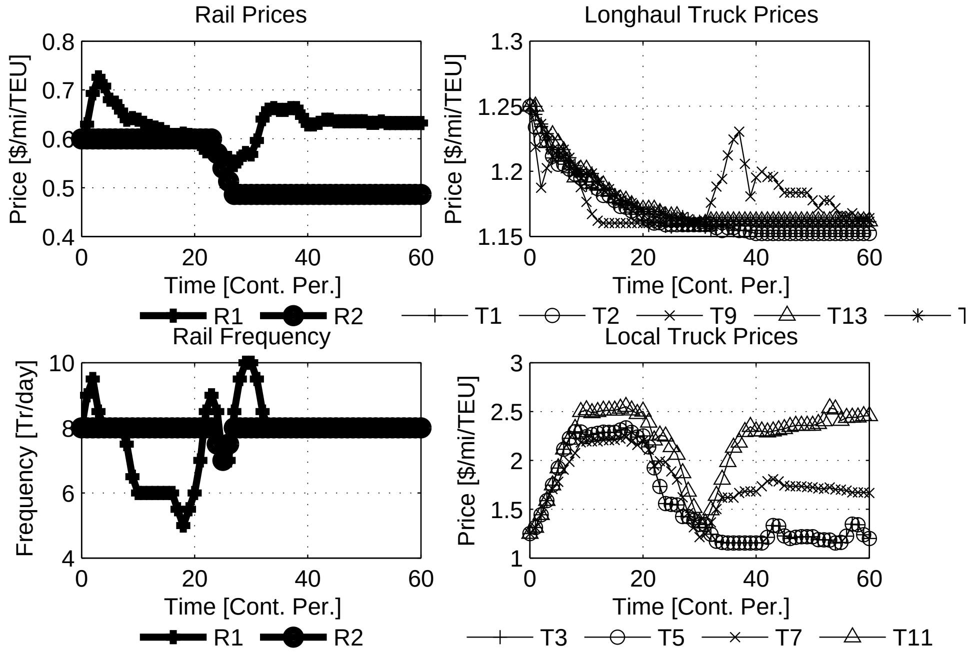
$$\begin{aligned} \max_{\tau} \quad & \left(\overbrace{\tau R + r_t - k_t}^{\pi_t} - \pi_t^d \right) \left(\overbrace{(1-\tau) R + r_r - k_r}^{\pi_r} - \pi_r^d \right) \\ \text{subject to} \quad & 0 \le \tau \le 1, \quad \pi_t \ge \pi_t^d, \quad \pi_r \ge \pi_r^d \end{aligned} \tag{6.8}$$

(e) The agreed split, τ^*, is kept for one year with the price of the cooperative route being adjusted as per the procedure step 1 above every quarter. After the year expires, the revenue split is renegotiated.

Limitations of cooperative agreements formulation

The Nash bargaining solution has several assumptions behind it that make it a potentially unrealistic[14]. Nash bargaining is essentially a static analysis and does not look at the offers and counter-offers that would occur in negotiation. While Nash still reveals the outcome of the bargaining, the procedure that led to that outcome is hidden. Rubinstein (1982) provides a model that looks at bargaining strategy not just outcome. Further alternative formulations of the bargaining problem are possible. The division of revenues could be handled via the Shapley value for the coordination game amongst the three (or more) parties (e.g. including the terminal operator(s) or a local trucker at either end of the route). The Shapley value (Shapley, 1953) is a way of equitably distributing the payoff gained by a coalition of players in a game such that players' shares reflect their importance in obtaining the payoff for the coalition (Binmore, 2007a). It is arrived at by imagining the coalition forming one player at a time and that player demanding to receive the incremental increase in the payoff that results from their participation and then averaging over all possible permutations of the order in which players join. It has certain properties that make it appear to be an equitable distribution method. This approach has been looked at before in Samet et al. (1984) for cost allocation in the transportation network flow problem in operations research. However, this work has not been well cited going forward and is worth renewed consideration as the problem of divisions is widely cited as a key struggle for intermodal freight transport (see review in Bontekoning et al., 2004). A more promising analogue is the Shapley value based cost allocation for interconnected Internet Service Providers developed by Ma et al. (2007). Key questions to be addressed include: How does one convince the parties to adopt the Shapley split? How does one ensure that the information provided to estimate contribution of each party to the coalition is accurate?

[14] While a full discussion of these assumption are beyond the scope here, they are listed here: efficiency – the whole value to be split is distributed; independence of irrelevant alternatives – the removal of an option that neither player would have chosen does not effect the outcome; symmetry – if the players utility functions are identical, they receive equal shares; and independence to a linear transformation (scaling) of the utility functions (Nash, 1950).



(a) Price and train frequency

(b) Traffic volume by type

FIGURE 6-18: Cooperative routes formation allowed at t=20

Results of allowing cooperatively formed routes are shown in Figure 6-18. Examining Figure 6-18(b) reveals that allowing cooperatively formed routes between long-haul truckers and rail roads has a dramatic impact on route choice by shippers. A substantial share of the market (78.65%) shifts towards using intermodal rail—either via the cooperative routes (horizontal shading in the figure) or via the forwarder formed routes (cross-hatch shading). The reason for this shift is that, with the introduction of the cooperative routes, there is finally competition for the short-haul truckers who, before the introduction of cooperative routes, had a local monopoly on access between the intermodal terminals and the origin and destination points (see Figure 6-5). Evidence for this can be found in the sharp price drop seen in the lower right panel of Figure 6-18(a). Such a price drop is not seen in any of the other cases. This further argues for the conclusion that it was the high prices of the local truckers that kept traffic away from the intermodal routes. Allowing cooperative routes opened up an alternative means to get to/from the intermodal terminals thereby breaking the local monopoly.

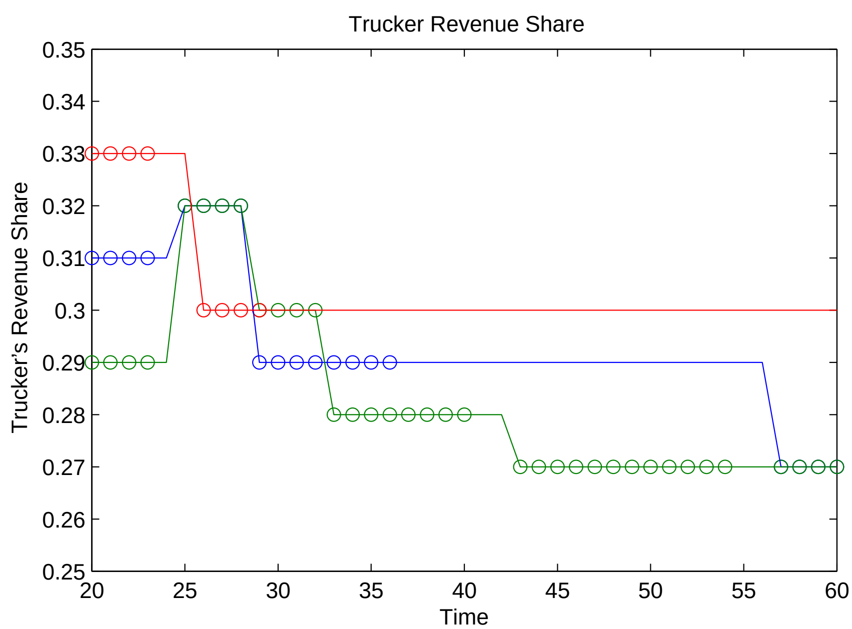


Figure 6-19: Share of cooperative route revenue going to the trucker

Looking more closely at the agreements being formed, one can see that, over time, as the market adjusts to the agreements, they become (relative to the non-cooperative and uni-modal) less profitable. Therefore, after t=30, there is less traffic on the cooperative routes, as some of the agreement are not renewed. One can see a shift of less revenue share on the cooperative routes going to the truckers. Figure 6-19 shows the revenue share to truckers for the agreements made by railroad R1 (i.e. the railroad between terminals A and B in Figure 6-5).

Overall, one can conclude that cooperative routes, while an effective means of shifting traffic onto road-rail intermodal routes, are sensitive to changes in the market brought on by their existance. The agreements depends upon the balance between what the carriers can get going-it-alone vs. cooperating. As the market adjusts to the presence of the agreements, this balance shifts. In particular, it would appear that the railroads are able to demand a larger share of revenue as more traffic moves their way (Figure 6-19). This dynamic sensitivity is reflective of the importance of capturing the feedback between constituent actions and SoS behavior in the AIR framework.

6.6 Reaction

The reaction phase in the AIR framework consists of the response of SoS constituents to influence applied by an SoS influencer. Since anticipation is imperfect, it is unlikely that the result of an influence will exactly match expectations. Rather, constituents will implement changes to the extent that they view the requests embodied in the influences as in their best interest. Some insight can be gained into this by looking at the revenue, cost, profit and market share results for each type of constituent (trucker and railroad) for each of influences described above (Table 6-4).

Table 6-4: Aggregate model results; Totals taken from t=20 to 60; Revenue, costs and profits are listed in billions (10^9) dollars.

	Truck Revenue	Railroad Revenue	Truck Cost	Railroad Cost	Truck Profit	Railroad Profit	Uni-modal Truck %	Inter-modal %
Baseline	2.454	0.937	2.119	0.991	0.335	-0.054	**53.381**	**46.619**
	Total	**3.391**	**Total**	**3.111**	**Total**	**0.281**		
Term. Subsidy	2.463	0.945	2.105	1.000	0.358	-0.056	**52.667**	**47.333**
	Total	**3.408**	**Total**	**3.106**	**Total**	**0.302**		
Term. Speedup	2.446	0.948	2.100	1.007	0.346	-0.059	**52.381**	**47.619**
	Total	**3.394**	**Total**	**3.107**	**Total**	**3.39**		
Road Tax	2.613	1.168	2.084	1.149	0.530	0.019	**38.500**	**61.500**
	Total	**3.781**	**Total**	**3.232**	**Total**	**0.549**		
Co-op	1.780	1.291	1.576	1.264	0.264	0.026	**21.345**	**78.655**
	Total	**3.071**	**Total**	**2.840**	**Total**	**0.291**		

The revenue columns of the table are the total revenue earned by all the carriers of a given mode over the period from t=20 to 60, i.e., when the influences were active. Note that since all the revenue comes from the shippers, the sub-totals indicated in the revenue columns reflect the total transportation cost paid by the shippers and is thus a surrogate measure for their view. The costs columns are the total costs incurred by the carriers over the same time period. The profit columns are simply revenue less cost. Finally, the market share columns provide the percentage of traffic volume that used uni-modal and inter-modal routes respectively.

For the truckers it would seem that, counter-intuitively, the preferred situation is that of the tax on road use as that maximizes their profit. The reason for this is that the portion of traffic that moves from long-haul truck to intermodal now has less of its travel attributed to road. This reduces the cost on truckers to a greater extent than the tax increases it. So, in balance, the truckers actually have less cost than in the baseline case. This reflects the tax having the desired effect of moving traffic to use rail for the long-haul portions of the journey. Note however that truck revenue also increases. This reflects a profit grab by the short-haul truckers increasing prices beyond that needed to cover the tax (increase in prices in the lower right graph in Figure 6-15(a)). Essentially, the short-haul truckers make more additional money with tax in place than the long-haul truckers lose and so the combined profit of all the truckers increases. In net, shippers end up paying more when the tax is in place.

For the railroads, the preferred alternative is the cooperative routes case. This makes sense as it shifts the most traffic onto intermodal routes. The shippers would also seem to prefer this alternative as they face the minimum total transport cost. The truckers, on the other hand, have a decrease in profit and may therefore be against this.

Given the different impact that each influence has on the constituents involved, implementation of one or more influences will depend upon how the constituents react. For the present case, reaction cannot be observed directly as there is no real system upon which the intervention are being imposed. Rather, a historical approach is used to examine past intermodal freight networks and suggest potential reactions that constituents may exercise in response to interventions. First the role of coalitions and alliances between constituents and influencer(s) is considered.

Tuimala and Lukka (1999) studied the strategic alliances in the intermodal freight transport network that serves southern Finland. They cite the role of alliances as supporting and fostering specialization among the members of intermodal chains. As alliances are strengthened and can be relied upon by chain members, they are more likely to adapt, that is specialize, their part of the chain to most efficiently match their partner's operations. They further found that while the benefit of alliances are well known, establishing and maintaining them can be difficult. Through a series of interviews with relevant members of different intermodal chains, they found that most of the alliances in their study area were operational in nature and lacked longer-term strategic stability. Operators are wary that alliances may remove flexibility that they need to respond to variability in future traffic volume/type/flows and consequent changes in the market.

Song (2002) takes a broader perspective looking at cooperation among terminals (in this case terminals are ports as the intermodal network being study connect maritime and land-based transport links). In looking at Hong Kong and its transport/trade relationship to neighboring provinces in southern China (in particular Schenzen), it is found that both cooperation and competition play a role in the intermodal network. Song observes that competition exists within the market in Hong Kong and in Southern China, while cooperation is more prevalent for arrangements that go between the two areas. He argues that forming alliances can be an effective way to combat market power that may exist for certain chain participants. Other participants can form coalitions to balance out the market. In the case study presented in this chapter, one sees this effect when the long-haul truckers and railroads cooperate to form routes that break the local monopoly of the short-haul truckers.

In the above cases, cooperation was found to be a key element to running the intermodal network. Enablers for cooperation included support from government authorities and the establishment of strategic partnerships to allow joint planning of future activities in addition to current operations. In reference to the Finnish case (Tuimala and Lukka, 1999), cooperation led to specialization and a consequent improvement in efficiency since partners could rely on each other to perform their respective functions. A possible reaction therefore to introducing the cooperative routes as in the current case would be for the cooperating constituents to change their operations to realize similar efficiency gains. Such an effect was not included in the current model, but is a fruitful area for future research. Such improvements will likely result in an even stronger movement to shippers using intermodal routes. Sometimes influencers need only take a limited role in fostering SoS change by establishing the conditions (for example, via the 5 Is) to support it and then can (and likely should) step back and let the constituents transform the SoS in line with their local needs.

Further examples of coordination mechanisms were reported by Van Der Horst and De Langen (2008). They looked 12 different real-world examples of coordination between members of intermodal chains connecting the port of Rotterdam to its surrounding hinterland. They identified four different coordination types of arrangements: "(1) Introduction of incentives, (2) Creation of an interfirm alliance, (3) Changing scope, and (4) Creating collective action" (Van Der Horst and De Langen, 2008). These correspond reasonably well to the 'Incentives', two different approaches to 'Integration, and, 'Institutional' change from the 5 Is. 'Interfirm alliances' refers to integration where the coordinating entities remain separate with clear boundaries of responsibility while creating a joint offering to the market. This is what occurs in the cooperative routes influence examined earlier. 'Changing scope', on the other hand, involves actually combining entities into a wholly new organizational structure as in vertical integration of chain members and the introduction of a chain manager. Van Der Horst and De Langen (2008) found that while changing scope results in tighter coordination and consequently larger efficiency gains, establishing and maintaining such coordination was difficult given the diversity of cultures and experiences among the involved chain members. While it may have been in the best interests of the firms to integrate, actually getting them to do so was difficult. Less permanent interfirm alliances that allowed for adjustment such as in that formed under the cooperative routes influence proved to be more robust both in terms of formation and sustainment.

6.7 Future work

Future work and extensions to this case study are presented below in two categories: model improvements and case study scope/scale improvements.

6.7.1 Model improvements

The overall structure of the model is governed by a quarterly decision cycle on which the shippers and carrier make their decision. In reality, of course, these decision are not so tightly coordinated. As the age/visibility of information is a relevant factor in decision making, allowing these decisions to be made at different times (and therefore with different information sets) could yield new dynamics.

Focusing in on truck carriers, several areas for improvement exist. First, there is the assumption that truck links are of infinite capacity. This is, of course, unrealistic. Replacing this assumption with a model of truck company operations would allow better representation of the costs (especially as they evolve over time) faced by truckers. A challenge in implementing this is the potentially large number of truckers who would need to be modeled. This could be mitigated by grouping truckers into fleets; however, that would introduce another layer of decision making, i.e., fleet management into the model.

Rail carrier modeling is also simplified. One key assumption that should be relaxed is that containers are placed on the next available train. As discussed in subsection 6.4.3, this assumption removed the need to model train formation and car/locomotive management. Adding those decisions would significantly expand the scope of the model and thus could require a different computational approach to be taken. However, that would also allow the railroad to offer different classes of service. Currently, railroads such as BNSF use class of service to segment their market by the varying needs of their customers. This allows, for example, a premium to charged to a customer who is more time sensitive. Real-world intermodal service pricing (especially for high value or time-critical goods) takes advantage of the ability to offer multiple classes of service. More complex tariff structures than the simple fixed price per mile per container used here would allow carriers to further differentiate their market.

Given the key role of intermodal terminals indicated in the literature, they should be treated in more detail and represented as active agents in the model. This would require adding a new set of decision makers to the model. Existing work in modeling terminal operations can be starting point for such an extension (Rizzoli et al., 2002). Initially this should target capturing the functional performance of the terminal in terms of a few simple parameters such as capacity and throughput delay to allow examination of congestion—a key issue in terminals. Not only would this improve the technical realism of the model, it would also enable additional influences such as congestion pricing (Hensher and Puckett, 2005). Congestion modeling could also be applied to the road links.

The most obvious limitation of the shipper model is the assumption that shippers always ship. There should be a demand model added that adjust the demand shippers faces based upon their prices (which in turn would need to connected to the costs they face). Furthermore, other types of shippers trying to solve other freight transport problems should be included. They may have different preferences on the service quality characteristics described earlier.

Finally, only one set of shippers and carriers was examined. While the parameters used were realistic, a sensitivity analysis should be performed.

6.7.2 Case study scale and scope

Currently the transport network under consideration is quite small. Real-world networks are much larger and often involve orders of magnitude larger number of actors. Simply adding additional actors to the existing model would not work as the computational requirement scale supra-linearly. For example, the number of evaluations of potential routes required scales as the product of the number of shippers and the number routes. One could look estimating group behavior at the cost of losing the ability to look at specific individuals. Such aggregated flows were used by Arnold et al. (2004).

In terms of study scope, additional transport modes could be added such as air freight. Additional traffic types such as passenger rail and bulk freight could be added as well to better represent real demands on a rail network. Real-world intermodal service also involves less-than-truckload (LTL) service. Expanding the model scope to include that would involve representing firms that combine LTL shipments into containers as active agents. Their requests of the transport network might be quite different from the existing shippers.

Finally, a key assumption in the model is that agents (shippers, carriers, etc.) solve a profit maximizing decision problem. Recent work by Meijer et al. (2012) has examined using actual human decision makers in a simulation/game setting similar to the this case study. Insights from such experiments could lead to the formulation of more accurate representation of agent decision making.

6.8 Implications with respect to transportation

From a transportation perspective, the case study results support the notion that external market interventions such as taxes and subsidies can be less effective that mechanisms that exploit self-interest such as allowing cooperation. The strong effect of cooperation is consistent with empirical studies of intermodal freight transport networks (Van Der Horst and De Langen, 2008). Of course such mechanisms may not always be available, but when extant, they should be carefully considered. As formulated, the case study model is quite simplified and so its results should be taken as behavioral and not predictive beyond showing potential trends. It can be extended to include a, larger, more realistic route network and more varied shipper and carrier populations such as in (Flodén, 2007) allowing better characterization of the effectiveness of the proposed strategies. Costs of implementing the influences were not considered as evaluating the cost of an institutional change in a way that is comparable with a technological change is difficult. Nash bargaining is only one approach to look at cooperation between constituents. Other game-theoretic approaches such as those developed in (Daniel Cooksey and Mavris, 2011) could be particularly useful as they can be applied to modeling the participation/cooperation decision in other SoS.

6.9 Implications for SoSE

The case study has several implications also for SoSE practitioners. First, with respect to problem formulation, the intermodal freight transport case is a good reminder that the SoSE challenge is not just technical in nature. As was seen in the case study, utilization of rail intermodal depended as much on the technical performance of the involved constituent systems as it did on economic decisions such as pricing and cooperative route formation. Therefore, as was supported by the literature in chapter 2, SoSE requires an expanded set of issues to be considered when compared to SE. Some of these new issues are evident in the case study. For example, terminal improvement while effective in a technical sense (transport times were decreased) did not have the desired impact overall. Shippers had already adjusted their inventory management to allow for longer transport times making the overall impact of the infrastructure improvement negligible. As decision making in SoS is distributed and not necessarily coordinated, such a situation arising in other SoS is entirely plausible. A well-intentioned SoS engineer might make a change, only to see its impact defused via the decisions of others. For example, in the case of the DoD document production system (Krygiel, 1999), underlying software elements were upgraded on a regular basis to keep them current. Deployment was incremental and had much localized control. Incompatibilities surfaced and the result was a need to develop systems that worked with the lowest common denominator infrastructure thus negating the the ability to take advantage of the upgrades. This is an example of a reaction to institutional influences requiring both implementation of new technologies while still maintaining compatibility.

A second implication for SoSE from the case is the role of agent-based modeling as a tool for understanding SoS. By using the agent-based approach, detailed analysis on individual constituents, such as that seen in the discussion of influence results above, could be done. The key idea to be incorporated from agent-based modeling into SE for use in SoSE is the importance of representing not just the behavior of the systems involved, but also the decision taken by the agents who control those systems, i.e., the decision making process of each constituent. A consequence of this more holistic formulation of the system model is that non-technical means of changing the SoS can be considered. Only one of the influences examined involved a technology change. All the others used various non-technical pressures and encouragements to change constituent decisions and induce the desired behavior from the SoS. Such influences are not typically within the purview of the systems engineer. Collaboration with the social scientists and political economists will be needed to fully understand how such approaches can be included in the SE process.

Having explored the use of the AIR framework in practice, the next and final chapter summarizes the key contributions made in the thesis and provides suggestions for future work.

Chapter 7

Summary, Contributions and Future Work

7.1 Summary

Recent decades have witnessed the emergence of a new class of systems engineering challenges that arise for systems of systems. As was demonstrated in chapter 2, through both empirical case example and supporting literature, distribution of decision making among a set of constituents presents a unique challenge creating and maintaining value delivery from an SoS. This review motivated two research questions:

> *What are the feedback relationships between the constituents and SoS influencers, and how do their influences result in changes in the constituent individually and the SoS as a whole?*

and

> *What approaches can be used by external SoS influencers to cause constituent decision makers to change constituent systems so as to induce a desired behavior from the SoS?*

Inherent to distributed decision making are the feedbacks introduced in chapter 3 between constituent decision makers and their respective systems. These feedbacks formed the basis of a new framework along with a new role—that of the SoS influencer introduced in chapter 4. The influencer has a unique role within systems engineering in that he can effect certain parts of the system he is trying to 'engineer' indirectly, i.e., via a set of self-interested third parties. Only through influences such as those introduced in chapter 5 can change those parts of his system (of systems).

Finally, the case study, chapter 6, revealed some of the real-world issues in implementing AIR and the 5 Is. Building models which contain multiple actors making decisions over multiple time scales present a unique set of challenges. Compromises must be made to account for limits on available information about constituents, the context and even the underlying physics of problem at hand. Even the highly simplified model presented was computationally taxing requiring nearly 50,000 optimizations of total logistics costs just for a single influence case. However, the case study did reveal several insights that may generalize to other SoS. Namely the fragility of cooperation (players not in the coalition can adjust their choices to make the coalition no longer viable) and the importance of ensuring that incentives are targeted at aspects of constituents' decision problems that are most valuable to the constituents.

7.2 Key contributions

The key contributions of this thesis are:

The AIR Framework A novel representation of decision making within an SoS that captures the indirect relationship between SoS influencers and constituent systems in their SoS.

The 5 Is An initial set of strategies that can be used to influence constituents.

They are now examined within the context of the research questions.

In regard to the first research question, a descriptive framework, known as AIR, for decision making in SoS was proposed. The framework involves three phases anticipation, influence and reaction. Anticipation involves gaining an understanding of the SoS and its constituent. This understating needs to extend past simply characterizing SoS behavior to also capturing the objectives, capabilities and constraints of the constituents whose decisions can change SoS behavior. The influence phase concerns an attempt to change the choices being made by the constituents so as to result in an SoS whose behavior is more in-line with the influencer's objectives. Finally, the reaction phase refers to the simple fact that both anticipation and influence are unlikely to be perfect and constituents may respond to influences in other ways than making the desired changes. The commonly used existing taxonomy of collaborative, virtual, directed and acknowledged SoS (Maier, 1999; Dahmann and Baldwin, 2008) was mapped into the framework. It was then shown that all four classes differ only by how the constituents decision makers interact with the influencer(s).

Building upon this descriptive foundation, a basis set of influence strategy types were proposed by treating constituent decision making as a value maximizing process. These types, known as the 5 Is, are Incentives, Information, Integration, Infrastructure and Institutions.

'Incentives' is rewarding/penalizing constituents for particular behavior that they would not do otherwise. 'Information' refers to providing constituents information to change the priors they use to make decision under uncertainty. 'Integration' is the re-assignment of particular SoS components to different constituents. A common example would be combining two systems into one. 'Infrastructure' refers to introducing new technology into the SoS. An example would be a new high-speed data network to facilitate higher bandwidth inter-connection between constituents. Finally, 'institutions' refer to the rules and regulations that constituents and influencer follow.

The use of the AIR framework is demonstrated in a case study of an intermodal freight transportation network. The purpose of the case study is to demonstrate a process that an SoS influencer could use to change the performance of the SoS via changes in constituent behavior. In doing so, first the intermodal freight transportation problem was characterized using the AIR framework and transportation literature. Underutilization of intermodal rail service was identified as a key concern in support of improving issues such as environmental pollution and road congestion. To gain a better understanding of this SoS problem, an example transportation system that uses both rail and truck routes was modeled using a agent-based simulation that represents shipper and carrier decision making over a 15 year period. The model was then used to examine several intervention strategies based upon the 5 Is. Different strategies can have vastly different impact on the constituents even though they produce similar behavior in the SoS. This result reinforces the need for SoS influencers to consider the effect of influences on constituents locally, not just on the SoS as whole.

These contributions, the AIR framework and 5 Is, fill a significant gap in the SoS Engineering literature. While the existing frameworks describing SoS identify the multi-stakeholder, multi-layer decision making structure as a key issue in SoS, they do not provide much proscriptive guidance to the systems engineer as to how to handle such a situation. The roles, interactions and processes described in AIR capture in a succinct form the key structures needed to understand SoS behavior where the constituent set and value proposition (at both the constituent and SoS levels) are fixed.

7.3 Impact upon system engineering practice

Having defined AIR and demonstrated its use, it is proposed that the AIR framework and 5 Is can have significant impact on systems engineering practice. They provide a simple, consistent representation of the key roles decision makers take in an SoS. At the highest level, these are the constituent and the influencer. While the notion of constituent is not new, the notion of an 'influencer' is novel. More often than not (e.g. managing a communication or transportation network) system of systems engineers find themselves in this influencing role and can only indirectly effect the constituent systems within the SoS. Traditional systems engineering is predicated on the ability of the highest level stakeholder to dictate requirements which determine decisions making at the lower levels. Such an approach would not work in SoS when there was a conflict between the needs of the influencer and that of the constituents. Rather strategies that account for the local needs of the constituents are required. The 5 Is are a useful first step towards developing such strategies.

As is demonstrated in the case study, counter-intuitive results can occur when attempting to intervene in systems of such significant decision-making complexity. Therefore modeling such as the agent-based approach used in the case is crucial to gaining a sufficient understanding of the SoS before intervening. Examples of this are replete in case studies of real SoS (Krygiel, 1999). When trying to modernize document production in the DoD, the need of for common standards was identified. In implementing these standards problems arose given the diverse areas in which the standards needed to be applied. Furthermore, making such changes without disturbing on-going operations was quite challenging. Even though the end-state was much better than the status quo, there was a need to ensure local buy-in to make the transitions happen.

7.4 Limitations and extensions

There are several significant limitations and opportunities for extension of AIR and the 5 Is. With respect to the AIR framework, one must keep in mind that AIR, on its own, is not sufficient for managing an SoS. It is best used in the context of broader frameworks such as those cited section 2.4. AIR only helps formulate strategies for changing constituent behavior. It does not aid in determining what the desired constituent behavior should be. That is the design problem of SoS and evidence of progress towards it has been documented by Crossley and Nusawardhana (2004). Simulation and modeling of SoS is required for the implementation of AIR. Progress in that area has been documented in Sloane et al. (2007b); Biltgen and Mavris (2007); Dagli and Kilicay (2007).

As developed thus far, AIR assumes a fixed constituent set. Changing this would require modeling a super-set of potential constituents and their respective life-cycles. In addition, scaling the agent-based modeling approach demonstrated in the case-study to very large numbers of constituents can be computationally challenging. For such large numbers, constituents may need to be represented as member of a class whose behavior is characterized statistically instead of considering individuals. System dynamics can be helpful in such a situation as was shown in (Shah et al., 2007b) where multiple satellite operators were aggregated.

Determining the costs associated with influence mechanism, especially those that are social in nature is quite challenging. Research from political and organizational science should be used in assessing such costs and managing trade-offs between constituents that arise during the reaction phase. Finally, influence strategies were discussed in isolation and were implemented as such in the case study. In reality, they will likely need to be used in combination to achieve the desired effect. How to form such combined strategies is an area for future research.

7.5 Other research areas that can be used to extend AIR and 5 Is

Several promising areas for extending AIR and 5 Is using other on-going research in the SE community are discussed below.

There are currently several on-going efforts to apply model-based systems engineering (MBSE) techniques to SoS. For example, the Comprehensive Modelling for Advanced Systems of Systems

(COMPASS) project in the EU is attempting to develop a comprehensive representation language and associated tools for SoSE (Woodcock et al., 2012). While such a language would be quite useful in easing communication and documentation within SoSs, these effort are in an early state and none have reached wide-spread adoption. AIR isn't meant as substitute for the MBSE efforts, which aim to create a much broader representation. It is suggested that the MBSE community should try to incorporate the three activities of anticipation, influence and reaction into the their work.

AIR assumed a fixed set of constituents. Recent work is making strides in describing how new constituents join SoS (Baldwin et al., 2012). Given reasonable boundaries on the set of potential constituents to consider, the participation decision could be included under the set of variables that a constituent controls. Prior to joining they would still exist in the shared SoS context (Shah et al., 2007a) but would not, as of yet, have any interfaces with constituents that are already participating. The shared context might be an avenue for discovery (both of the SoS by the potential constituent and vice versa).

In the intermodal freight transport case study, only one-time, permanently applied influences were considered. The real world examples, by contrast, used multiple influences in concert. As was suggested by Ames et al. (2011), the formulation of true dynamic strategies that adjust influences based upon reaction and post-facto feedback will be needed. How should one formulate such strategies? Is there a structured manner in which it could be done?

7.6 Conclusion

Systems of systems represent a new and challenging area of research for the systems engineering community. History demonstrates that with each passing year, as the world becomes more interconnected, most if not all systems will be SoS or will be constituents within SoS or both. Such a situation calls into question the foundational assumptions that underlie traditional systems engineering. Distribution of decision making as seen in SoS is particularly challenging. Systems engineers and architects will need to find ways to not only manage this aspect but leverage it to foster new, emergent SoS. They will need to become SoS influencers. AIR and the 5 Is are some of the first tools they can use in that new role.

References

AAR (2010). *Railroad Facts 2010*. American Association of Railroads.

Ackoff, R. L. (1971). Towards a system of systems concepts. *Management Science*, 17(11):pp. 661–671. doi:10.1287/mnsc.17.11.661.
URL http://mansci.journal.informs.org/content/17/11/661.abstract

AF/SAB (2005a). Domain integration. Tech. Rep. SAB-TR-05-03, United States Air Force Scientific Advisory Board.

AF/SAB (2005b). System-of-systems engineering for Air Force capability development. Tech. Rep. SAB-TR-05-04, United States Air Force Scientific Advisory Board.

Ames, A. L., Glass, R. J., Brown, T. J., Linebarger, J. M., Beyeler, W. E., Finley, P. D., Moore, T. W. (2011). Complex adaptive system of systems (CASoS) engineering framework version 1.0. SAND Report SAND 2011-8793, Sandia National Labs.

Anson, P., Cummings, D. (1991). The first space war: The contribution of statellites to the gulf war. *The RUSI Journal*, 136(4):pp. 45–53.

Arnold, P., Peeters, D., Thomas, I. (2004). Modelling a rail/road intermodal transportation system. *Transportation Research Part E: Logistics and Transportation Review*, 40(3):pp. 255–270.

Baldwin, W. C., Ben-Zvi, T., Sauser, B. J. (2012). Formation of collaborative system of systems through belonging choice mechanisms. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 42(4):pp. 793 –801. doi:10.1109/TSMCA.2011.2164522.

Ballis, A., Golias, J. (2002). Comparative evaluation of existing and innovative rail–road freight transport terminals. *Transportation Research Part A: Policy and Practice*, 36(7):pp. 593 – 611. doi:10.1016/S0965-8564(01)00024-6.

Berry, B. J. L. (1964). Cities as systems within systems of citles. *Papers in Regional Science*, 13(1).

Beyeler, W. E., Glass, R. J., Bech, M., Soramäki, K. (2006). Congestion and cascades in payment systems. Federal Reserve Bank of New York Staff Reports No. 259.

Biltgen, P. T., Mavris, D. N. (2007). Capability-based quantitative technology evaluation for systems-of-systems. In *SoSE '07*, pp. 1–6.

Binmore, K. (2007a). The shapley value. In *Playing for Real*, pp. 530–531. Oxford University Press.

Binmore, K. (2007b). Taking charge. In *Playing for Real*, chap. 20. Oxford University Press.

Bjelkemyr, M., Semere, D., Lindberg, B. (2007). An engineering systems perspective on system of systems methodology. In *1st Annual IEEE Systems Conference*, pp. 1–7.

Blanchard, B. S., Fabrycky, W. J. (2005). *Systems Engineering and Analysis*. Prentice-Hall International Series in Industrial and Systems. Prentice-Hall, 4th edn.

Boardman, J., Sauser, B. (2006). System of systems – the meaning of of. In *IEEE/SMC International Conference on System of Systems Engineering*, pp. 118–123.

Bodeau, D. (1994). System-of-systems security engineering. In *Computer Security Applications Conference, 1994. Proceedings., 10th Annual*, pp. 228 –235. doi:10.1109/CSAC.1994.367304.

Boehm, B. (2000). *Spiral Development: Experience, Principles, and Refinements*. University of Southern California.

Boehm, B., Lane, J. A. (2006). 21st century processes for acquiring 21st century software-intensive systems of systems. *Crosstalk*.

Bontekoning, Y. M., Macharis, C., Trip, J. J. (2004). Is a new applied transportation research field emerging?—-a review of intermodal rail–truck freight transport literature. *Transportation Research Part A*, 38:pp. 1–34.

Brandimarte, P., Zotteri, G. (2007). *Introduction to Distrubution Logisitcs*. Wiley-Interscience.

Cammons, D., Tisserand, J. B., III, Williams, D. E., Seise, A., Lindsay, D. (2005). *Operations*, vol. 1 of *Network Centric Warfare Case Study: U.S. V Corps And 3rd Infantry Division (Mechanized) During Operation Iraqi Freedom Combat Operations (Mar-Apr 2003)*. U.S. Army War College.

Caris, A., Macharis, C., Janssens, G. K. (2008). Planning problems in intermodal freight transport: Accomplishments and prospects. *Transportation Planning and Technology*, 31(3):pp. 277–302.

Carlock, P. G., Fenton, R. E. (2001). System of systems (sos) enterprise systems engineering for information-intensive organizations. *Systems Engineering*, 4(4):pp. 242–261.

Charles, D. (2006). TR35: Paul Rademacher: The man who opened up the map. *Technology Review*. URL http://www.technologyreview.com/tr35/Profile.aspx?TRID=437

Chatterjee, L., Lakshmanan, T. R. (2008). Intermodal freight transport in the United States. In R. Konings, H. Priemus, P. Nijkamp, eds., *The Future of Intermodal Freight Transport*, chap. 3. Edward Elgar Publishing Ltd.

Chattopadhyay, D., Ross, A. M., Rhodes, D. H. (2008). A framework for tradespace exploration of systems of systems. In *CSER 2008*.

Chen, P., Clothier, J. (2003). Advancing systems engineering for systems-of-systems challenges. *Systems Engineering*, 6(3):pp. 170–183.

Collens, J. R., Krause, B. (2005). Theater battle management core system: Systems engineering case study. Tech. rep., Center for Systems Engineering AFIT.
URL https://acc.dau.mil/CommunityBrowser.aspx?id=37601

Cowley, S. (2005). Level 3, cogent resolve peering dispute, renew deal. *COMPUTERWORLD*.
URL http://www.computerworld.com/s/article/105790/Level_3_Cogent_resolve_peering_dispute_renew_deal

Crainic, T. G., Kim, K. H. (2006). Intermodal transportation. In C. Barnhart, G. Laporte, eds., *Handbooks in Operations Research & Management Science: Transportation*. North Holland.

Crainic, T. G., Laporte, G. (1997). Planning models for freight transportation. *European Journal of Operational Research*, 97(3):pp. 409 – 438. doi:10.1016/S0377-2217(96)00298-6.

Cripwell, J. P. (2001). OSCAR V (operational simplified costing analysis for railways). In *SARA Workshop*.
URL http://www.railwaycosting.com/presentations/oscoscds/oscoscds.htm

Crossley, W. A., Nusawardhana, M. M. (2004). Variable resource allocation using multidisciplinary optimization: Initial investigations for system of systems. In *10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*. AIAA/ISSMO.

Cunningham, J. (1991). The role of satellites in the gulf war. *The Journal of Practical Applications in Space*, 2(3):pp. 45–69.

Dagli, C., Kilicay, N. (2007). Understanding behavior of system of systems through computational intelligence techniques. In *1st Annual IEEE Systems Conference*, pp. 1–7.

Dahmann, J. S., Baldwin, K. J. (2008). Understanding the current state of us defense systems of systems and the implications for systems engineering. In *IEEE Systems Conference*. IEEE.

Daniel Cooksey, K., Mavris, D. (2011). Game theory as a means of modeling system of systems viability and feasibility. In *Aerospace Conference, 2011 IEEE*, pp. 1–11. IEEE.

Danielis, R., Marcucci, E., Rotaris, L. (2005). Logistics managers stated preferences for freight service attributes. *Transportation Research Part E*, 41:pp. 201–215.

DeLaurentis, D., Callaway, R. K. (2004). A system-of-systems perspective for public policy decisions. *Review of Policy Research*, 21(6):pp. 829–837.

den Hengst, M. (2008). Interorganizational coordination: the role of information technology. In R. Konings, H. Priemus, P. Nijkamp, eds., *The Future of Intermodal Freight Transport*, chap. 13. Edward Elgar Publishing Ltd.

DeWitt, W., Clinger, J. (2000). Intermodal freight transportation. In *Transportation in the new Millenium*. TRB.

DiMario, M., Boardman, J., Sauser, B. (2009). System of systems collaborative formation. *Systems Journal, IEEE*, 3(3):pp. 360 –368. doi:10.1109/JSYST.2009.2029661.

DoD (2008). *Systems Engineering Guide for System of Systems, Version 1.0*. Office of the Under Secretary of Defense for Acquisition, Technology and Logistics (OUSD AT&L).

Eisner, H. (1994). A systems engineering approach to architecting a unified system of systems. *Systems, Man, and Cybernetics, 1994. 'Humans, Information and Technology'., 1994 IEEE International Conference on*, 1:pp. 204–208 vol.1. doi:10.1109/ICSMC.1994.399837.

Eisner, H., Marciniak, J., McMillan, R. (1991). Computer-aided system of systems (s2) engineering. *Systems, Man, and Cybernetics, 1991. 'Decision Aiding for Complex Systems, Conference Proceedings., 1991 IEEE International Conference on*, 1:pp. 531–537. doi:10.1109/ICSMC.1991.169739.

Eliot, Christian (2005). Planning for the global earth observation system of systems (GEOSS). *Space Policy*, 21(2):pp. 105 – 109. doi:10.1016/j.spacepol.2005.03.002.

Ellison, R., Linger, R., Longstaff, T., Mead, N. (1999). Survivable network system analysis: a case study. *Software, IEEE*, 16(4):pp. 70–77.

Ellison, R. J., Fisher, D. A., Linger, R. C., Lipson, H. F., Longstaff, T. (1997). Survivable network systems: An emerging discipline. Tech. Rep. CMU/SEI-97-TR-013, CMU SEI.

Ellison, R. J., Goodenough, J., Weinstock, C., Woody, C. (2008). Survivability assurance for system of systems. Tech. Rep. CMU/SEI-2008-TR-008, CMU SEI.

Fernández L., J., de Cea Ch., J., Soto O., A. (2003). A multi-modal supply-demand equilibrium model for predicting intercity freight flows. *Transportation Research Part B: Methodological*, 37(7):pp. 615 – 640. doi:10.1016/S0191-2615(02)00042-5.

Flodén, J. (2007). *Modelling Intermodal Freight Transport*. Ph.D. thesis, Göteborg University.

Forsberg, K., Mooz, H., Cotterman, H. (2005). *Visualizing Project Management: Models and Frameworks for Mastering Complex Systems*. Wiley, 3rd edn.

Fowkes, A., Nash, C., Tweddle, G. (1991). Investigating the market for inter-modal freight technologies. *Transportation Research Part A: General*, 25(4):pp. 161–172.

Gambardella, L. M., Rizzoli, A. E., Funk, P. (2002). Agent-based planning and simulation of combined rail/road transport. *SIMULATION*, 78(5):pp. 293–303. doi:10.1177/0037549702078005551. URL http://sim.sagepub.com/content/78/5/293.abstract

Gibbons, R. (1992). *Game theory for applied economists*. Princeton University Press.

Goldman, A. (2005). The cogent-level 3 dispute. *ISP-Planet*. URL http://www.isp-planet.com/business/2005/cogent_level_3.html

Google, Inc. (2009). Google maps/google earth APIs terms of service. URL http://code.google.com/apis/maps/terms.html

Gourevitch, P. (1978). The second image reversed: the international sources of domestic politics. *International Organization*, 32(4):pp. 881–912.

Hensher, D. A., Puckett, S. M. (2005). Refocusing the modelling of freight distribution: Development of an economic-based framework to evaluate supply chain behaviour in response to congestion charging. *Transportation*, 32(6):pp. 573–602.

Holland, B. (2005). In a conductorless orchestra, players vote with their bows. *The New York Times*. URL http://www.nytimes.com/2005/05/12/arts/music/12orph.html

Hsieh, M.-H. (2007). *Standards as Interdependent Artifacts: The case of the Internet*. Ph.D. thesis, MIT Engineering Systems Division.

Hurley, W. J., Petersen, E. R. (1994). Nonlinear tariffs and freight network equilibrium. *Transportation Science*, 28(3):pp. 236–245.

INCOSE (2006). Systems engineering handbook v.3. Tech. Rep. INCOSE-TP-2003-002-03, INCOSE.

Jamshidi, M. (2005). System of systems engineering—a definition. In *IEEE SMC 2005*.

Janic, M. (2007). Modelling the full costs of an intermodal and road freight transport network. *Transportation Research Part D*. doi:10.1016/j.trd.2006.10.004.

Karcanias, N., Hessami, A. G. (2011a). System of systems and emergence part 1: Principles and framework. In *Fourth International Conference on Emerging Trends in Engineering & Technology*. IEEE Comp. Soc. doi:10.1109/ICETET.2011.51.

Karcanias, N., Hessami, A. G. (2011b). System of systems and emergence part 2: Synergetic effects and emergence. In *Fourth International Conference on Emerging Trends in Engineering & Technology*. doi:10.1109/ICETET.2011.52.

Keating, C., Rogers, R., Unal, R., Dryer, D., Sousa-Poza, A., Safford, R., Peterson, W., Rabadi, G. (2003). System of systems engineering. *Engineering Management Journal*, 15(3):pp. 36–45.

Keeney, R., Raiffa, H. (1993). *Decisions with multiple objectives: Preferences and value tradeoffs*. Cambridge Univ Press.

Khalsa, S., Nativi, S., Geller, G. (2009). The GEOSS interoperability process pilot project (ip3). *Geoscience and Remote Sensing, IEEE Transactions on*, 47(1):pp. 80 –91. doi:10.1109/TGRS.2008.2007304.

Kotov, V. (1997). Systems of systems as communicating structures. Tech. Rep. HPL-97-124, Hewlett-Packard.

Krag, T., Büettrich, S. (2004). Wireless mesh networking. *O'Reilly Wireless Dev Center*. URL http://www.oreillynet.com/pub/a/wireless/2004/01/22/wirelessmesh.html

Krasner, S. D. (2010). Revisiting "The Second Image Reversed". In *A Conference in honor of Peter Gourevitch*.

Krygiel, A. J. (1999). *Behind the Wizards's Curtain: An Integration Environment for a Systems of Systems*. CCRP.

Kwon, O. K. (1994). *Managing Heterogeneous Traffic On Rail Freight Networks Incorporating The Logistics Needs Of Market Segments*. Ph.D. thesis, MIT Civil and Environmental Engineering.

Kwon, O. K., Martland, C. D., Sussman, J. M. (1998). Routing and scheduling temporal and heterogeneous freight car traffic on rail networks. *Transportation Research Part E: Logistics and Transportation Review*, 34(2):pp. 101 – 115. doi:10.1016/S1366-5545(97)00022-7.

Kyamakya, K., Jobman, K., Meincke, M. (2000). Security and survivability of distributed systems: an overview. In *MILCOM 2000. 21st Century Military Communications Conference Proceedings*, vol. 1, pp. 449–454. doi:10.1109/MILCOM.2000.904993.

Le, L., Donohue, G., Chen, C. (2004). Auction-based slot allocation for traffic demand management at hartsfield atlanta international airport: A case study. *Transportation Research Record*, 1888:pp. 50–58.

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., Postel, J., Roberts, L. G., Wolff, S. (2003). A brief history of the internet. Tech. rep., Internet Society.
URL http://www.isoc.org/internet/history/brief.shtml

Lukasik, S. J. (1998). Systems, systems of systems, and the education of engineers. *Artif. Intell. Eng. Des. Anal. Manuf.*, 12(1):pp. 55–60. doi:10.1017/S0890060498121078.

Ma, R. T., Chiu, D., Lui, J. C., Misra, V., Rubenstein, D. (2007). Internet economics: The use of shapley value for isp settlement. In *CoNEXT'07*. ACM.

Magee, C. L., de Weck, O. L. (2004). Complex systems classification. In *Fourteenth Annual International Symposium of the International Council On Systems Engineering (INCOSE)*.

Maier, M. W. (1999). Architecting principles for systems-of-systems. *Systems Engineering*, 1:pp. 267–284.

Manthorpe, W. H. J., Jr. (1996). The emerging joint system of systems: A systems engineering challenge and opportunity for apl. *Johns Hopkins APL Technical Digest*, 17(3).

Meijer, S. A., Mayer, I. S., van Luipen, J., Weitenberg, N. (2012). Gaming rail cargo management. *Simulation & Gaming*, 43(1):pp. 85–101. doi:10.1177/1046878110382161.

Mekdeci, B., Ross, A. M., Rhodes, D. H., Hastings, D. E. (2011). Examining survivability of systems of systems. In *INCOSE International Symposium 2011*.

Miller, R. (2009). Peering disputes migrate to IPv6. *Data Center Knowledge*.
URL http://www.datacenterknowledge.com/archives/2009/10/22/peering-disputes-migrate-to-ipv6/

Morganwalp, J., Sage, A. P. (2003). A system of systems focused enterprise architecture framework and an associated architecture development process. *Information Knowledge Systems Management*, 3.

Murphy, P. R., Daley, J. M. (1998). Some propositions regarding rail-truck intermodal: An empirical analysis. *Journal Of Transportation Management*.

Nash, J. (1950). The bargaining problem. *Econometrica*, 18(2):pp. pp. 155–162.
URL http://www.jstor.org/stable/1907266

Nightingale, D. J. (2003). Lean enterprises – a systems perspective. In *ESD Internal Symposium ESD-WP-2003-01.14*.

Norman, D., Kuras, M. L. (2004). Engineering complex systems. Tech. Rep. 04-0043, MITRE.

Oliveira, P., von Hippel, E. (2011). Users as service innovators: The case of banking services. *Research Policy*, 40(6):pp. 806 – 818. doi:10.1016/j.respol.2011.03.009.
URL http://www.sciencedirect.com/science/article/pii/S0048733311000527

Pei, R. S. (2000). Systems of systems integration (SoSI) — a smart way of acquiring army C4I2WS systems. In *Proceedings of the Summer Computer Simulation Conference*.

Plant, J. F. (2002). Railroad policy and intermodalism: Policy choices after deregulation. *The Review of Policy Research*, 19(2):pp. 13–32.

Rademacher, P., Marks, M. (2012). A conversation with Paul Rademacher and Mano Marks, Google Maps API Office Hours.
URL http://www.youtube.com/watch?v=0YWjgtqQLSc

Rao, M., Ramakrishnan, S., Dagli, C. (2006). Modeling net-centric system of systems using the systems modeling language. In *CSER 2006*. CSER.

Rijsenbrij, J. (2008). Container handling in mainports: a dilemma about future scales. In R. Konings, H. Priemus, P. Nijkamp, eds., *The Future of Intermodal Freight Transport*, chap. 7. Edward Elgar Publishing Ltd.

Rizzoli, A. E., Fornara, N., Gambardella, L. M. (2002). A simulation tool for combined rail/road transport in intermodal terminals. *Mathematics and Computers in Simulation*, 59(1–3):pp. 57–71. doi:10.1016/S0378-4754(01)00393-7.

Ross, A. M., Rhodes, D. H. (2008). Using natural value-centric time scales for conceptualizing system timelines through epoch-era analysis. In *INCOSE International Symposium 2008*.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica*, 50(1):pp. 97–109.

Sage, A. P., Biemer, S. M. (2007). Processes for system family architecting, design, and integration. *IEEE Systems Journal*, 1(1).

Sage, A. P., Cuppan, C. D. (2001). On the systems engineering and management of systems of systems and federations of systems. *Information Knowledge Systems Management*, 2:pp. 325–345.

Samet, D., Tauman, Y., Zang, I. (1984). An application of the aumann-shapley prices for cost allocation in transportation problems. *Mathematics of Operations Research*, 9(1):pp. 25–42.

Sappington, D. E. M. (1991). Incentives in principal-agent relationships. *The Journal of Economic Perspectives*, 5(2):pp. 45–66.

Schneeweiss, C. (2003). Distributed decision making–a unified approach. *European Journal of Operational Research*, 150(2):pp. 237–252. doi:10.1016/S0377-2217(02)00501-5.

Schneeweiss, C., Zimmer, K. (2004). Hierarchical coordination mechanisms within the supply chain. *European Journal of Operational Research*, 153(3):pp. 687–703. doi:10.1016/S0377-2217(02)00801-9.

Sgouridis, S. P. (2007). *Symbiotic Strategies in Enterprise Ecology: Modeling Commercial Aviation as an Enterprise of Enterprises*. Ph.D. thesis, MIT Engineering Systems Division.

Shah, N. B., Hastings, D., Rhodes, D. (2007a). Systems of systems and emergent system context. In *5th Conference on Systems Engineering Research*.

Shah, N. B., Richards, M., Broniatowski, D., Laracy, J., Springmann, P., Hastings, D. (2007b). System of systems architecture: The case of space situational awareness. In *AIAA Space 2007*. AIAA.

Shapley, L. S. (1953). A value for n-person games. In H. Kuhn, A. Tucker, eds., *Contributions to the Theory of Games II*, vol. 28 of *Annals of Mathematical Studies*, pp. 307—317. Princeton University Press.

Simon, H. (1996). *The Sciences of the Artificial*. MIT Press.

Singel, R. (2008). ISP quarrel partitions internet. *Wired.com: Threat Level*.
URL http://www.wired.com/threatlevel/2008/03/isp-quarrel-par/

Sloane, E., Way, T., Gehlot, V., Beck, R., Solderitch, J., Dziembowski, E. (2007a). A hybrid approach to modeling soa systems of systems using cpn and mesa/extend. In *1st Annual IEEE Systems Conference*, pp. 1–7.

Sloane, E., Way, T., Gehlot, V., Levitin, A., Beck, R. (2007b). SoSE modeling and simulation approaches to evaluate security and performance limitations of a next generation national healthcare information network (nhin-2). In *SoSE '07*, pp. 1–6.

Song, D.-W. (2002). Regional container port competition and co-operation: the case of hong kong and south china. *Journal of Transport Geography*, 10(2):pp. 99–110. doi:doi:10.1016/S0966-6923(02)00003-0.

Stahlbock, R., Voss, S. (2008). Operations research at container terminals: a literature update. *OR Spectrum*, 30:pp. 1–52.

Sussman, J. (2000). *Introduction to Transportation Systems*. Arthouse.

Sussman, J. M. (2002). Collected views on complexity in systems. Working Paper ESD-WP-2003-01.06, MIT ESD.
URL http://esd.mit.edu/WPS/internal-symposium/esd-wp-2003-01.06.pdf

Team Register (2005). Level 3 depeers cogent. *The Register*.
URL http://www.theregister.co.uk/2005/10/06/level3_cogent/

Tisserand, J. B. (2005). *Network Centric Warfare Insights*, vol. 3 of *Network Centric Warfare Case Study: U.S. V Corps And 3rd Infantry Division (Mechanized) During Operation Iraqi Freedom Combat Operations (Mar-Apr 2003)*. U.S. Army War College.

Transportation Research Board (1998). *Policy Options For Intermodal Freight Transportation*. Transportation Research Board.

Tsamboulas, D., Kapros, S. (2000). Decision-making process in intermodal transportation. *Transportation Research Record: Journal of the Transportation Research Board*, 1707(-1):pp. 86–93.
URL http://dx.doi.org/10.3141/1707-11

Tuimala, J., Lukka, A. (1999). The importance of strategic alliances in intermodal transport, theoretical implications and an empirical pre-evaluation of the phenomena in the south-eastern finland. Research report, Lappeenranta University of Technology, Industrial Engineering, Logistics.
URL http://urn.fi/URN:ISBN:952-214-272-7

Unger, D. W. (2003). *Product Development Process Design: Improving Development Response to Market, Technical, and Regulatory Risks*. Ph.D. thesis, MIT Engineering Systems Division.

Valerdi, R., Axelband, E., Baehren, T., Boehm, B., Dorenbos, D., Jackson, S., Madni, A., Nadler, G., Robitaille, P., Settles, S. (2008). A research agenda for systems of systems architecting. *International Journal of System of Systems Engineering*, 1(1/2):pp. 171–188.

Van Der Horst, M. R., De Langen, P. W. (2008). Coordination in hinterland transport chains: A major challenge for the seaport community. *Maritime Econ Logistics*, 10(1-2):pp. 108–129.

von Neumann, J., Morgenstern, O. (1953). *Theory of Games and Economic Behavior*. Princeton University Press, 3rd edn.

Walden, D. D. (2007). The changing role of the systems engineer in a system of systems (sos) environment. In *1st Annual IEEE Systems Conference*, pp. 1–6.

Wardrop, J. G. (1952). Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers*, 1:pp. 325–378.

Wiegmans, B., Nijkamp, P., Rietveld, P. (2008). Container handling quality. In R. Konings, H. Priemus, P. Nijkamp, eds., *The Future of Intermodal Freight Transport*, chap. 6. Edward Elgar Publishing Ltd.

Woodcock, J., Cavalcanti, A., Fitzgerald, J., Larsen, P., Miyazawa, A., Perry, S. (2012). Features of cml: a formal modelling language for systems of systems. In *Proceedings of the 7th International Conference on System of System Engineering*.

Yan, Z., Haimes, Y. Y. (2007). Multiobjective resource allocation framework for risk management of large-scale hierarchical systems. In *SoSE '07*, pp. 1–6.

Yang, K., Chen, Y., Lu, Y., Zhao, Q. (2010). The study of guided emergent behavior in system of systems requirement analysis. In *System of Systems Engineering (SoSE), 2010 5th International Conference on*, pp. 1 –5. doi:10.1109/SYSOSE.2010.5544033.