A Theory of Objective Sizing

Ricardo Valerdi
Massachusetts Institute of Technology
Cambridge, MA 02139
rvalerdi@mit.edu

Abstract
The process of building parametric models to estimate the cost of large scale complex systems have recently uncovered unanticipated challenges. The most difficult of which includes the ability to define the boundary of the system being estimated. This boundary is an essential step towards determining the size of the system; a major input into parametric models. In this paper, we build on a concept from psychology known as the *moon illusion* to develop a theory of objective sizing. This theory has two main benefits: it helps explain why stakeholders have different views of systems and it provides an approach for how these differences can be reconciled. Ultimately it will help technical communities arrive at a more objective way for measuring system size which will ultimately improve the accuracy and influence of parametric models.

Introduction

In the last few hundred years human experiences have led to the development of science and supporting theories. These theories, while occasionally proven wrong, have helped explain phenomena surrounding the world in which we live. If adequately validated, these theories evolve to be socially accepted and applied in different contexts. In the twentieth century, a multitude of theories have emerged in engineering and management that have been socially accepted. Some of these include systems theory, cybernetics theory, complexity theory, chaos theory, game theory, approximation theory, and network theory. Even with these theories, there are unexplained phenomena that exist which have not been adequately understood.

Today we are surrounded by many new design system paradigms, increasingly complex enterprises, and large-scale infrastructures that require new methods, tools, and theories that explain their behavior. Specifically, experience developing COSYSMO[1] introduced new issues that previous cost models did not address. These issues were centered around blurred technical boundaries within which systems engineerswork. The existence of multiple layers of systems engineering introduced significant difficulties in using the model since users had diverse perspectives of the system boundary. This new dimension to the cost model required the introduction of clear counting rules for its size drivers which were: system requirements, system interfaces, and operational scenarios. It was observed that the diversity in interpretation of these drivers was caused by the perspective of users and the level of granularity they were interested in estimating. Coupling these with measures of system complexity resulted in a useful approach that helped explain how the size of a system could be arrived at objectively. Thus, the

[1] The Constructive Systems Engineering Cost model

proposed *Theory of Objective Sizing*, which will simply be referred to as *Theory S*. A necessary important step towards an explanation of this theory involves establishing a related concept which is a theory in itself: the stimulus-relation theory.

The Moon Illusion

The moon seen close to the horizon appears to be quite a bit larger to the human eye than when it is viewed at its zenith. The same effect can be observed when viewing the sun or constellations of stars. This odd phenomenon is called the *moon illusion* and there are records of it dating as far back as seventh century A.D. (Minnaert 1940). This illusion was later verified through astronomical experiments and has motivated other experiments in areas of psychology (Kaufman & Rock 1962) and visual perception (Baird & Wagner 1982), eventually leading to its formalization as the stimulus-relation theory. The basic idea behind the theory is that our perception of the size of the moon at the horizon and at the zenith varies despite the physical distance to the moon remaining essentially the same.

Through the ages, many writers have speculated that the illusion has a physical basis. It was suggested that the horizon moon really is closer to the observer, and therefore seen larger, until Newton's description of the nature of the moon's orbit showed the contrary to be true: the image of the moon at the horizon is in fact slightly smaller than the moon high up in the sky because the moon is 1.67% closer to the observer at its zenith, as illustrated in Figure 1.

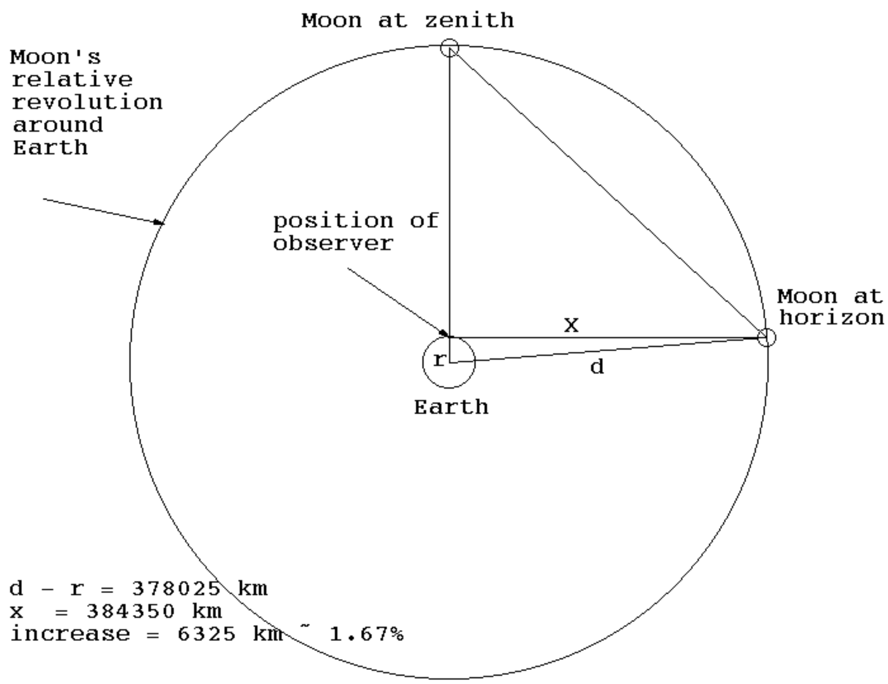

Figure 1. The Moon's Orbit Around the Earth (Borghuis 1999)

As shown in Figure 1, the distance from the earth to the moon is approximately 378,025 kilometers. The distance from an observer standing on earth and viewing the horizon moon is 384,350 which is a difference of less than 2%. Even with the insignificant difference in the distance to the moon, human subjectivity leads to different interpretation of the moon's size depending on its location. Experiments have been performed to measure the magnitude of this illusion. One experiment showed that observers estimated the moon to be 3.5 times as large near the horizon (Minnaert 1940).

This theory holds true in the physical world where familiar reference points such as the earth's horizon are used to compare the size of the horizon moon and the zenith moon. The subjectivity of this process is the main source of the magnitude of the variability of this illusion. A similar phenomenon exists in the functional world which provides insight into the development of a new theory for sizing systems.

The Functional World

In contrast to the physical world, the functional world requires a different conceptual model for size simply because distance, volume and weight do not adequately capture system complexities. Complex systems such as the Future Combat System, iPods, the internet need measures for size that consider their inherent attributes independent of physical size. Functional size is important is because it is an adequate predictor of development effort and cost. Most parametric cost models agree that there is a positive relationship between functional size and effort by following the general form:

$$Effort = f(Size)$$

The ISPA community and associated groups have worked towards identifying such measures of size that are reliable predictors of cost such as software lines of code (Boehm 1981) and function points (Albrecht & Gaffney 1983) for software systems. These predictors, together with associated counting rules, help measure system size from a functional standpoint and are fundamental components of parametric cost models.

These size drivers are generally adequate for systems that have clear boundaries such as an ERP[2] software system. For example, by using the COCOMO II model it is possible to estimate the amount of software development effort in person-months required for the delivery of an ERP system (Boehm 2000). One of the possible size drivers used to arrive at this estimate is SLOC. The effort and schedule estimate provided by COCOMO II is bounded by the scope of the ERP "system-of-interest" which is a notion introduced in the ISO 15288 *System Life Cycle Processes* (ISO 2002) shown in Figure 2. The idea behind the standard is that the system-of-interest is a flexible construct that can be applied to any part of the system. In Figure 2, the system of interest can extend as far as the entire hierarchical chart or as specific as a subsystem as indicated by the dashed rectangle. While the system-of-interest diagram is applicable to organizational, political, environmental, or other types of systems, the focus of this discussion is specific to engineering systems[3] and socio-technical systems[4] because of their applicability in the field of cost modeling.

---

[2] Enterprise Resource Planning is used as a generic example for a Human Resource application

[3] Engineering Systems are large and complex systems designed by humans having some purpose; will have a management or social dimension as well as a technical one (i.e., the telephone system)

[4] Socio-technical systems are is an approach to complex organizational work design that recognizes the interaction between people and technology in workplaces (Cherns 1976)

As these complex engineering systems emerge there is a need for cost models to cover systems-of-interest that reach beyond the boundaries of traditional systems. In the case of the COSYSMO model, users were frequently faced with the challenge of estimating multiple layers of systems engineering effort for a single system; often in the case of a request for proposal. This is partly due to the different layers in which systems engineering can be applied at different layers of complex systems. As in the ERP example, it is common for systems engineering activities to take place at the Human Resources level that focus solely on the ERP system, as well as the corporate IT level for business planning, and the corporate level for strategic planning. This presents new cost modeling challenges that cannot be addressed by existing cost models.

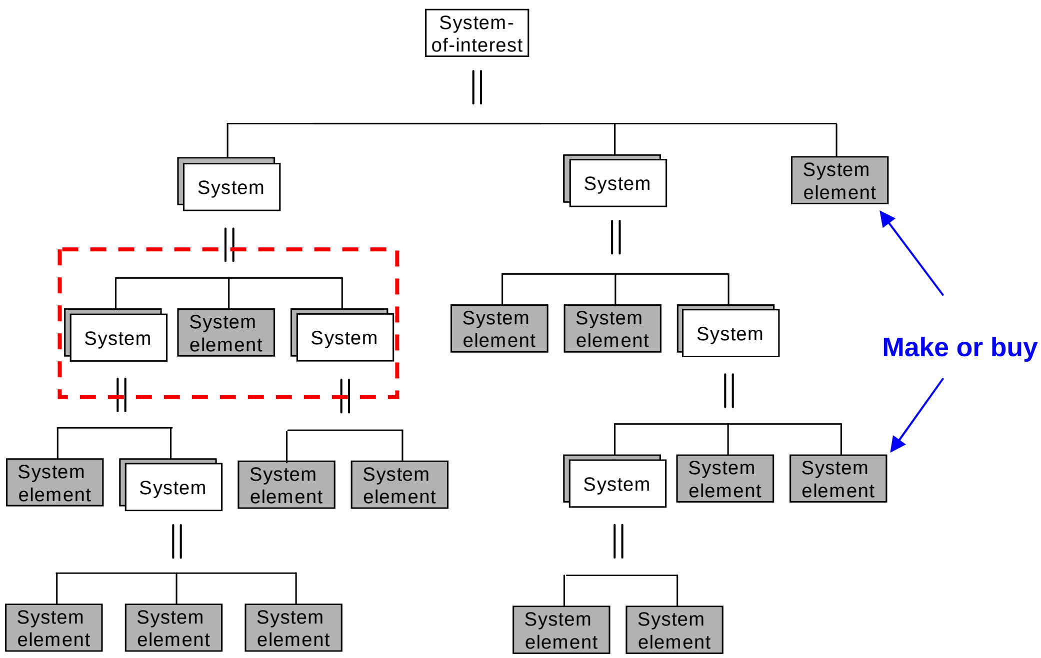

Figure 2. ISO/IEC 15288 System of Interest Structure (ISO 2002)

The proposed theory aims to improve the use of size measures and thereby improve their predictive power which can lead to a consequential significant improvement in the accuracy of parametric cost models.

Theory S

A theory is an attempt at formulating an explanation for a certain class of experiences. It is also an argument that is based on empirical observation which requires some type of evidence to be considered valid. The arguments proposed by the theory are understandable only to those that can stir up similar experiences as the empirical basis for testing the truth of theory. Unless the theory triggers corresponding experiences at least to a degree, it will create the impression of empty talk or will perhaps be rejected as an irrelevant expression of subjective opinions.

While the topic of sizing has consistently appeared in previous cost estimation conferences it has mostly focused on systems that have a well defined system boundary. *Theory S* grew out of a need to size large-scale complex systems with dynamic system boundaries. This class of systems is inevitable due to the increased sophistication of technology and needed capabilities in the military and commercial arenas.

The "objectiveness" of the theory reiterates the main goal: to develop a framework for measuring the size of a system using methods that rely less on human opinion and as a result have higher reliability and accuracy. Theory S consists of three underlying elements: perspective, granularity, and complexity. Each element provides a different trait of the theory and its applicability to the functional world.

<u>Perspective</u>

The point of reference of a success critical stakeholder is a critical determinant of how the functional size of a system is measured. Independent of the size metric used, stakeholders will take on different perspectives of the system based on who they are and their relative role in the system. The best example of this concept is in the different perspectives that can be adopted when describing the boundaries of an aerospace enterprise. While it deviates slightly from the topic of pure technical systems it shows the applicability of *Theory S* in socio-technical systems. In Figure 3, three system-of-interest perspectives from the lean manufacturing industry are shown (Piepenbrock 2005) which demonstrate increasingly larger views of enterprises. In Figure 3a, the perspective of an enterprise is illustrated as a company with external linkages to customers, government, labor unions, and suppliers. Aerospace companies could appropriately see themselves from this perspective and as a result maintain that the size of their enterprise is based on the company alone. In Figure 3b, an expanded view of the same organization is illustrated from a broader perspective to include associated customers and suppliers as key elements of the enterprise. In this variation, the size of the enterprise is leads to a larger functional size of the system-of-interest.

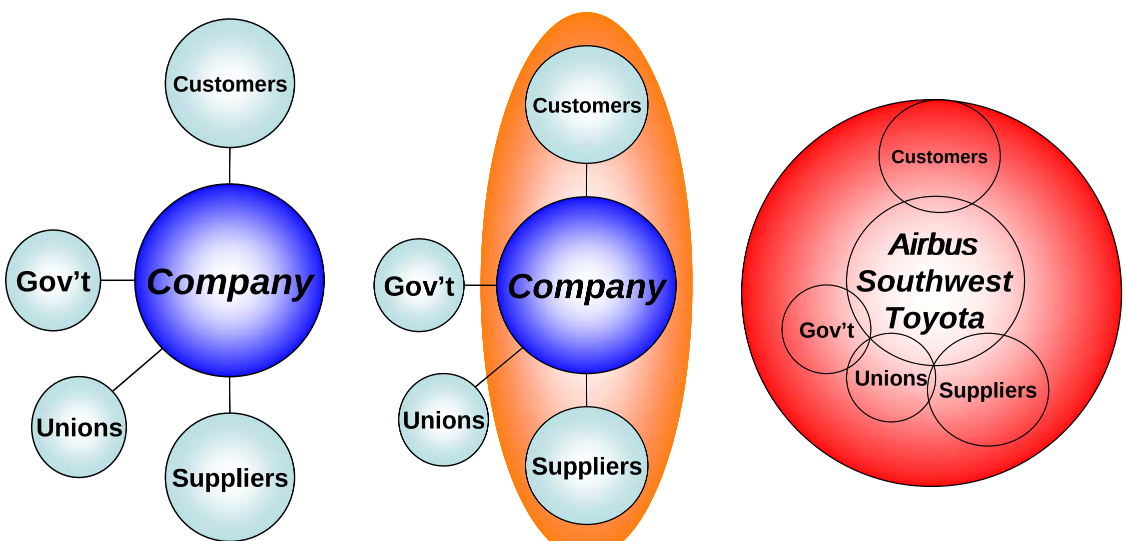

Figure 3a. The Company as a System of Interest

Figure 3b. Enterprise with Supply Chain Focus

Figure 3c. Extended Enterprise Focus

In Figure 3c, the perspective of the extended enterprise is broadened to include complementary partner companies tightly coupled with customers, government, labor unions, and suppliers. This yields an increased functional size that is driven by the increased complexity and capability of the enterprise. Communicating which perspective is being used is the first and most important element of *Theory S* because it helps define the boundary of the system; whether social or technical.

Granularity

The idea that one person's system may be another's subsystem is the catalyst for the variability in granularity within a system. The concept of granularity originates from the idea that the texture of grain changes as the visual distance changes. As a key component of *Theory S*, the perspective must be determined before the preferred granularity of the system is identified.

In the world of software development, use cases have been a popular and effective way to describe the functions of software before actual coding begins. When software is treated as a system it too can have multiple layers of complexity that can be described with different levels of granularity. A creative example of this concept appears in the book *Writing Effective Use Cases* by Alistair Cockburn. The author provides a framework for categorizing five different levels of software use cases: cloud level, kite level, sea level, fish level, and clam level (Cockburn 2001). As shown in Figure 4, it is a useful metaphor for demonstrating software functionality in increasing detail as described by use cases.

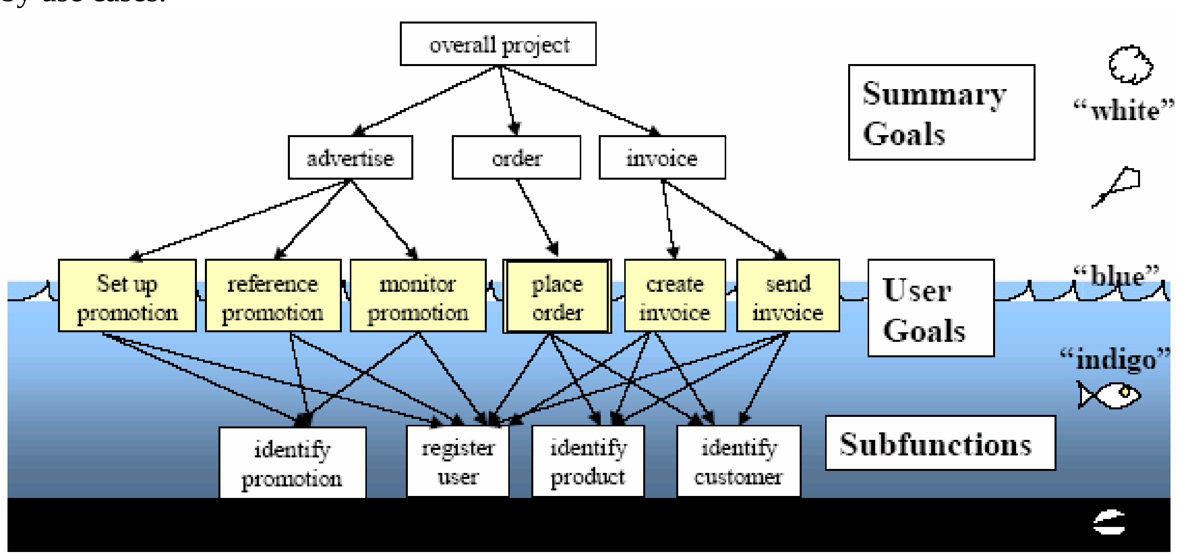

Figure 4. Three named goal levels framework for Use Cases (Cockburn 2001)

This key element of *Theory S* is tightly coupled with the perspective element. The two elements are implicit assumptions in system sizing which are often taken for granted and often not communicated. In the domain of complex systems, these must be explicit because so many things can vary between them as the system approaches a steady state. More interestingly, they evolve together as needs evolve. Once they perspective and granularity are solidified, the question of system complexity and how it is measured should be addressed next.

Complexity

Together with the functional size of a system there is an inherent characteristic of complexity. These are both quantifiable but orthogonal characteristics. For instance, it can be shown that a space satellite represents a larger complexity than a toaster. But in order to differentiate the two, a reliable measure of complexity is needed. The third element of *Theory S* postulates that functional size is a function of system complexity:

$$Size = f(Complexity_{SYSTEM})$$

It is unrealistic to expect a single metric of complexity for all types of systems. Naturally, the type of metric used is motivated by the perspective and granularity of the system-of-interest under consideration. It also depends on the type of complexity being measured. The viewpoint of structural complexity, for example, can be used to measure the number of units in a system and the connections between them. This construct is applicable to higher level views of the system and can be calculated with:

$$S = \frac{\sum_{1}^{n} f_i^2}{n}$$

where: f_i = number of calls (or "fanout" from unit "*i*")

v_i = number of variables (or data items) processed by unit "*i*")

n = number of units in design

Alternatively, a cyclomatic complexity viewpoint could be used at a very low level of a software system to determine the count of the nonrepeating paths through the decision structure of the pseudocode or code which can be calculated with:

$$v(G) = e - n + 2$$

where: v = cyclomatic number of graph

e = number of edges

n = number of nodes

The type of complexity being measured deliberately complements the other two elements of *Theory S* and helps arrive at a more objective measure of system size.

Conclusion

Just as the moon can seem larger at the horizon than at the zenith, sizing approaches for complex systems can vary dramatically. When orchestrated, the three elements of *Theory S* can drastically improve the process of sizing which can subsequently lead to more accurate cost estimates of complex systems. This is due to the fact that these elements can help objectively define the unit of analysis (perspective), the level of detail (granularity), and complexity metric(s) needed to quantify complex systems. Realistically, however, there is a certain degree of subjectivity involved in this exercise by nature of involving humans in the process of selecting perspectives and levels of granularity. It is possible to partially alleviate this through the use of architecture products such as the twenty-six views provided in DODAF[5]. Unfortunately, the use of this approach may be limited to the DoD community due to the military bias of its taxonomy.

The next step for *Theory S* is to apply it to situations where all three elements can be realized. The theory can only be validated and trusted to the extent that it is tested with experiments. In the end, the hope is to have a theory that is useful to users and developers of parametric models so that they can influence decision making processes in the development of complex systems.

Acknowledgements

The author appreciates enthusiastic feedback received from JK Srinivasan, Chris Miller, and Briana Valerdi.

References

Albrecht, A. J., Gaffney, J., Software function, source lines of code, and development effort prediction: A software Science validation, *IEEE Transactions on Software Engineering*, Vol. SE-9, pp. 639-648, 1983.

Boehm, B. W. *Software Engineering Economics*, Prentice-Hall, 1981.

Boehm, B. W., Abts, C., Brown, A. W., Chulani, S., Clark, B., Horowitz, E., Madachy, R., Reifer, D. J. and Steece, B., *Software Cost Estimation With COCOMO II*, Prentice Hall, 2000.

Borghuis, B., *The Moon Illusion*, Literature Thesis, University of Pennsylvania Vision Research Center, January 1999.

Baird, J.C., Wagner, M., The moon illusion: I. How high is the sky?, *Journal of Experimental Psychology*, 81, 376-380, 1982.

Cherns, A.E., Principles of Sociotechnical Design, *Human Relations*, 29,(8) 783-792, 1976.

Cockburn, A., *Writing Effective Use Cases*, Addison-Wesley, 2001.

Gilinsky, A.S., The Effect of Attitude Upon the Perception of Size, *American Journal of Psychology*, 68, 173-192, 1955.

Gutzwiller, M.C., The Oldest Three-Body Problem, *Reviews of Modern Physics*, 70 (2), 589-639, 1998.

[5] Department of Defense Architecture Framework

Hershenson, M., *The Moon Illusion*, Hillsdale, NJ, Erlbaum, 1989.

ISO/IEC, *ISO/IEC 15288:2002(E) Systems Engineering - System Life Cycle Processes*, 2002.

Minnaert, M., *Light & Colour in the Open Air*, G. Bell & Sons, London, 1940.

Rock, I., Ebenholtz, S., The Relational Determination of Perceived Size, *Psychol. Review*, Vol. 66, pp. 387-401, 1959.

Kaufman, L., Rock, I., The moon illusion: I, *Science*, Vol. 136, No. 3520, pp. 953-61, June 1962.

Piepenbrock, T., "Enterprise Ecosystems: Architectural Forms, Structural Dynamics, Co-Evolutionary Paths, and Competitive Outcomes," *MIT Engineering Systems Division Seminar*, 2005.

Biography

Ricardo Valerdi is a Research Associate at the Lean Aerospace Initiative at MIT, a Member of the Technical Staff at the Aerospace Corporation in the Economic & Market Analysis Center, and a Visiting Associate at the Center for Software Engineering at USC. He earned his BS in Electrical Engineering from the University of San Diego, MS and PhD Systems Architecture & Engineering from USC. Formerly he was a Systems Engineer at Motorola and at General Instrument Corporation.