A Rule-Based Decision Support Tool for Architecting Earth Observing Missions

Daniel Selva
Massachusetts Institute of Technology
77 Massachusetts Ave, Room 33-409
Cambridge, MA 02139
617-682-6521
dselva@mit.edu

Edward F. Crawley
Massachusetts Institute of Technology
77 Massachusetts Ave, Room 33-413
Cambridge, MA 02139
617-253-7510
crawley@mit.edu

Abstract—**A decision support tool is presented that is especially tailored for architecting Earth observing missions and programs. The tool features both a cost model and a performance model. This paper focuses on the description of the performance model. Indeed, while considerable effort has been put into the development of cost estimating models, comparably much less effort has been put into the development of quantitative methods to assess how well Earth Observing Mission satisfy scientific and societal needs. A literature review revealed that existing methods include a commercial approach, a value-of-information approach, end-to-end simulation, assimilation in Observing System Simulation Experiments, and simple expert judgment. Limitations of these methods include limited applicability, computational complexity, low modeling fidelity (e.g. abstraction of synergies between measurements), and subjectivity. Our method uses a knowledge-based system to store and manage large quantities of expert knowledge in the form of rules-of-thumb that replace expensive computations. Scientific and societal measurement requirements and instrument capabilities are expressed in the form of logical rules and data structures. An efficient pattern matching algorithm performs the comparison of the measurement requirements and the measurement capabilities on the basis of 64 different measurement attributes. The system is demonstrated on the Earth Science Decadal Survey. While the system is still under development, it shows great potential to enhance traceability in the modeling of scientific and societal value of Earth observing missions. Furthermore, the recursive nature of rule-based systems shows potential to model synergies between instruments and measurements, at a sufficient level of fidelity for architectural trade studies, especially for the ones conducted in committees with experts such as Decadal Surveys.**

TABLE OF CONTENTS

1. INTRODUCTION

This paper is concerned with the assessment of scientific and societal merit of Earth Observing Missions or more generally Earth Observing Satellite Systems (EOSS), in the context of early architectural trade studies. The methodology presented in this paper is particularly relevant for semi-automatic assessment using man-in-the-loop decision support tools.

This first section of the paper starts by providing the necessary background on system architecture, decision support tools, EOSS, and the particular problem of scientific benefit assessment of EOSS. A review of prior methods for scientific merit assessment is provided. A gap in the literature, and subsequent research goals, are identified. Knowledge-based systems (KBS) and in particular rule-based expert systems (RBES) are then introduced as an alternative approach. The necessary background on RBES is provided, including a short history and a critique of RBES. The section concludes with a review of the structure of the rest of the paper.

System Architecting and Decision Support Tools

System Architecting— In the late 80's, researchers started to realize that some concepts from traditional architecture and civil engineering were being used by engineers in charge of designing and building unprecedented, large, complex systems [1]. These concepts included the creation of a separate position for a lead systems engineer at the interface between the client and the design team, a more direct engagement of the client in the high-level design of the system, and a holistic, value-centered, lifecycle view of the system. Rechtin was arguably the first to formalize this concept, and he coined the term "systems architecting" [2]. His book with Maier is, perhaps, still the best introduction to the field [1].

Crawley defines system architecture as "the embodiment of concept, and the allocation of physical/informational function (process) to elements of form (objects) and definition of structural interfaces among the objects" [3]. Essentially, the architecture of a system is its highest level design. However, it takes a holistic view that goes beyond traditional design. More precisely: 1) it takes into account technical and non-technical factors; 2) it is centered in delivering value to stakeholders as opposed to optimizing performance or cost; 3) it takes into account all the phases of the lifecycle including manufacturing, testing, operations and disposal.

978-1-4577-0557-1/12/$26.00 ©2012 IEEE

System architecting is the process of creating a system architecture. Fundamentally, the system architecting process is a decision making process, where the decisions to be made concern the highest level design of the system [4]. Architectural decisions are done the earliest in the design process, and mostly because of this, they are different from other design decisions, in several points: 1) they have to be done in a highly ambiguous context because very few decisions have been made; 2) they commit the largest part of the lifecycle cost of the system; 3) they have the largest impact on subsequent design decisions; 4) they have the largest impact on performance, risk, flexibility and other figures of merit. Once the main architectural decisions are made, the concept or essence of the system is fixed.

The typical system architecting process is similar to that of a trade study, and consists of three steps: 1) the system architect(s) select a handful of candidate architectures; 2) each architecture is assessed, typically in terms of cost and performance; 3) one or two architectures are selected for further studies. This process, or more precisely the decision making parts of this process, is far from being ideal for several reasons, mostly related to the bias that the human system architect(s) and their organization bring from their previous experience and expertise.

Decision Support Tools— Today, it is widely accepted that decision support tools (DST) can help improve the system design and architecting processes by providing a rigorous framework for objective and consistent evaluation of a larger number of architectures under a variety of scenarios using several figures of merit derived directly from stakeholder needs, as well as some guidance for the search and selection processes [5]. These tools include a mix of simulation tools, combinatorial optimization algorithms, artificial intelligence search algorithms, utility theory, and decision making under uncertainty.

DSTs for system architecting perform three tasks: synthesis or **enumeration** of feasible architectures, **evaluation** of architectures, and **selection** of preferred architectures. Modeling is required throughout the process to: a) encapsulate architectural decisions in an enumerable data structure; b) create objective functions that represent the needs of the stakeholders; c) create the appropriate filters for down-selection of the preferred architectures.

There are a variety of examples of DSTs successfully applied to architecting of large-scale aerospace systems. Simmons and Koo used Object Process Networks to study the architecture of the lunar exploration program [6]. Ross, Hastings et al used the Multi-Attribute Tradespace Exploration methodology to study several projects, for instance, the Terrestrial Observer Swarm (X-TOS) [7]. Bayley et al used a genetic algorithm to optimize the architecture of a launch vehicle [8]. De Weck et al also used real options to analyze architectures for constellations of communications satellites [9]. In the context of the DARPA F6 project for fractionated spacecraft, several DST were developed by partnerships between industry and academia [10], [11], [12], [13], [14].

Architecting Earth Observing Missions— Basic system architecture theory teaches us that, when trying to identify the architecture of a system, it is necessary to ask ourselves how the system provides value to the stakeholders. We propose that ultimately, EOSS provide value to their stakeholders because they take valuable **measurements** that are processed into **data products,** which in turn are then transformed into information and knowledge, which actually bear the value. Therefore, most stakeholder needs and goals can be projected into a set of measurement requirements. Using words from the system architecting jargon, measurements are the fundamental elements of function of EOSS. This does not mean that they are the only elements of function, since for example data processing is required to transform raw measurements into different data products, from Level 0 to Level 4.

These measurements are taken by instruments flying on spacecraft, which in our model are the fundamental elements of form of EOSS. An instrument-centric view of an EOSS, as opposed to a mission-centric view, provides the right level of abstraction for architectural studies: going one level deeper in detail would imply delving into instrument subsystem design, which is not needed at this early stages of mission development; on the other hand, staying at the mission level would preclude the study of some architectural decisions such as that of deciding whether two instruments should share a common bus or rather fly on dedicated buses [15]. A pictorial summary of our architectural model for EOSS is provided in Figure 1.

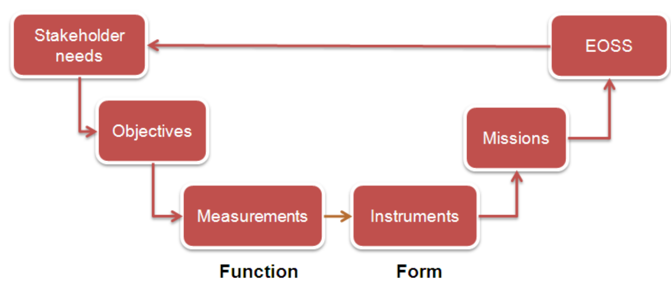


Figure 1: Simplified architectural model for an EOSS, in which data processing functions have been abstracted out.

Scientific and Societal Value of EOSS

Motivation— When conducting architectural trade studies of any system, it is customary to compare architectures in terms of their cost-effectiveness. This implies the use of at least two figures of merit: one related to performance, and one related to cost. Cost estimating methodologies for space systems are far from being ideal, but they have been thoroughly studied (see for example [16], [17]).

In some cases, providing a quantitative assessment of system performance may actually be at least as challenging as estimating the cost of a mission. Indeed, scientific or societal value is hard to quantify, and often necessarily subjective due to large amounts of uncertainty in any scientific endeavor. It is hard to predict the value of an EOSS because it is hard to predict whether a certain scientific experiment will be more successful than another one. Moreover, additional scientific value may come from unforeseen uses of a given technology for measurements other than their primary application. Still, having the capability to quantitatively assess the scientific value of an EOSS is not only useful for architectural trade studies, but arguably a required step for any large project using taxpayer money, as it usually is the case for EOSS.

Literature Review—Scientific performance can be described as the ability of the EOSS to produce data sets that will advance the state of the art of a certain scientific discipline. Science Traceability Matrices [18] provide a way to trace mission and instrument requirements back to scientific requirements, but they are not by themselves a tool to quantify scientific merit. Several different approaches have been proposed that treat all or part of the problem: a) commercial approach; b) value-of-information approach; c) end-to-end simulation approach; d) assimilation approach; e) expert assessment approach.

The commercial approach computes the benefit of an EOSS as the price in the market of the total volume of data produced by the EOSS. This is mostly relevant to imaging systems, because they generate data products that are of interest to the commercial sector (e.g. Google maps). However, the commercial approach fails to capture broader benefits of certain types of measurements, for which there is currently no market (e.g. Earth radiation budget).

In the value-of-information approach, the benefit of science is typically also monetary, although it can sometimes be expressed in "utils", the adimensional unit in which utility functions are assumed to be implicitly expressed. Typically, the problem is framed as a situation in which there are several possible outcomes and a decision needs to be made under uncertainty (e.g. climate policy under the uncertainty of magnitude and effects of global change). Reducing the uncertainty in the range of possible outcomes can increase the expected outcome of the situation (e.g. expected consumption from an economic perspective). The value of a certain remote sensing data product can thus be computed as the difference in expected outcome (e.g. consumption in $ or "utils") between the scenario where the data product is available (i.e. reduced uncertainty), and that where it is not available (i.e. full uncertainty). In other words, the value in this case comes from the use of the data produced by the EOSS to make early, informed decisions that have an impact on society. This approach is well illustrated in Nordhaus and Popp's paper on the value of scientific knowledge in the context of global climate change [19]. This approach is the preferred one by economists to capture the societal benefit of EOSS whenever this decision making framework can be applied. However, this approach fails again to capture part of the purely scientific value.

In the end-to-end simulation approach, models are built for the observable, instrument, satellite, ground station, and data processing components, so that the different levels of the data product are simulated, from the unprocessed data at full resolution (Level 0) to the completely processed gridded images of the variable of interest in physical units (Level 3), or even of derived variables (Level 4). This allows numerically computing the changes in the quantity and quality (i.e. the utility) of the Level 2-4 data products resulting from a certain design or architectural change in the EOSS (e.g. drop a certain spectral band, or go from 2 to 4 satellites). End-to-end simulators are arguably the preferred method for engineers to assess the scientific value of an EOSS. End-to-end simulators are currently developed for most Earth observing missions at ESA [20], [21], [22], as a means of assessing mission effectiveness.

In some cases, the data produced by the EOSS is to be assimilated into some kind of forward running model, for example a numerical weather prediction (NWP) model. In this case, the value of an EOSS can be calculated as the marginal improvement in the performance of the forward model (e.g. improvement in anomaly correlation in weather forecasting) due to the assimilation of that EOSS dataset. Such experiments are called Observing System Simulation Experiments (OSSEs). They were first proposed by Atlas [23], and they have become very popular in the last decade (see for example ESA references [24], [25] and NASA references [26], [27]). OSSEs are the preferred method by scientists to quantify the marginal scientific merit of an EOSS for NWP. However, OSSEs suffer of two main limitations: a) scope: not all disciplines of the Earth sciences use forecasting models, and for some of them that do use them, it can be very hard to create a "nature run", i.e. a "truth" data set with which to compare the simulated dataset; b) computational complexity: OSSEs are extremely computationally expensive, and therefore it is virtually impossible to use them to compare hundreds, dozens, or even a handful of different EOSS architectures.

The last approach is simply the use of expert assessment. In this approach, experts are directly questioned on the scientific or societal merit of an instrument, mission, or EOSS, and their answers are either directly used, or lightly processed to create, scores that represent the scientific merit of an EOSS. Such called Science Value Matrices (SVM) are sometimes used for organizing and combining expert assessment of EOSS. SVMs are similar to science traceability matrices, except they do contain explicit scores for the scientific merit of an EOSS [18], [28]. This approach is the preferred one for higher level managers and decision makers, and typically used in expert committees such as Decadal surveys. An example of this approach at NASA can be found in [29]. This method is preferred by decision makers and program managers. The use of SVMs is very computationally efficient, because in the worst case a few matrix operations are made to evaluate an EOSS. However, the reasoning process behind an expert assessment, which is largely based on the experience of the scientist with objective simulation tools such as OSSEs or end-to-end simulators, but also contains a subjective component, is abstracted out in this methodology. This lack of traceability is a limitation of most current SVMs. Furthermore, the process of eliciting the knowledge from the experts is hard for several reasons: experts may be biased in their assessments by their own expertise, and different people may have different evaluation scales resulting in inconsistent scores. The latter problem can be at least partially mitigated if a formal method from the social sciences is applied to obtain the scores. Examples of such methods are the Delphi method [30], or Saaty's Analytic Hierarchy Process [31].

Research Goals

The literature review showed that several different approaches have been proposed to assess the scientific benefit of an EOSS. However, none of them satisfy the following characteristics: a) they are universally applicable to all EOSS, all the disciplines of the Earth sciences, as well as societal needs; b) they are fast enough computationally to be used for early architectural trades; c) they are rigorous and objective.

Universality is important for two reasons: first, it provides a common evaluation scale for multi-discipline Earth observing programs and campaigns such as the Decadal survey; second, using the same tool for multiple EOSS reduces the development cost of DSTs, which is a time and resource consuming activity. Low computational complexity is important because it enables a more exhaustive exploration of the architecture tradespace, which reduces the probability of leaving an interesting architecture unexplored. Rigor and objectivity are always required in mission design, especially in public endeavors.

In this paper we propose an alternative methodology that: a) is universal because it is based on a hierarchical decomposition of stakeholder needs; b) is computationally fast enough to be used for architectural trade studies; c) compensates the subjectivity with an explanation facility to enhance the transparency of the tool. This approach is knowledge-intensive, i.e. it uses large quantities of expert knowledge, which are organized in an RBES. Since part of this knowledge will necessarily be inexact or subjective heuristics, the RBES has a built-in explanation facility to enable full transparency between the tool and the user. In other words, the user can trace back a particular result to the set of inputs that are responsible for it. This in turn facilitates sensitivity analyses in which the values of uncertain inputs (e.g. the presence of a certain precursor mission in the future, or the evolution of the maturity of a certain technology) can be changed and the model rerun to study the sensitivity the effect on the results in a matter of minutes.

This method is not meant to replace any of the alternative methods identified in the literature review. However, its characteristics make it particularly suitable for expert committees such as Decadal Surveys, and therefore the method could be used to enhance and complement SVMs.

Rule-based Expert Systems

Definitions— An expert system is "a computer program designed to model the problem-solving ability of a human expert" [32]. In order to do that, an expert system uses large bodies of heuristic – expert – knowledge. In a rule-based expert system (RBES), expert knowledge is encapsulated in the form of logical rules. In the words of Feigenbaum, considered the major creator of the first expert system, these rules map the knowledge "over from its general form (first principles) to efficient special forms (cookbook recipes)" [33]. This is in opposition to other kinds of expert systems that primarily use different data structures to store expert knowledge, such as frames in frame-based expert systems (FBES), which are very similar to objects in object-oriented programming [34].

In RBES, a logical rule is composed of a set of conditions in its left-hand side (LHS), and a set of actions in its right-hand side (RHS). The actions in the RHS are to be executed if the conditions in the LHS are all satisfied. An example of a logical rule is the following: LHS:="if the car won't start", RHS := "then check the electrical engine". An RBES infers information from rules in one of two ways: forward-chaining, when logical rules are used from the data to the goal in a deductive process; backward-chaining, when the rules are applied working backwards from a target goal [35].

It has been noted that many expert systems, and ours is not an exception, are actual mixtures of RBES and FBES, as they use both rules and frames to represent the knowledge. We consider our expert system for assessing scientific value of EOSS an RBES, since the primary means – not the only means - of knowledge representation is that of logical rules.

For an excellent introduction to rule-based expert systems, see [36]. A short history of RBES, a description of their structure, an introduction to the CLIPS language for developing RBES, and a critique of RBES are all provided in the Appendix.

Paper Structure

The rest of the paper is organized as follows. Section II describes in details the approach to the problem including the RBES developed for the assessment of scientific merit of EOSS. In section III, the RBES is applied to the Decadal Survey, and the results obtained are presented and discussed. Finally, in the conclusion (section IV), a summary of the main contributions and limitations of this work and some guidelines for potential improvements of the tool are discussed.

2. APPROACH

Overview—In the literature review we described several approaches to the problem that come from looking at the problem from different perspectives: science, engineering, economics, and management.

Our methodology proposes a system architecture approach to the problem, which combines aspects of all of the above, but is arguably closest to the management standpoint. From the system architecture perspective, the problem of assessing the scientific or societal value of an EOSS is an instance of the more general problem of assessing value of a system architecture, and therefore, according to the state-of-the-art of system architecture, it shall start with the identification of the stakeholders and their needs. In other words, scientific value is defined as satisfaction of a hierarchy of scientific needs.

In the case of EOSS, scientists of different disciplines (e.g. hydrology) and the people are two key stakeholder groups. An example of a need of hydrologists is to "improve our understanding of the water cycle", which decomposes into several more specific needs, one of which could relate to reduce the uncertainty in the total amount of surface water on the Earth". A decomposition of these needs, and their subsequent projection on the domain of the system at hand, provides two levels of requirements, that we called system objectives and subobjetives. For instance, one can conclude that in order to meet the needs of hydrologists in terms of surface water, it is required to measure surface soil moisture at a global scale, with an accuracy of 5%, a spatial resolution of 10km, and a temporal resolution of 1 to 2 days. Requirements expressed in this form always make reference to a particular measurement or to a set of measurements or data products, such as gridded images of a certain parameter (e.g. soil moisture).

Measurement attributes such as accuracy, spatial resolution, and temporal resolution determine their value to the stakeholders. For instance, if an EOSS provides soil moisture with an accuracy of 50% instead of 5%, the value that the EOSS provides to hydrologists will be severely reduced. Therefore, our strategy is based on an attribute-level comparison of two measurement sets: the required measurement set, which depends on the stakeholders and their needs; and the achieved measurement set, which depends on the EOSS architecture, and will in general differ from the required measurement set.

This comparison is essentially a pattern-matching process, and therefore could in principle be efficiently solved using RBES. Thus, we define rules that assess total or partial satisfaction of individual requirements, by comparing the attributes of the required and achieved measurement sets. These rules are called **requirement satisfaction rules**.

If stakeholders provide information about the relative importance of their requirements, individual requirement satisfaction can be aggregated to provide a single scientific/societal benefit metric. This information is contained in the **value aggregation rules**.

Requirement rules and value aggregation rules, together with stakeholder needs, objectives, and requirements, are all in the functional domain. The boundary between the functional domain and the formal domain is crossed when the instruments are introduced. The achieved measurement set is generated by applying the **instrument capability rules** to the set of instrument and satellite facts. For example, instrument capability rules can suggest that L-band passive radiometers can measure soil moisture and ocean salinity. The attributes of these measurements will be determined by the attributes of both the instrument and the mission through the **attribute inheritance rules**. For instance, the horizontal spatial resolution of a soil moisture measurement will be determined by the angular resolution of the instrument, its off-nadir angle, and the orbit characteristics. Finally, **synergy rules** take into account the emergent capabilities of measurements. For instance, a synergy rule may codify the fact that disaggregation schemes can produce a data product of high spatial resolution and high temporal resolution from the combination of two measurements of the same parameter: a sparse measurement with high spatial resolution, and a frequent measurement with low spatial resolution. All of these rules are described in further detail in the next paragraphs.

A Language for Architecting EOSS—In RBES, rules match patterns on facts in working memory. Therefore, rules and facts need to be specified in a common "language", namely a set of templates or data structures with slots or attributes.

In order to be able to express requirements for all the disciplines of the Earth Sciences, it is necessary to define a set of templates that has enough modeling breadth and depth. The goal is to maximize the expressivity of the language by adding as many attributes as needed.

The major fact types and corresponding templates used in the RBES are: measurements, instruments, missions (considered instruments and manifested instruments), orbits (for a mission analysis database), subobjective satisfaction and objective satisfaction. A few slots of the templates for missions, instruments, and measurements, are provided in Table 1. Overall, 64 different slots or attributes are defined in the measurement template, 102 attributes are defined in the instrument template, and 45 attributes are defined in the mission template. These numbers and Table 1 give an idea of the richness and complexity of the rules that can be expressed with such abstractions. Some of these attributes are only relevant to a particular scientific discipline (e.g. sensitivity-in-low-troposphere-and-PBL for atmospheric chemistry), or to a particular technology (e.g. number-of-looks# for radar). Sometimes, as a result of an expert interview, a new slot was added to the instrument or measurement template. This process can be done in a matter of seconds without affecting the rest of the code.

Table 1: Templates for the 3 main types of facts used in the RBES: missions, instruments, and measurements. Attributes that finish with the # symbol are numerical; the others are descriptive or semi-quantitative.

Mission	Instrument	Measurement
launch-vehicle	Concept	avg-revisit-time-cold-regions#
Lifetime#	Illumination source	avg-revisit-time-global#
mechanisms-penalty	num-of-SWIR-channels#	avg-revisit-time-US#
mission-architecture	num-of-TIR-channels#	Horizontal-Spatial-Resolution-Along-track#
mission-cost#	num-of-UV-channels#	Horizontal-Spatial-Resolution-Cross-track
Name	num-of-VNIR-channels#	signal-to-noise-ratio#
num-of-planes#	Penetration	Spectral-resolution
num-of-sats-per-plane#	Pointing-capability	Vertical-Spatial-Resolution
operations-cost#	Radiometric-accuracy#	Radiometric-accuracy#
orbit-altitude#	scanning	sensitivity-in-cirrus
orbit-anomaly#	sensitivity-in-cirrus	sensitivity-in-low-troposphere-PBL
orbit-inclination	sensitivity-in-low-troposphere-PBL	NEP-NEDT

Attribute Inheritance Rules— Mission, instrument, and measurement attributes, are populated at different moments in the evaluation processes, and have different origins: the user, other evaluation models (rule-based or not), an optimization algorithm, a database (e.g. a mission coverage attributes are retrieved from a mission analysis database), a parent system element (e.g. an instrument attribute may be inherited from its mission), or a combination of attributes from parent system elements (e.g. a measurement attribute may be computed from a mission attribute and an instrument attribute).

Attribute inheritance rules describe all these relationships between system attributes. A closer look at Table 1 will show that instruments and measurements share a few attributes such as radiometric accuracy. An attribute inheritance rule exists for each of these cases that dictates that the measurement attributes be inherited from the parent instrument. A more sophisticated type of attribute inheritance rule combines information from two different hierarchical levels in the architecture to populate a low-level attribute. An example of this case is illustrated in Figure 2, which corresponds to the calculation of horizontal spatial resolution from orbit characteristics (mission attribute) and angular resolution (instrument attribute). Note that some control over the flow of rule execution is required, since mission->instrument inheritance needs to occur before instrument->measurement inheritance, and inheritance from database needs to occur before both of them. Therefore, attribute inheritance rules are divided in a three groups with different rule priority (a.k.a. "salience" in CLIPS/Jess [54]).

```
(defrule compute-hsr-cross-track
  "Compute horizontal spatial resolution from instrument angular resolution and orbit altitude"
  ?instr🡐( Manifested-instrument (orbit-altitude# ?h) (Angular-resolution-azimuth# ?ara) (Horizontal-Spatial-Resolution-Cross-track# nil))
🡒
  (modify ?instr (Horizontal-Spatial-Resolution-Cross-track# (* ?h (* ?ara (/ (pi) 180)) )) )
  )
```

Figure 2: CLIPS code of the attribute inheritance rule that computes horizontal spatial resolution from orbit altitude and instrument angular resolution. Note that this is a simplification, and the actual rule contains more involved trigonometric calculations.

Instrument Capability Rules— The main relationships between functional elements (measurements) and formal elements (instruments) are encoded in instrument capability rules. In the actual RBES, in order to increase flexibility and usability, instrument-specific instrument capability rules are automatically generated in a pre-processing step from a spreadsheet containing the user-input measurements - and their attributes that are independent of the mission attributes - that each instrument can take. Instrument capability rules have the structure presented in Figure 3. Whenever a new instrument is asserted, several measurement facts are asserted. With an ideal knowledge of the state-of-the-art of every discipline of remote sensing, instrument capability rules could be completely independent of the EOSS at hand. For example, L-band radiometers are likely to always be capable of measuring soil moisture, regardless of whether it is a NASA mission (e.g. SMAP) or an ESA mission (e.g. SMOS). However, the framework is flexible enough to allow by-passing this feature for cases where knowledge is incomplete, or simply the user wants to increase computational speed. Therefore, instrument-specific capability rules can also be used.

```
(defrule AIRS-measurements
  "Define measurement capabilities of AIRS instrument "
 (Manifested-instrument (Name AIRS) (orbit-altitude# ?h) (orbit-anomaly# ?ano) (orbit-RAAN ?raan) (flies-in ?miss))

🡒
  (assert (Measurement (Parameter "2.5.1 Surface temperature -land-") (Temporal-resolution High-1h-1day) (Accuracy High) (flies-in ?miss) (Id AIRS1) (Instrument AIRS) …)
  (assert (Measurement (Parameter "1.2.1 Atmospheric temperature fields (Temporal-resolution High-1h-1day) (Accuracy High) (flies-in ?miss) (Id AIRS2) (Instrument AIRS) …)
…
)
```

Figure 3: CLIPS code of an instrument capability rule

Synergy Rules—Instrument capability and attribute inheritance rules are not enough to define the emergent behavior that appears when dealing with satellite measurements. After solving the inverse problems, i.e. the retrieval of some property of interest from the magnitude measured by the instrument, scientists can sometimes apply data processing algorithms such as assimilation algorithms, or disaggregation schemes, which will produce new data products with properties that none of the initial data products has. Since these emergent data products can satisfy requirements that the initial data products cannot, it is very important to be able to model them. This is the role of the synergy rules. An example of a synergy rule is provided in Figure 4. [59]**.** This particular rule will apply a disaggregation scheme as the one previously described. Note that even though there are only a few synergy rules declared, they are extremely powerful in generating new measurement facts. Synergy rules are effectively the way the RBES framework models systems emergence. Synergy rules are EOSS-independent, i.e. they do not depend on the EOSS being considered, but they may be measurement dependent, i.e. some rules modeling particular data processing algorithms may only apply to a certain measurement (e.g. soil moisture). Note that this piece of pseudo-code assumes the existence of the functions: a) cross-registered, which makes sure that the measurements are cross-registered; b) fuzzy-max and fuzzy-avg, which compute the maximum and average values of two fuzzy attributes (e.g. max(High, Low) = High, avg(High,Low) = Medium). Furthermore, this piece of pseudo-code has been simplified for the sake of clarity, and therefore does not strictly follow CLIPS/Jess syntax in some details.

```
(defrule spatial-disaggregation
"A frequent coarse spatial resolution measurement can be combined with a sparse high spatial resolution measurement to produce a frequent high spatial resolution measurement with average accuracy"

  ?m1 ← (Measurement (Parameter ?p) (Temporal-resolution ?tr1) (Horizontal-Spatial-Resolution ?hsr1) (Accuracy ?a1) (Id ?id1))
  ?m2 ← (Measurement (Parameter ?p) (Temporal-resolution ?tr2) (Horizontal-Spatial-Resolution ?hsr2) (Accuracy ?a2) (Id ?id2))
  (test (cross-registered (measurements ?id1 ?id2)))

→

  (duplicate ?m1 (Parameter ?p) (Temporal-resolution (eval (fuzzy-max Temporal-resolution ?tr1 ?tr2)))
        (Horizontal-Spatial-Resolution (eval (fuzzy-max Horizontal-Spatial-Resolution ?hsr1 ?hsr2)))
        (Accuracy (eval (fuzzy-avg ?a1 ?a2))))
)
```

Figure 4: CLIPS code of a synergy rule modeling a generic spatial disaggregation scheme such as the one proposed for the SMAP mission

Requirement Satisfaction Rules—Requirement satisfaction rules express stakeholder needs and goals in the form of measurement requirements. Thus, when the requirement rules are executed, the RBES performs the attribute-level comparison between the measurements performed by the EOSS (the instrument capability, attribute inheritance, and synergy rules have all already been applied), and the measurement requirements. There are two types of requirement satisfaction rules: full satisfaction rules, and partial satisfaction rules. Full satisfaction rules express the level of data quality and quantity (i.e. measurement attributes) required for full satisfaction of a particular stakeholder objective. Partial satisfaction rules cover many degraded cases in which one or more attributes are not at the required level for full satisfaction, resulting in a partial loss of benefit.

The general structure of a requirement rule is provided in Figure 5. Like instrument capability rules, requirement satisfaction rules are imported from a spreadsheet containing the numerical data for each objective. Requirement satisfaction rules are obviously EOSS-specific.

```
(defrule full-satisfaction-subobjective-C1-2
  "Defines measurement requirements for full satisfaction of subobjective 1.2 of the climate panel in the Decadal"

(Measurement (Parameter "Soil Moisture") (Horizontal-Spatial-Resolution ?hsr) (Temporal-Resolution ?tr) (taken-by ?instr)
(test (SameOrBetter Horizontal-Spatial-Resolution ?hsr Low-1km-10km))
(test (SameOrBetter Temporal-Resolution ?tr Medium-1day-3days))

→

 (assert (full-satisfaction (of-objective C1-2) (by ?instr)))
  )
```

Figure 5: CLIPS code of a requirement satisfaction rule. This piece of pseudo-code assumes the existence of the function SameOrBetter, which tests whether a fuzzy attribute value is better than another one (e.g. High temporal resolution is better than low temporal resolution)

Value Aggregation Rules—The current implementation of the RBES allows for four different hierarchical levels of value decomposition: 1) overall EOSS value: a single number representing the aggregated scientific and societal benefit of the EOSS architecture; 2) panel value: a set of numbers representing the scientific and societal benefit of the EOSS architecture to the major stakeholder groups; 3) objective satisfaction: a set of numbers representing the degree of satisfaction of individual stakeholder objectives; 4) subobjective satisfaction: a set of numbers representing the degree of satisfaction of the different subobjectives of which an objective consists.

The comparison between the achieved and required measurement sets is done at the subobjective level, i.e. three levels below the overall EOSS value in the hierarchy. We propose that this removes some of the subjectivity in the knowledge elicitation process, since experts are less reluctant to make an assessment at this higher level of fidelity, and they will probably do it more accurately. Value aggregation rules combine individual subobjective requirement satisfaction into objective satisfaction, objective satisfaction into panel satisfaction, and panel satisfaction into EOSS value. Value aggregation rules are also EOSS-specific. In their simplest form, value aggregation rules can use weighted averages of subobjective satisfaction to infer objective satisfaction, weighted averages of objective satisfaction to infer panel satisfaction, and a weighted average of panel satisfaction as metric for overall EOSS value. More sophisticated value aggregation rules may include non-linear terms such as Boolean expressions (and, or, not), maximum and minimum values for certain metrics, or non-linear utility functions (e.g. a logarithmic utility curve to account for risk aversion) amongst others.

Fuzzy Attributes Rules—Traditionally, RBES are naturally capable of dealing with inexact reasoning and uncertain statements [36]. While our RBES does not have a full fuzzy reasoning capability, some rules were created in order to deal with the simultaneous presence of both quantitative and semi-quantitative information. For example, a scientist may express a requirement for horizontal spatial resolution using a numerical value (e.g. 250m), an interval (e.g. anything between 100 and 400m), or a fuzzy or ambiguous value (e.g. very high). The RBES needs to be able of going back and forth from the quantitative and semi-quantitative world. In order to do that, fuzzy attribute rules are defined so that a mapping can be specified by the user to make the link between fuzzy attributes and numerical attributes. Note that these mappings can be in some cases completely dependent on the application. For example, high spatial resolution can be 100m for hydrology, or 1m for disaster monitoring.

Fact Databases—In addition to all the rule sets presented in this section, there are two fact databases that are added into working memory at the initialization of the RBES: a) mission analysis database; b) instrument characteristics database. The mission analysis database is used by the attribute inheritance rules in order to compute the temporal resolution of a satellite or constellation. It contains coverage figures of merit, calculated off-line with dedicated simulation software (AGI's Satellite Tool Kit ® a.k.a. STK) for the most common orbits for satellites and constellations in EOSS. This includes sun-synchronous orbits, true polar orbits, and low inclination orbits between 265km and 1000km of altitude, all of them circular. For each orbit, revisit time was computed using STK for several different coverage grids (global, US, tropical regions, cold regions), and then then the average over time, and worst-case on latitude-longitude is reported on the database. During execution of the attribute inheritance rules in the RBES, the revisit time information corresponding to the relevant satellite or constellation is retrieved. If the information is not available in the database, an STK session can be opened to calculate the necessary data and save it in the database for later execution. The instrument characteristics database is also used by the attribute inheritance rules that copy attribute values from the instrument database to the actual manifested instrument facts. When an instrument is manifested, its characteristics are copied from the database. Note that the instrument characteristics database may be specific to a given case study, or it can also be used as a way to store information about past, present, and future instruments from several agencies and organizations.

Summary—As a summary of the section, Figure 6 shows the flow of execution of the RBES. From the EOSS architecture, the corresponding mission and instrument facts are asserted. Instrument capability rules assert the measurement facts associated to each of the manifested instruments. Attribute inheritance rules regulate how attributes are inherited between missions, instruments, and measurements: instrument attributes are inherited either from their mission or from the corresponding instrument in the database; measurement attributes are inherited from their instrument, their mission, or they are derived from combinations of other mission and instrument attributes. The initial set of measurement capabilities is completed with new and modified measurements through the synergy rules. This new set of measurement capabilities is compared against measurement requirements defined in the requirement rules, which assert subobjective satisfaction facts (full or partial satisfaction). Subobjective satisfaction facts are logically combined to produce objective satisfaction facts, and objective satisfaction facts are combined to produce panel satisfaction metrics. These two steps occur through the value aggregation rules. Finally, panel satisfaction metrics are weighted to produce an overall EOSS score.

3. APPLICATION

The RBES framework was applied to the Earth Science Decadal Survey as a case study. This section is organized as follows: first, the context of the Decadal Survey is described; second, an overview of the rules that are specific to this case study (i.e. requirement satisfaction rules, value aggregation rules, and instrument capability rules) is provided; finally, the results of applying the RBES to the Decadal Survey are discussed, focusing on the scientific and societal value of the different instruments and missions.

Context— In 2004, the NASA Office of Earth Science, the National Oceanic and Atmospheric Administration (NOAA) National Environmental Satellite Data and Information Service (NESDIS), and the U.S. Geological Survey (USGS) Geography Division asked the National Research Council (NRC) Space Studies Board (SSB) to "conduct a decadal survey to generate consensus recommendations from the Earth and environmental science and applications communities regarding a systems approach to space-based and ancillary observations that encompasses the research programs of NASA; the related operational programs of NOAA; and associated programs such as Landsat, a joint initiative of USGS and NASA" [60]. In response to this request, an ad-hoc NRC committee consisting of experts from different disciplines of Earth sciences produced a report known in the community as the "Decadal Survey" (DS). The DS lays out a reference architecture for an integrated Earth Observation Satellite System for the next decade that will fulfill the needs of all the scientific communities in terms of space-based measurements, while also providing essential societal benefits [60].

This reference architecture consists of 15 missions for NASA and 2 missions for NOAA. A total of 39 instruments are assigned to these 17 missions on the basis of a variety of technical, scientific, and programmatic factors including synergies between instruments, data continuity, orbit compatibility, different instrument maturity levels, and expected yearly budget. For each mission, the report provides a description of the scientific objectives fulfilled by the mission, the physical parameters measured, the instruments used, the orbit required, a rough estimation of the lifecycle mission cost, and the expected mission launch date.

However, some of the assumptions used in the study such as mission cost, yearly budget and precursor missions have now changed. Mission cost estimates have grown on average by 64% between the time where the report was issued in 2007 and today, as shown in Figure 7. The standard deviation is 59%, and the range of increases goes from 2% to 200%. Note that the latest cost estimates may be unfair in that missions that have not started their development cannot possibly have any schedule slippage or cost overrun. Second, yearly budget has decreased by about 50% with respect to the $750M/yr used in the Decadal Survey. And finally, some of the precursor missions have failed or have been severely delayed (e.g. the Orbiting Carbon Observatory mission or OCO, and the National Polar-orbiting Operational Environmental Satellite System, or NPOESS.) Therefore, the question arises whether this is still the best architecture possible given the current assumptions. The purpose of this paper is to lay out a methodology to explore alternative architectures in a systematic way.

Value Aggregation Rules for the Decadal Survey—The Decadal Survey was organized in six panels: climate, weather, land and ecosystems, water, health, and solid Earth. Each panel provided a list of scientific and societal objectives, and each objective typically concerns more than one measurement. A 3-layer structure for value decomposition was used based on this, with 6 panels in the first level, 35 objectives in the second level, and 186 subobjectives in the third level (see Figure 8). Numerical values for the relative importance of panels, objectives within a panel, or subobjectives within an objective, were taken from previous work on the Decadal Survey [28], [61], which in turn inferred their values from various sources including the Decadal Survey report itself, several interviews with NASA, and other sources such as congress hearings. Despite the rigor and effort put into the obtention of these numerical values, we acknowledge the uncertainty and subjectivity that may surround this process. In order to partially mitigate this limitation, a spreadsheet-based user interface was built in order to allow for rapid what-if analysis. "What-if" analyses are a common form of sensitivity analysis between decision makers and program managers. A snapshot of this spreadsheet is provided in Figure 8. These numerical values are used to compute the scientific-societal benefit of a Decadal instrument or mission from the set of subobjective satisfaction variables.

Requirement Satisfaction Rules for the Decadal Survey—For each of the 186 subobjectives, a full satisfaction rule and several partial satisfaction rules were defined, totaling 1503 rules. The numerical values for subobjective satisfaction related to each partial satisfaction rule are also introduced through a spreadsheet in order to facilitate what-if analysis. A snapshot of this spreadsheet is given in Figure 9. These rules were obtained from several sources: a) the appendices to the Decadal Survey report containing detailed information about each panel; b) several interviews with senior scientists at MIT and NASA Goddard Space Flight center; c) an extensive literature review.

Instrument Capability rules for the Decadal Survey—For each of the 39 instruments in the Decadal Survey, a set of instrument capability rules was created based on the information available in the Decadal Survey report, several publications devoted to individual instruments. Instrument capability rules are also introduced through a spreadsheet to allow for fast sensitivity analysis. A snapshot of this spreadsheet is shown in Figure 10. Figure 10 consists of two parts: a) a list of the attributes of each instrument, some of which are inherited by all its measurements such as spectral band, or day-night capability; b) a list of the measurements taken by each instrument (e.g. soil moisture, atmospheric temperature fields), together with the measurement attributes (spatial resolution, temporal resolution, signal-to-noise ratio).

Results—The RBES was applied to each mission in the Decadal Survey in isolation, i.e., accounting for synergies between instruments on the same mission, but not for synergies across missions. The results of this analysis are shown in Figure 11. Note that, for each mission the six panel scores were weighted according to the panel weights defined in the aggregation rules presented in Figure 8: The RBES provides a full explanation of the scores of all instruments to all panels. Since this is too much data to include in the context of a paper, one mission, namely the SMAP mission, is commented in detail. The SMAP mission carries two instruments sharing a common dish: an L-band passive radiometer, and an L-band radar. The characteristics of both instruments are provided in Table 2.

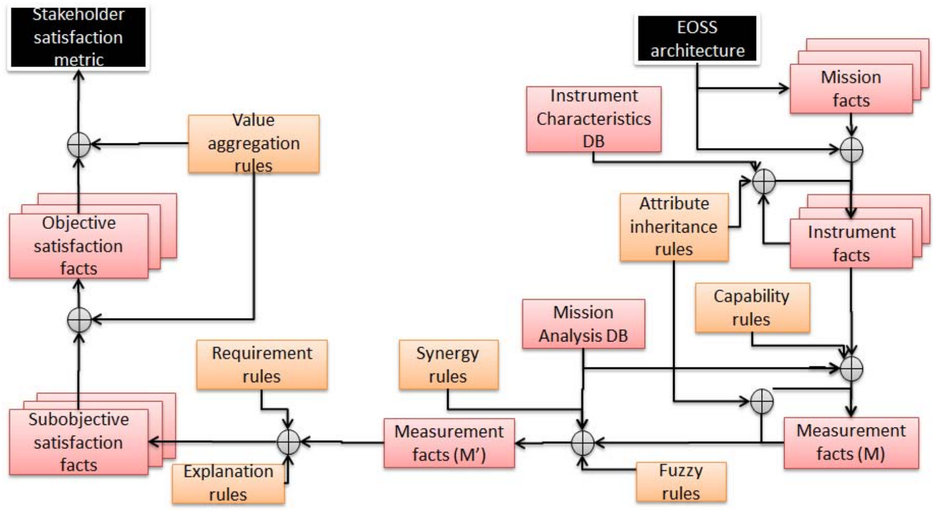


Figure 6: Flow of execution in the Rule-based Expert System for assessing the scientific value of EOSS

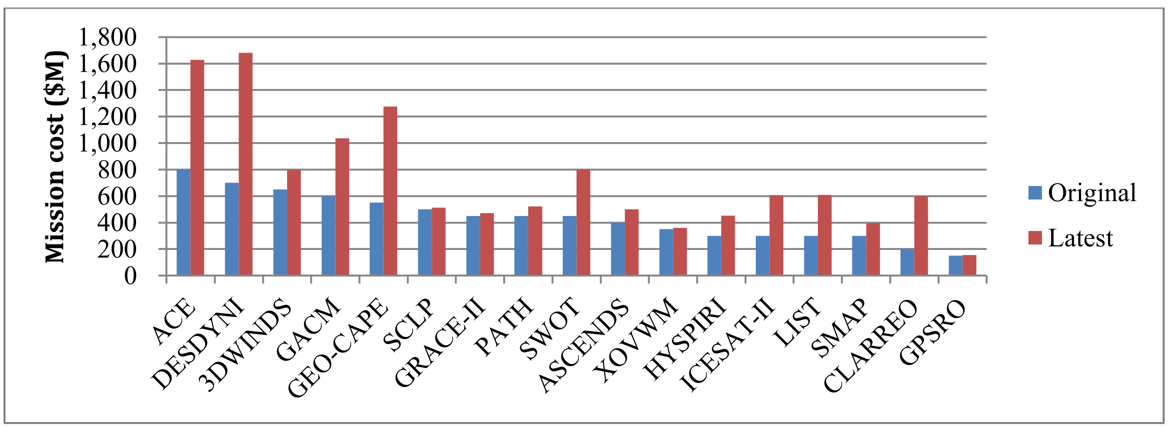


Figure 7: Decadal Mission Cost Estimates

1st LEVEL OF STAKEHOLDER NEEDS DECOMPOSITION (PANELS)

Panel	Id	Description	Weight
Weather	WE	Weather	21%
Climate	C	Climate	21%
Land	ECO	Land and Ecosystems	21%
Water	WA	Water	16%
Health	HE	Human health	11%
Solid Earth	SE	Solid Earth	11%
			100%

2nd LEVEL OF STAKEHOLDER NEEDS DECOMPOSITION (OBJECTIVES)

Weather panel

Objec	Id	Description	Weight
1	WE1	Atmospheric winds	19%
2	WE2	High temporal resolution air pollution	15%
3	WE3	All-weather temperature and humidity profiles	12%
4	WE4	Comprehensive global tropospheric aerosol characte	8%
5	WE5	Radio Occultation	19%
6	WE6	Comprehensive global tropospheric O3 measuremer	15%
7	WE7	Aerosol-cloud discovery	12%
			100%

Climate panel

Objec	Id	Description	Weight
1	C1	Aerosol-Cloud Forcing	24%
2	C2	Ice Sheet, Sea Ice Volume and Ice Dynamics	24%
3	C3	Carbon Sources and Sinks	24%
4	C4	Radiance Calibration and Time-Reference Observato	18%
5	C5	Ocean Circulation, Heat Storage, and Climate Forcing	12%
			100%

Figure 8: Spreadsheet for introducing value aggregation rules in the Decadal Survey case study

Table 2: SMAP instrument capabilities used for the RBES. Note that the numerical values are transformed by the RBES into semi-quantitative fuzzy values (e.g. “High”, “Medium”) through domain-specific fuzzy attribute rules.

Characteristic	**Radar**	**Radiometer**
Horizontal-Spatial-Resolution#	3km	40km
Accuracy#	1dB (space-averaged at @10km)	0.64K in brightness temperature = 0.01-0.04 cm^3/cm^3 @ 40km (soil moisture)
Polarizations	HH, VV, HV	H, V
Soil penetration	5cm	5cm
Illumination	Active	Passive
Swath	1000km	1000km
Frequency	1.26-1.29 GHz	1.41 GHz

This information was gathered from: a) Decadal Survey report itself [60]; b) several papers describing the mission, the instruments, and the retrieval algorithms [62], [63], [64], [59].as well as similar missions such as ESA’s SMOS mission [65], [20]; c) an interview with one of the senior scientists of the mission at MIT.

Note that the radar has higher spatial resolution but lower accuracy or sensitivity than the passive radiometer. Both instruments by themselves are capable of measuring soil moisture, but none of them by themselves are capable of fully meeting the soil moisture requirement of 4% accuracy and 10km spatial resolution (e.g. a simple resampling of the radiometer data will yield a 5% accuracy [59]). It is necessary to combine the data produced by the two instruments and process it to achieve the scientific requirement.

This emergent or synergistic behavior is captured by the synergy rules. A synergy rule was created that models at a high level the algorithms proposed in [59] or [20], i.e. the combination of a high spatial resolution. In addition to the soil moisture measurement, the radar adds the capability of measuring freeze-thaw state, for which there is also a requirement in the Decadal Survey, and to a lesser extent ocean surface wind speed and direction. As for the radiometer, in addition to soil moisture, it can also measure to a lesser extent ocean surface salinity, ocean surface wind speed and direction, and provide useful information for ice and snow cover retrieval.

Taking all these capabilities into account, and comparing to the requirements set by the Decadal Survey, the scores in Table 3 are obtained for the SMAP mission.

Table 3: SMAP panel scores obtained by the RBES. The weighted sum of the panel scores using the weights defined in Figure 8 yields SMAP’s score on Figure 12.

Panel	Score = % panel objectives satisfied
Weather	0.0%
Climate	1.0%
Land and Ecosystems	0.0%
Water	28.7%
Human health	6.0%
Solid Earth	0.0%
Weighted score	**5.4%**

As shown in Table 3, according to our recollection of the Decadal Survey objectives, the SMAP mission provides value to the water and human health panels primarily, and only marginal value to the climate panel.

At this point the users may ask themselves about the reasoning behind these scores. As mentioned before, RBES have a built-in capability to answer this question because the simple trace of the rules that were fired contains all the relevant information. The explanation provided by the RBES about SMAP’s scores is given in Table 4. Table 4 explicitly shows the traceability between subobjectives and data products from instruments or combinations of instruments. Note for example that the soil moisture subobjectives from the water panel are fully satisfied by a synergistic product that comes from applying a disaggregation scheme to the combination of the radar and radiometer individual products. The human health requirements in terms of soil moisture are as stringent in spatial and temporal resolution, but less stringent in accuracy, and therefore they are fully satisfied by the radar alone.

Finally, climate subobjectives concerning sea surface winds are partially satisfied by wind data products potentially produced using SMAP’s radar and polarimetric radiometer. The explanation facility shows that these objectives are not fully satisfied because the products do not have enough accuracy compared to the required one (the baseline for these measurement comes from the Decadal XOVWM that carries several SAR scatterometers operating at more sensitive frequencies.

Note that the individual scores assume the presence of all static and dynamic ancillary data, namely soil and surface air temperature, vegetation water content, sand and clay fraction, urban area, percentage of permanent open water, crop type, land cover class, precipitation, snow, mountainous area, permanent ice, vegetation parameters, and soil roughness [64]. The tool user may decide to make the presence of this ancillary data a requirement that must be fulfilled by other missions in the Decadal Survey.

.

Subobject	rule	value	description	Measurement	Attribute-value4	Attribute-value5
WE1-1	nominal	100%	"Conditions for full satisfaction "	"1.4.1 atmospheric wind speed"	SameOrBetter Temporal-resolution High-1h-1day	Accuracy High
WE1-1	degraded-1	67%	"Only most of the region of interest is covered"	"1.4.1 atmospheric wind speed"	SameOrBetter Temporal-resolution High-1h-1day	Accuracy High
WE1-1	degraded-2	33%	"Only some of the region of interest is covered"	"1.4.1 atmospheric wind speed"	SameOrBetter Temporal-resolution High-1h-1day	Accuracy High
WE1-1	degraded-3	50%	"Insufficient accuracy"	"1.4.1 atmospheric wind speed"	SameOrBetter Temporal-resolution High-1h-1day	**Accuracy Medium**
WE1-1	degraded-4	50%	"Insufficient vertical spatial resolution"	"1.4.1 atmospheric wind speed"	SameOrBetter Temporal-resolution High-1h-1day	Accuracy High
WE1-1	degraded-5	50%	"Insufficient temporal resolution (1-3 days)"	"1.4.1 atmospheric wind speed"	**SameOrBetter Temporal-resolution Medium-1day-3c**	Accuracy High
WE1-1	degraded-6	25%	"Insufficient temporal resolution (over 1 week)"	"1.4.1 atmospheric wind speed"	**SameOrBetter Temporal-resolution Low-3days-1-we**	Accuracy High
WE1-1	degraded-7	50%	"Insufficient sensitivity at high latitudes"	"1.4.1 atmospheric wind speed"	SameOrBetter Temporal-resolution High-1h-1day	Accuracy High
WE1-2	nominal	100%	"Conditions for full satisfaction "	"1.4.2 atmospheric wind direction"	SameOrBetter Temporal-resolution High-1h-1day	Accuracy High
WE1-2	degraded-1	67%	"Only most of the region of interest is covered"	"1.4.2 atmospheric wind direction"	SameOrBetter Temporal-resolution High-1h-1day	Accuracy High
WE1-2	degraded-2	33%	"Only some of the region of interest is covered"	"1.4.2 atmospheric wind direction"	SameOrBetter Temporal-resolution High-1h-1day	Accuracy High
WE1-2	degraded-3	50%	"Insufficient accuracy"	"1.4.2 atmospheric wind direction"	SameOrBetter Temporal-resolution High-1h-1day	**Accuracy Medium**
WE1-2	degraded-4	50%	"Insufficient vertical spatial resolution"	"1.4.2 atmospheric wind direction"	SameOrBetter Temporal-resolution High-1h-1day	Accuracy High
WE1-2	degraded-5	50%	"Insufficient temporal resolution (1-3 days)"	"1.4.2 atmospheric wind direction"	**SameOrBetter Temporal-resolution Medium-1day-3c**	Accuracy High
WE1-2	degraded-6	25%	"Insufficient temporal resolution (over 1 week)"	"1.4.2 atmospheric wind direction"	**SameOrBetter Temporal-resolution Low-3days-1-we**	Accuracy High
WE1-2	degraded-7	50%	"Insufficient sensitivity at high latitudes"	"1.4.2 atmospheric wind direction"	SameOrBetter Temporal-resolution High-1h-1day	Accuracy High

Figure 9: Spreadsheet for introducing requirement satisfaction rules in the Decadal Survey case study. Each line in this spreadhseet will be transformed into one requirement satisfaction rule. Note the presence of both full satisfaction rules (labeled as "nominal") and partial satisfaction rules (labeled as "degraded").

Parameter	Attribute-Value1	Attribute-Value2	Attribute-Value3
"2.2.1 surface deformation"	Horizontal-Spatial-Resolution High-10-100m	Temporal-resolution Low-3days-1-week	Accuracy High
"2.2.2 Hi-res topography"	Horizontal-Spatial-Resolution High-10-100m	Temporal-resolution Low-3days-1-week	Accuracy Medium
"2.3.2 soil moisture"	Horizontal-Spatial-Resolution High-10-100m	Temporal-resolution Low-3days-1-week	Accuracy Medium
"2.4.1 vegetation type and structure"	Horizontal-Spatial-Resolution High-10-100m	Temporal-resolution Low-3days-1-week	Accuracy High
"2.4.4 canopy density"	Horizontal-Spatial-Resolution High-10-100m	Temporal-resolution Low-3days-1-week	Accuracy High
"2.6.1 land use"	Horizontal-Spatial-Resolution High-10-100m	Temporal-resolution Low-3days-1-week	Accuracy High
"2.6.2 landcover status"	Horizontal-Spatial-Resolution High-10-100m	Temporal-resolution Low-3days-1-week	Accuracy High
"2.6.3 disaster monitoring"	Horizontal-Spatial-Resolution High-10-100m	Temporal-resolution Low-3days-1-week	Accuracy High
"2.6.4 hydrocarbon reservoir monitoring"	Horizontal-Spatial-Resolution High-10-100m	Temporal-resolution Low-3days-1-week	Accuracy High
"4.1.4 Ice sheet velocity"	Horizontal-Spatial-Resolution High-10-100m	Temporal-resolution Low-3days-1-week	Accuracy High
"4.1.5 Ice Sheet topography"	Horizontal-Spatial-Resolution High-10-100m	Temporal-resolution Low-3days-1-week	Accuracy Medium
"4.2.4 snow cover"	Horizontal-Spatial-Resolution High-10-100m	Temporal-resolution Low-3days-1-week	Accuracy Low
"4.3.1 Sea ice thickness"	Horizontal-Spatial-Resolution High-10-100m	Temporal-resolution Low-3days-1-week	Accuracy Low
"4.3.2 Sea ice cover"	Horizontal-Spatial-Resolution High-10-100m	Temporal-resolution Low-3days-1-week	Accuracy Low

Name	Attribute pair1	Attribute pair2	Attribute pair3	Attribute pair4	Attribute pair5
ACE_CPR	Intent "Cloud profile and rain radars"	Illumination Active	All-weather yes	Day-Night Day-and-night	Penetration Medium-other-MW
ACE_ORCA	Intent "Imaging multi-spectral radiometers -passive optical-"	Illumination Passive	All-weather no	Day-Night Only-day	Penetration Low-optical
ACE_POL	Intent "Imaging multi-spectral radiometers -passive optical-"	Illumination Passive	All-weather no	Day-Night Only-day	Penetration Low-optical
ACE_LID	Intent "Laser altimeters"	Illumination Active	All-weather no	Day-Night Day-and-night	Penetration Low-optical
ASC_LID	Intent "Laser altimeters"	Illumination Active	All-weather no	Day-Night Day-and-night	Penetration Low-optical
ASC_GCR	Intent "Gas filter correlation radiometer"	Illumination Passive	All-weather no	Day-Night Day-and-night	Penetration Low-optical
ASC_IRR	Intent "IR Atmospheric temperature and humidity sounders"	Illumination Passive	All-weather no	Day-Night Day-and-night	Penetration Low-optical
ASC_LAS	Intent "Laser altimeters"	Illumination Active	All-weather no	Day-Night Day-and-night	Penetration Low-optical
CLAR_TIR	Intent "Earth radiation budget radiometers"	Illumination Passive	All-weather no	Day-Night Only-day	Penetration Low-optical
CLAR_VNIR	Intent "Earth radiation budget radiometers"	Illumination Passive	All-weather no	Day-Night Only-day	Penetration Low-optical
CLAR_GPS	Intent "GNSS-Sun-Moon-Stars-Radio Occultation instruments"	Illumination Passive	All-weather yes	Day-Night Day-and-night	Penetration High-P-or-L-band
DESD_SAR	Intent "Imaging MW radars -SAR-"	Illumination Active	All-weather yes	Day-Night Day-and-night	Penetration High-P-or-L-band

Figure 10: Spreadsheets for introducing instrument capability rules in the Decadal Survey case study. The top snapshot corresponds to the list of parameters that a certain instrument can measure. One line corresponds to one parameter being measured, and any number of measurement attributes (e.g. horizontal spatial resolution) may be specified. The bottom snapshot corresponds to the instrument attributes. Each line contains the characteristics of one instrument. Some of these characteristics will be inherited by corresponding measurements through attribute inheritance rules.

A similar process to the one described for the SMAP mission was conducted for each mission, leading to the results shown in Figure 11. The highest scientific score was thus achieved by the DESDYNI and HYSPIRI missions. The combination of the thermal and visual imaging spectrometers in the HYSPIRI mission was modeled as being able to measure cloud type, sea ice cover (with medium accuracy), vegetation type, vegetation state, land surface temperature, soil moisture (only the first mm of the surface). These measurements, with their attributes, satisfy objectives from all panels except for the weather panel, which explains the high score. Concerning DESDYNI, the L-band SAR was modeled as being able to measure primarily surface deformation, and vegetation structure, and a variety of other measurements to a lesser extent, including topography, land use and land cover status, hydrocarbon reservoirs, disaster monitoring, ice cover, and snow cover. The lidar was modeled as being able to perform vegetation measurements as well as some topography, and aerosol and cloud properties. This synergistic combination of measurements satisfies objectives of 5 out of 6 panels, leaving only the weather panel unsatisfied.

The missions with the lowest science score are XOVWM and 3D-WINDS. XOVWM provides high spatial resolution ocean surface wind vectors (speed and direction), which is relatively important for the climate panel because of ocean circulation, and heat storage, but it provides only marginal or no value to all other panels. It should be noted that part of the value of the XOVWM mission is on the fact that it continues a long data series that would otherwise be compromised. This is not taken into account in this study, even though this capability is being added into the RBES. Concerning 3D-WINDS, it will be the first space-based dual coherent/non-coherent Doppler wind lidar. This will result in the first global, high accuracy three dimensional direct measurements of atmospheric winds, with high sensitivity in the troposphere. However, the requirements for atmospheric winds have high temporal resolutions that are unachievable by lidar. As a consequence, the 3D-WINDS mission gets a low value because all the corresponding subobjectives are only partially satisfied. This situation is partially reversed when the 3D-WINDS mission is considered in the context of the overall EOSS, i.e. when synergies across missions are considered. That habilitates the use of 3D-WINDS data with other lower accuracy, more frequent datasets to yield new data products that can fully satisfy wind objectives. While one may not agree with one or more of these modeling assumptions, the RBES is transparent and has the capability to explain its results. Furthermore, one can change the assumptions and rerun the model in a matter of seconds or a few minutes.

Missions were then ranked according to their cost-effectiveness, defined as the ration between the weighted scientific-societal benefit derived from the RBES and a mission cost estimate. Two mission cost estimates were utilized: the initial NRC estimates in the Decadal Survey, and the latest available cost estimates from a variety of sources at NASA, as indicated in Figure 7. The results of the cost-effectiveness analysis are provided in Figure 12. While the weighted science scores of the missions do not vary too much when changing from the weights in Figure 7 to uniform weights, the great variability in cost increase leads to changes in cost-effectiveness between the two charts in Figure 12. Missions with cost increase greater than the average cost increase of 64% such as CLARREO and DESDYNI generally lost positions in the ranking, while missions with a lower cost increase such as SMAP and SCLP typically gained a few positions. It should be noted again that this cost-effectiveness ranking may be biased by the fact that some missions have not yet started their development and therefore their cost estimate has stayed almost constant.

4. CONCLUSION

Summary—A solution has been presented to assess the scientific and societal value of EOSS using rule-based systems. This solution is computationally fast enough to allow for automatic evaluation of hundreds of different mission concepts as opposed to other methods such as OSSEs, and it has better modeling fidelity and expressivity than simple SVMs. While it still deals with subjective expert assessments, we argue that it is less subjective than SVMs due to the fact that the expert makes the assessment at a higher level of detail, where more information is available. The processes of making and aggregating the judgments at the attribute level are separated. More importantly, its explanation facility ensures transparency and enhances the human-machine interaction.

Limitations—In addition to the limitations inherent to RBES discussed in the introduction, an underlying assumption of this methodology is that the vast majority of scientific and societal needs for EOSS can be expressed in terms of measurement requirements using the attributes defined in the measurement templates. While this is arguably the case for most system requirements that appear on formal mission documents, it is also true that there might be unwritten needs and objectives that are only marginally related to measurement requirements, such as those coming from policy directions. As explained in the body of the paper, the numerical values used in the case study come from a variety of sources including the Decadal Survey report, interviews with senior scientists at MIT and NASA, and several other publications. While standard interview guidelines and rubrics were used, and significant effort was put into obtaining reasonable values, no formal social science method was utilized to elicit the information from the interviewees. Finally, note that in this particular application, only synergies within missions and not across missions were considered, which as it was shown in the case of the 3D-WINDS mission, is an important issue to take into account in the future because every single addendum to the current global EOSS needs to consider both the existing capabilities of the global EOSS and the synergies between the addendum and the global EOSS.

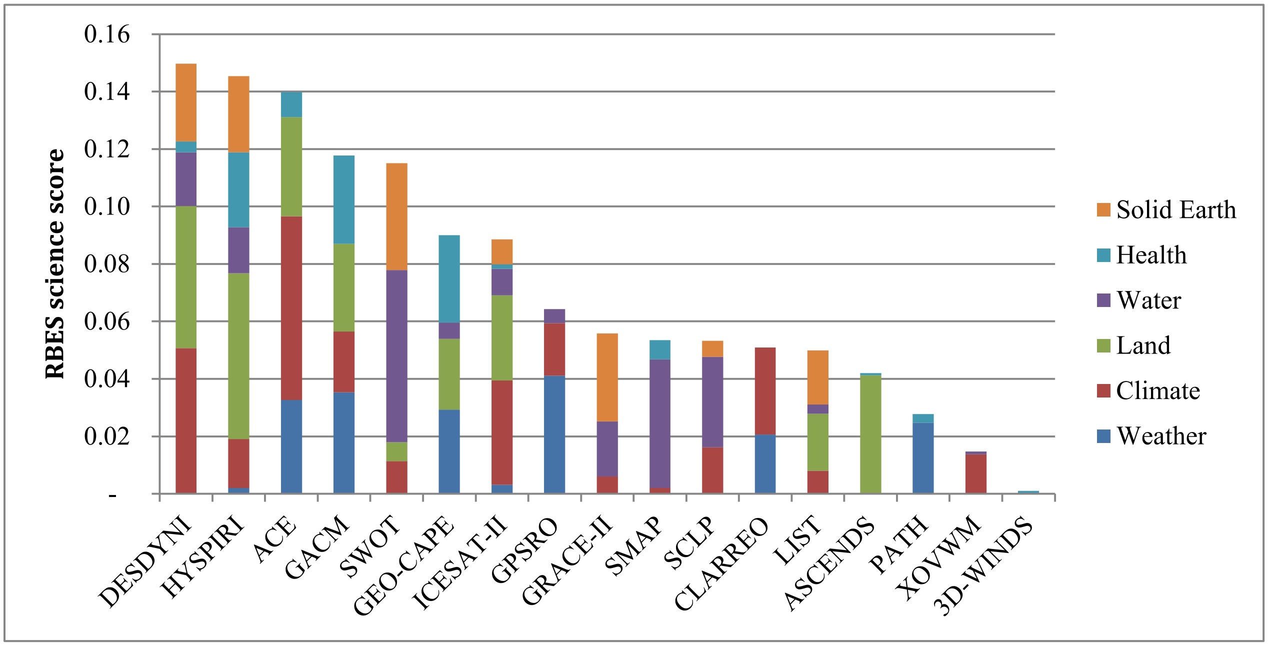
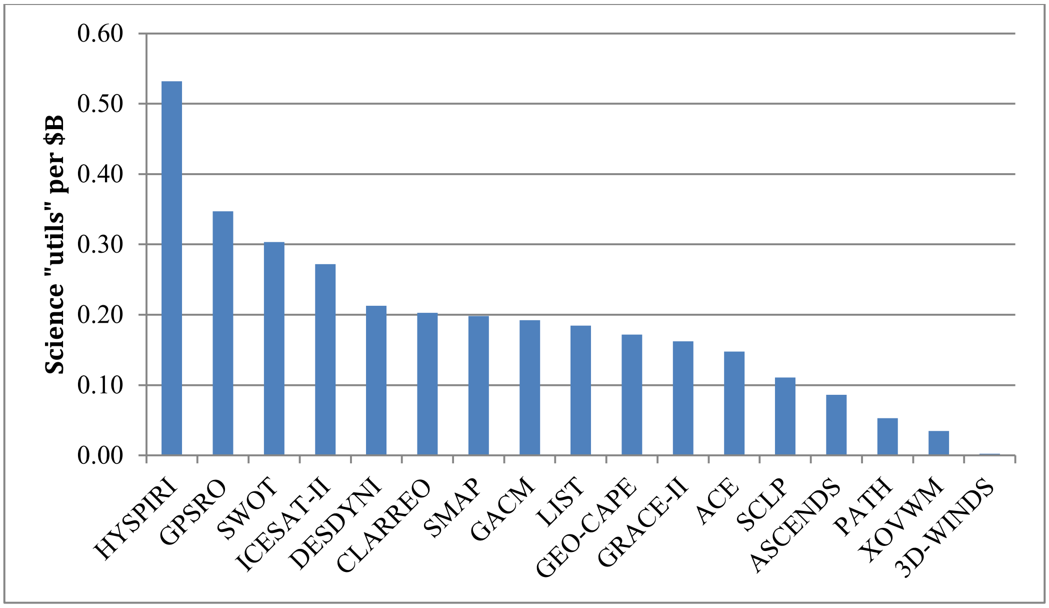


Figure 11: RBES science scores for the 17 Decadal Survey missions, ranked in order of decreasing value. The scores represent a weighted average of the percentage of panel objectives satisfied by the mission.

Table 4: Traceability of results for SMAP mission obtained from explanation facility

Subobjective	Parameter	Objective	Subobjective satisfaction	Data product
C5-3	3.4.1 Ocean surface wind speed	Ocean Circulation, Heat Storage, and Climate Forcing	Partial (insufficient accuracy)	SMAP_RAD-
C5-3	3.4.1 Ocean surface wind speed	Ocean Circulation, Heat Storage, and Climate Forcing	Partial (insufficient temporal resolution)	SMAP_RAD-time-averaged-time-averaged
C5-4	3.4.2 Ocean surface wind direction	Ocean Circulation, Heat Storage, and Climate Forcing	Partial (insufficient accuracy)	SMAP_RAD
C5-4	3.4.2 Ocean surface wind direction	Ocean Circulation, Heat Storage, and Climate Forcing	Partial (insufficient temporal resolution)	SMAP_RAD-time-averaged-time-averaged
HE2-2	2.3.2 soil moisture	Heat Stress and Drought	full	SMAP_RAD
HE2-2	2.3.2 soil moisture	Heat Stress and Drought	full	SMAP_MWR-SMAP_RAD-disaggregated
HE6-4	2.3.2 soil moisture	Vector-borne and Zoonotic Disease	full	SMAP_RAD
HE6-4	2.3.2 soil moisture	Vector-borne and Zoonotic Disease	full	SMAP_MWR-SMAP_RAD-disaggregated
WA1-1	2.3.1 Freeze/thaw state	Soil moisture and freeze-thaw state	full	SMAP_RAD
WA1-2	2.3.2 soil moisture	Soil moisture and freeze-thaw state	full	SMAP_MWR-SMAP_RAD-disaggregated



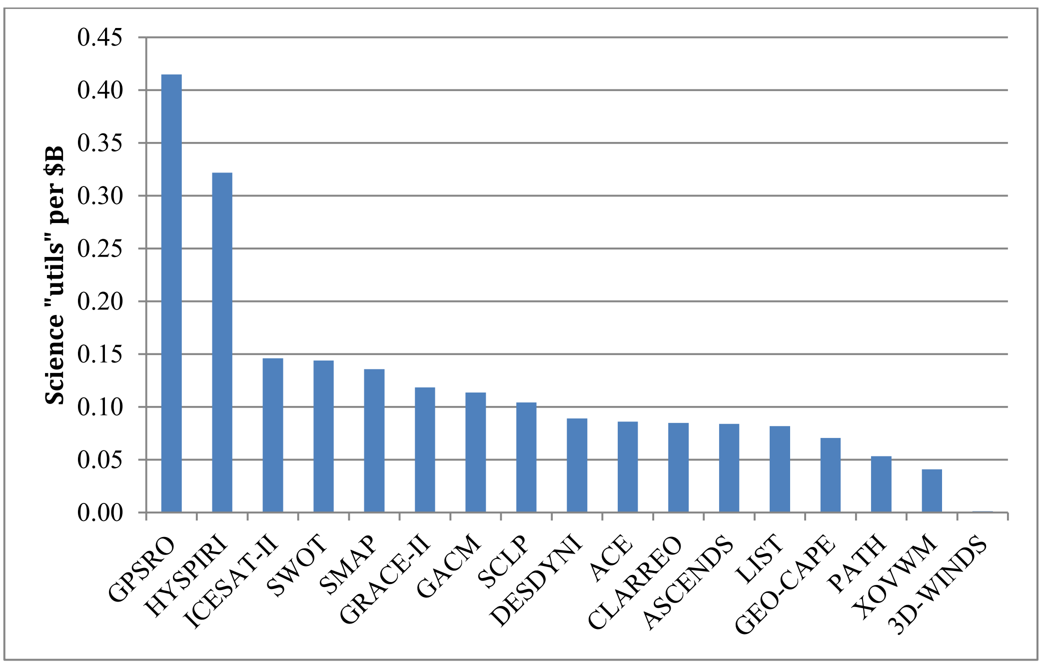


Figure 12: Decadal Missions cost-effectiveness in "utils" per $B, with original costs and unweighted panels (top); updated costs and weighted panels (bottom). "Utils" represent a weighted average of the percentage of panel objectives fully or partially satisfied by the mission.

Future Work—The performance of any RBES is determined by the quantity and quality of rules in the rule database. Therefore, our primary effort going forward will be in the continuous improvement of the different rule bases.

There are several sources of uncertainty in these results: a) the list of measurement capabilities in each instrument's capability rules may be wrong or incomplete; b) the attributes of the same list may be wrong or incomplete; c) the values in the partial satisfaction requirement rules may be wrong or incomplete; d) the relative weights determined in the value aggregation rules may be inaccurate. In this sense, the methodology would benefit from the incorporation of a formal social science method such as the Delphi method [30] to elicit the information from stakeholders. Furthermore, all of these sensitivity analyses could be done automatically using the RBES. Such study could be conducted as follows: first, the score of the EOSS is calculated using the RBES, with the best possible guess for each instrument characteristic given the available information. Then, for each instrument characteristic, all plausible values are identified. For each different value of each different attribute, one at a time, a new science score is calculated using the RBES. Thus, for each attribute, a curve showing the sensitivity of the score to that particular attribute around the reference point can be constructed. These values are similar in nature to the partial derivatives of the score to the different instrument attributes.

Furthermore, while our system provides a basic capability to deal with fuzzy attributes through a set of rules that transform back and forth from semi-quantitative to quantitative values, a more formal application of state-of-the-art fuzzy set theory as first proposed by Zadeh [66] would certainly be beneficial to treat stakeholder satisfaction.

The development of a machine learning layer on top of the fuzzy rule-based system would significantly improve its performance. This machine learning layer could be used for example to perform multi-attribute regression analysis on the mission analysis database and infer simple relationships that allow for fast prediction of the revisit times of a certain constellation on different regions of the Earth

Finally, it was also suggested by some of the interviewees that a policy set of rules be added to the current RBES in order to take into account important rules that might be driving EOSS architecture while being almost independent of science performance.

REFERENCES

[1] M. W. Maier and E. Rechtin, *The Art of Systems Architecting*. New York: CRC press, 2000.

[2] E. Rechtin, *System Architecting, Creating & Building Complex Systems*. Englewood Cliffs, NJ: Prentice Hall, 1991.

[3] E. Crawley et al., "The influence of architecture in engineering systems - Engineering Systems Monograph," 2004.

[4] W. L. Simmons, "A Framework for Decision Support in Systems Architecting," Massachusetts Institute of Technology, 2008.

[5] J. P. Shim, M. Warkentin, J. F. Courtney, D. J. Power, R. Sharda, and C. Carlsson, "Past, present, and future of decision support technology," *Decision support systems*, vol. 33, no. 2, pp. 111-126, Jun. 2002.

[6] W. L. Simmons, B. H. Y. Koo, and E. F. Crawley, "Architecture generation for Moon-Mars exploration using an executable meta-language," in *Proceedings of the AIAA Space 2005 Conference & Exhibition*, 2005, vol. 30.

[7] A. M. Ross, D. E. Hastings, J. M. Warmkessel, and N. P. Diller, "Multi-attribute Tradespace Exploration as Front End for Effective Space System Design," *Journal of Spacecraft and Rockets*, vol. 41, no. 1, 2004.

[8] D. J. Bayley, "Design Optimization of Space Launch Vehicles Using a Genetic Algorithm," Auburn University, 2007.

[9] O. De Weck, R. De Neufville, and M. Chaize, "Staged deployment of communications satellite constellations in low Earth orbit," *Journal of Aerospace Computing, Information, and Communication*, vol. 1, no. 3, pp. 119–136, 2004.

[10] D. McCormick, B. Barrett, and M. Burnside-Clapp, "Analyzing Fractionated Satellite Architectures using RAFTIMATE – A Boeing Tool for Value-Centric Design," in *AIAA SPACE 2009 Conference & Exposition*, 2009.

[11] D. B. Maciuca et al., "A Modular, High-Fidelity Tool to Model the Utility of Fractionated Space Systems," in *AIAA SPACE 2009 Conference & Exposition*, 2009.

[12] E. Eichenberg-bicknell, M. J. Wisniewski, S. W. Choi, and D. M. Westley, "Using a Value-centric Tool Framework to Optimize Lifecycle Cost, Value, and Risk of Spacecraft Architectures," in *AIAA SPACE 2009 Conference & Exposition*, 2009.

[13] J. M. Lafleur and J. H. Saleh, "Exploring the F6 Fractionated Spacecraft Trade Space with GT-FAST," in *AIAA SPACE 2009 Conference & Exposition*, 2009, no. September, pp. 1-22.

[14] L. Hill, K. Leavitt, W. Voshell, and A. Weigel, "A Value Centric Design Approach to Earth Observing Missions," in *AIAA SPACE Conference & Exposition*, 2009.

[15] D. Selva and E. F. Crawley, "Integrated Assessment of Packaging Architectures in Earth Observing Programs," in *Proceedings of the 2011 IEEE Aerospace Conference*, 2010.

[16] H. Apgar, D. Bearden, and R. Wong, "Cost Modeling," in *Space Mission Analysis and Design*, 3rd ed., W. J. Larson and J. R. Wertz, Eds. Microcosm, 1999, pp. 783-820.

[17] Cost Analysis Division - NASA Headquarters, "2008 NASA Cost Estimating Handbook,"

Washington DC, 2008.
[18] J. R. Weiss and W. D. Smythe, "Science Traceability," in *Proceedings of the 2005 IEEE Aerospace Conference*, 2005, pp. 1-8.
[19] W. Nordhaus, "What is the value of scientific knowledge? An application to global warming using the PRICE model," *The Energy Journal*, vol. 39, no. 1, p. 65, Jan. 1998.
[20] A. Camps et al., "The SMOS end-to-end performance simulator: description and scientific applications," in *Proceedings of the 2003 IEEE International Geoscience and Remote Sensing Symposium*, 2003, pp. 13-15.
[21] N. Olsen et al., "The Swarm End-to-End mission simulator study : A demonstration of separating the various contributions to Earth's magnetic field using synthetic data," *Earth, Planets, and Space*, vol. 58, pp. 359-370, 2006.
[22] G. Catastini et al., "The GOCE end-to-end system simulator," in *Proceedings of the 2003 EGS - AGU - EUG Joint Assembly*, 2003.
[23] R. M. Atlas, "Observing System Simulation Experiments: methodology, examples and limitations," in *Proceedings of the WMO Workshop on the Impact of various observing systems on Numerical Weather Prediction*, 1997.
[24] D. G. H. Tan, E. Andersson, M. Fisher, and L. Isaksen, "Observing-system impact assessment using a data assimilation ensemble technique: application to the ADM–Aeolus wind profiling mission," *Quarterly Journal of the Royal Meteorological Society*, vol. 133, no. 623, pp. 381–390, 2007.
[25] a Stoffelen, G. J. Marseille, F. Bouttier, D. Vasiljevic, S. de Haan, and C. Cardinali, "ADM-Aeolus Doppler wind lidar Observing System Simulation Experiment," *Quarterly Journal of the Royal Meteorological Society*, vol. 132, no. 619, pp. 1927-1947, Jul. 2006.
[26] W. T. Crow et al., "An observing system simulation experiment for hydros radiometer-only soil moisture products," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 6, pp. 1289-1303, Jun. 2005.
[27] W. a Lahoz et al., "An observing system simulation experiment to evaluate the scientific merit of wind and ozone measurements from the future SWIFT instrument," *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 606, pp. 503-523, Jan. 2005.
[28] T. Seher, "Campaign-level Science Traceability for Earth Observation System Architecting," Massachusetts Institute of Technology, 2009.
[29] M. Adler et al., "Rapid Mission Architecture Trade Study of Enceladus Mission Concepts," in *Proceedings of the 2011 IEEE Aerospace Conference*, 2011.
[30] N. Dalkey and O. Helmer, "An Experimental Application of the Delphi Method to the Use of Experts," *Management Science*, vol. 9, no. 3, pp. 458-467, 1963.
[31] T. L. Saaty, "Decision Making With the Analytic Hierarchy Process," *International Journal of Services Sciences*, vol. 1, no. 1, pp. 83-98, 2008.
[32] J. Durkin, *Expert Systems: Design and Development*. New York: Macmillan Publishing Company, 1994.
[33] E. A. Feigenbaum, B. G. Buchanan, and J. Lederberg, "On generality and problem solving: a case study using the DENDRAL program," in *Machine Intelligence 6*, B. Meltzer and D. Michie, Eds. Edinburgh, Scotland: , 1971, pp. 165-190.
[34] M. Minsky, "A framework for representing knowledge." Cambridge, MA, 1974.
[35] B. G. Buchanan and E. H. Shortliffe, *Rule-based Expert Systems: the MYCIN experiments of the Stanford Heuristic Programming Project*. Addison-Wesley, 1984.
[36] J. C. Giarratano and G. D. Riley, *Expert Systems: Principles and Programming, Fourth Edition*. Course Technology.
[37] R. Lindsay, B. G. Buchanan, and E. A. Feigenbaum, "DENDRAL: A Case Study of the First Expert System for Scientific Hypothesis Formation," *Artificial Intelligence*, vol. 61, no. 2, pp. 209-261, Jun. 1993.
[38] R. 0 Duda, J. G. Gaschnig, and P. E. Hart, "Model Design in the PROSPECTOR Consultant System for Mineral Exploration," in *Expert Systems in the Microelectronic Age*, D. Michie, Ed. Edinburgh, Scotland: Edinburgh University Press, 1979, pp. 153-167.
[39] W. A. Woods, "Progress in natural language understanding — An application to lunar geology," in *Proceedings of the national computer conference and exposition*, 1973.
[40] J. McDermott, "R1: A Rule-Based Configurer of Computer Systems," *Artificial lntell., 19: 39*, vol. 19, no. 1, pp. 39-88, Sep. 1982.
[41] W. J. Clancey, "Knowledge-Based Tutoring: The GUIDON program," in *MIT Press Series in Artificial Intelligence*, Cambridge, MA: The MIT Press, 1987.
[42] W. J. Clancey, R. Letsinger, and S. U. C. S. Dept, "NEOMYCIN: Reconfiguring a Rule-based Expert System for Application to Teaching," in *Readings in Medical Artificial Intelligence: The First Decade*, W. J. Clancey and E. H. Shortliffe, Eds. Reading, MA: Addison-Wesley, 1984.
[43] K. J. Healey, "Artificial Intelligence Research and Applications at the NASA Johnson Space Center," *AI Magazine*, vol. 7, no. 3, pp. 146-152, 1986.
[44] M. Dincbas, "A knowledge-based expert system for automatic analysis and synthesis in CAD," in *IFIP Congress*, 1980, pp. 705–710.
[45] G. Stiny and J. Gips, "Shape Grammars and the Generative Specification of Painting and Sculpture," *Information Processing*, vol. 71, no. 1460-1465,

1972.
[46] J. Lighthill, "Artificial Intelligence: A General Survey," in *Artificial Intelligence: a Paper Symposium*, B. S. R. Council, Ed. 1973.
[47] P. Russell, Stuart; Norvig, *Artificial Intelligence: A Modern Approach*. Englewood Cliffs, NJ: Prentice Hall, 1995.
[48] H.-J. Zimmermann, *Fuzzy sets, decision making, and expert systems*. Boston, MA: Kluwer Academic Publishers, 1987.
[49] G. Stiny, "Introduction to shape and shape grammars," *Environment and Planning B: Planning and Design*, vol. 7, no. 3, pp. 343-351, 1980.
[50] C. L. Forgy, "Rete: A fast algorithm for the many pattern/many object pattern match problem," *Artificial Intelligence*, vol. 19, no. 1, pp. 17-37, Sep. 1982.
[51] G. Riley, C. Culbert, R. T. Savely, and F. Lopez, "CLIPS: An expert system tool for delivery and training," in *3rd Conference Artificial Intelligence for Space Applications*, 1987, pp. 53-57.
[52] E. J. Friedman-Hill, "Jess, The Java Expert System Shell," Livermore, CA, 2000.
[53] G. L. Steele, *COMMON LISP: the language*. Digital Press, 1990, p. 1029.
[54] E. Friedman-Hill, *JESS in Action*. Manning, 2003.
[55] R. Studer, "Knowledge Engineering: Principles and Methods," *Data & Knowledge Engineering*, vol. 25, no. 1-2, pp. 161-197, Mar. 1998.
[56] G. Schreiber et al., *Knowledge Engineering and Management: The CommonKADS Methodology*. Cambridge, MA: MIT Press, 2000.
[57] A. Newell and H. A. Simon, *Human Problem Solving*. Englewood Cliffs, NJ: Prentice Hall, 1972.
[58] M. Suwa, A. C. Scott, and E. H. Shortliffe, "Completeness and Consistency in Rule-based Systems," in *Rule-based Expert Systems*, Reading, MA: Addison-Wesley, 1984.
[59] N. N. Das, D. Entekhabi, and E. G. Njoku, "An Algorithm for Merging SMAP Radiometer and Radar Data for High-Resolution Soil-Moisture Retrieval," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 99, pp. 1–9, 2011.
[60] NRC Committee on Earth Science and Applications from Space, "Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond," Washington DC, 2007.
[61] T. A. Sutherland, "Stakeholder Value Network Analysis for Space-Based Earth Observations," Massachusetts Institute of Technology, 2009.
[62] D. Entekhabi, "The Soil Moisture Active Passive (SMAP) Mission," in *Proceedings of the IEEE*, 2010, vol. 98, no. 5, pp. 704-716.
[63] M. Piles, D. Entekhabi, and A. Camps, "A change detection algorithm for retrieving high-resolution soil moisture from SMAP radar and radiometer observations," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 47, no. 12, pp. 4125–4131, Dec. 2009.
[64] P. E. O'Neill, E. Podest, and E. G. Njoku, "Utilization of ancillary data sets for SMAP algorithm development and product generation," in *Proceedings of the 2011 IEEE International Geoscience and Remote Sensing Symposium*, 2011, pp. 2436-2439.
[65] H. Barre, B. Duesmann, and Y. H. Kerr, "SMOS: The Mission and the System," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 3, pp. 587–593, Mar. 2008.
[66] L. A. Zadeh, "Fuzzy Sets," *Information and Control*, vol. 8, no. 3, pp. 338-353, Jan. 1965.

BIOGRAPHIES

Daniel Selva *is a PhD candidate in the department of Aeronautics and Astronautics at MIT. His research interests focus on the application of multidisciplinary optimization and artificial intelligence techniques to space systems engineering and architecture, with a strong emphasis on Earth observing missions. Prior to MIT, Daniel worked for four years in Kourou (French Guiana) as a member of the Ariane 5 Launch team. In particular, he worked as a specialist in operations concerning the guidance, navigation and control subsystem, and the avionics and ground systems. Daniel has a dual background in electrical engineering and aeronautical engineering, with MS degrees from Universitat Politecnica de Catalunya in Barcelona, Spain, and Supaero in Toulouse, France. He is a 2007 la Caixa fellow, and received the Nortel Networks prize for academic excellence in 2002.*

Dr. Ed Crawley *received an Sc.D. in Aerospace Structures from MIT in 1981. His early research interests centered on structural dynamics, aeroelasticity, and the development of actively controlled and intelligent structures. Recently, Dr. Crawley's research has focused on the domain of the architecture and design of complex systems. From 1996 to 2003 he served as the Department Head of Aeronautics and Astronautics at MIT, leading the strategic realignment of the department. Dr. Crawley is a Fellow of the AIAA and the Royal Aeronautical Society (UK), and is a member of three national academies of engineering. He is the author of numerous journal publications in the AIAA Journal, the ASME Journal, the Journal of Composite Materials, and Acta Astronautica. He received the NASA Public Service Medal. Recently, Prof Crawley was one of the ten members of the presidential committee led by Norman Augustine to study the future of human spaceflight in the US.*

APPENDIX

In this appendix we complement our paragraph on RBES with a succinct review of their history, a discussion of their structure, a short introduction to one of the most prevalent languages for the development of RBES (the CLIPS language), and a critique of RBES.

Short History of RBES— The first experiments with expert systems and RBES in particular were DENDRAL and MYCIN, both conducted at Stanford in the late sixties and early seventies in Buchanan and Feigenbaum's research group [35], [37]. The DENDRAL experiment was a project in organic chemistry. DENDRAL was capable of inferring the molecular structure of a substance from its mass spectra and other experimental data. The MYCIN experiment leveraged from the early experience of DENDRAL, and focused on the diagnostic of the type of bacteria causing an infection in a patient, and the subsequent prescription of the right antibiotic, using fuzzy rules, i.e. rules with certainty factors. With 450 rules, MYCIN was able to perform better than senior doctors, and considerably better than junior doctors.

After DENDRAL and MYCIN, the development of RBES spread around the world, to many disciplines of the sciences and engineering. Duda et al developed the PROSPECTOR system, which was capable of identifying ore deposits [38]. Woods et al developed the LUNAR systems, an RBES that answered geology questions about rock samples brought back from the Apollo missions [39]. The early success of RBES allowed for their commercialization, starting with the R1 (later called XCON) system developed by McDermott at CMU to assist in the configuration of DEC's VAX systems using 2500 rules [40]. Clancey adapted the MYCIN program for teaching and tutoring and created NEOMYCIN and GUIDON [41], [42]. Marsh and Healey developed several RBES for the NASA Johnson Space Center in the 1980's (NAVEX, RPMS, MCCSSES, MRDB) including applications to on-board navigation, electric power management, planning, requirements management, and monitoring radar data from the Space Shuttle in real time and estimate its position and velocity amongst others [43]. Dincbas developed an RBES to design digital circuits [44]. The field of architecture also embraced RBES as a means for automatic form synthesis using shape grammars, first developed at MIT by Stiny in the early seventies [45].

The success of early systems such as MYCIN and DENDRAL made some experts think that AI and RBES had no limits. Then, the promises of life-changing technology of the late sixties and early seventies were wiped out by Lighthill in his infamous paper known as the Lighthill report. The Lighthill report gave a very pessimistic view of AI in general and RBES in particular, which led to funding cuts in many research labs in the UK and around the world [46]. As a matter of fact, as Russell and Norvig point out that "to save embarrassment, a new field called IKBS (Intelligent Knowledge-Based Systems) was defined because Artificial Intelligence had been officially cancelled" [47]. Thus, expert systems became knowledge-based systems (KBS), and the term knowledge engineering was coined to designate the process of developing a KBS.

It wasn't until the late 80's and early 90's, that KBS and RBES in particular started being used in decision making, as a decision support tool [48]. The number and types of RBES has grown continuously during the last 40 years, and RBES are present today in almost all science and engineering disciplines, education, finance, and architecture [49].

We end this section by noting that the role of NASA, in particular of JSC, in the history of RBES is far from being negligible: DENDRAL, NAVEX, and LUNAR were both NASA-funded projects, and CLIPS was entirely developed at JSC.

Structure of RBES— RBES consist of three major elements: a fact database, a rule database, and an inference engine. The fact database contains relevant pieces of information about the specific problem at hand called facts. Information in facts is organized according to predetermined data structures similar to C structures and Java Beans, with properties and values (e.g. a fact of type car may contain a property make, a property model, and a property price, amongst others). These data structures are called templates in many RBES development tools. Facts can be asserted, modified, and retrieved from the database anytime. The rule database contains a set of logical rules that contain the domain knowledge. The LHS of these rules may match one or more facts in the working memory. The RHS of these rules define the actions to be executed for each of these matches, which typically include asserting new facts, modifying the matching facts, performing calculations, or showing some information to the user.

The inference engine performs three tasks in an infinite loop: a) pattern matching between the facts and the LHS of the rules in the working memory and creation of activation records (also known as the conflict set); b) while there remain activation records, select the next rule to execute, or fire (conflict resolution); c) execute the selected rules' RHS (deduction). Most current rule engines are based on the Rete algorithm, developed by Forgy in 1982 [50]. The Rete algorithm is faster than other algorithms because it "remembers" prior activation records in a network in memory called the Rete network. The Rete network is very efficient in speeding up the search process because most of the time, the network does not change much between iterations. Note that the improvement in computational time comes at the price of increased use of memory.

The CLIPS and Jess languages— CLIPS (C-language integration production system) is a public language to write expert systems developed in 1985 at NASA Johnson Space Center as an alternative to the proprietary ART*inference [51]. Ten years later in 1995, Dr Friedman-Hill at Sandia National Labs developed an expert system shell in Java based on CLIPS, especially tailored for RBES [52]. As any other RBES, Jess deals with rules and facts, and its inference engine is based on the Rete algorithm. CLIPS/Jess syntax is very similar to common LISP, one of the earliest programming languages in artificial intelligence [53]. In CLIPS/Jess, the properties in templates are defined using the deftemplate command and properties are listed using the keywords slot (single element attributes) or multislot (multiple element attributes). Rules are defined using the defrule command. A fact is added into working memory, modified, or removed from working memory using the assert, modify, and retract commands respectively. Whenever the working memory is modified (e.g. a new fact is added or modified), the Rete network is recalculated. Matching rules are fired in the order determined by the Rete algorithm once the run command is sent. The general structure of the definitions of a template, a fact, and a rule are provided in Figure 13. For further information on the Jess language, a good overview can be found in [54].

```
(deftemplate my-template
   "Description of the template"
   (slot slot1-name) (slot slot2-name) …
(multislot multislot1-name) (multislot multislot2-name) …
)

(assert my-template (slot1-name slot1-value) (slot2-name
slot2-value) …
   (multislot1-name ms1-value1 ms1-value2 …)
   (multislot2-name ms2-value1 ms2-value2 …) …
   )

(defrule my-rule
   "Description of the rule"
   ;; LHS = conditions
   ?this_fact ← (my-template (slot1-name ?x) (slot2-name
?y) …)
    ;; other conditional elements using variables ?x and ?y
   →
   ;; RHS = actions, e.g. see below
    (bind ?newx (+ ?x 1))
    (modify ?this_fact (slot1-name ?newx)
   )
```

Figure 13: Definition of a template, assertion of a fact, and defintion of a rule in the CLIPS and Jess languages

Our RBES was developed in Jess rather than in other more common progamming languages for artificial intelligence such as LISP or PROLOG. The main reason for that is that RBES development tools such as CLIPS and Jess allow the user to focus on the application and forget about having to “reinvent the wheel” since much code that is needed to create an RBES is already available. Furthermore, Jess is written in Java, which facilitates the integration with the Matlab environment, very popular amongst both scientists and engineers.

Critique of RBES—Reasons for criticism toward RBES can be grouped in four categories: a) problems in knowledge elicitation; b) problems in knowledge representation; c) problems in computational efficiency; d) problems in overall performance.

Each of these groups is succinctly developed in the next paragraphs. During the development of the early MYCIN and DENDRAL systems, it became obvious that a bottleneck in the development of RBES was the process of eliciting the knowledge from experts. One of the major challenges is the large quantity of tacit knowledge, which the expert is not aware that he or she has. Hence, the primitive view of KBS as a knowledge transfer process was challenged. Today, there is wide agreement that the process of building a KBS is a modeling process rather than a knowledge transfer process [55]. A review of modern knowledge engineering techniques can be found in [56]. While much progress has been made since then, the problem of eliciting expert knowledge is still a major challenge for the RBES developer. In terms of knowledge representation, Newell and Simon postulated in their classic text “Human Problem Solving” that most human expert knowledge can be expressed in the form of logical rules [57]. Earth science is not an exception to this. However, some knowledge is more naturally expressed using other more generic types of data structures such as frames (see for example Minsky’s work [34]). Related to this problem is the difficulty to guarantee the completeness and consistency of the rule base [58]. In terms of computational efficiency, we mentioned before that the Rete algorithm is the most efficient inference algorithm for rule engines, and it generally improves the naïve approach of comparing all facts with all rules in working memory, which yields a time complexity of O(RFP), R being the number of rules, F the number of facts, and P the average number of clauses in the LHS of the rules. Indeed, the Rete algorithm achieves a time complexity of O(R’F’P’), where R’, F’, and P’ are less than or equal to R, F, and P [50]. However, in the worst case, this still an exponential problem, which can lead to combinatorial explosion very fast in real-time applications.

Finally, in terms of overall performance, RBES have been successful when applied to relatively narrow domains, but have often failed in much larger applications [55]. Also, expert systems in general are sometimes called weak methods in the AI community because they use “weak” (uncertain) information about the domain. Paraphrasing Russell and Norvig, “one might say that to solve a hard problem, you almost have to know the answer already” [47].

Our choice of RBES over other KBS architectures for this work, despite all the aforementioned limitations, was driven by superior characteristics in terms of scalability thanks to the separation between the knowledge (the rules) and the flow control (the inference engine), as well as their transparency achieved through their built-in explanation facilities, two attributes that we consider of outmost importance, and often overlooked by DST developers in system architecting.