TRADESPACE EXPLORATION APPROACH FOR ARCHITECTURAL DEFINITION OF IN-SPACE TRANSPORTATION INFRASTRUCTURE SYSTEMS FOR FUTURE HUMAN SPACE EXPLORATION

Alexander Rudat
Massachusetts Institute of Technology, Aero/Astro Department, USA, arudat@mit.edu

Jonathan Battat[1], Alessandro Aliakbargolkar[2], Morgan Dwyer[3], Dr. Bruce Cameron[4], Prof. Edward Crawley[5]

Abstract

Current planning activities for human spaceflight development are centered on exploration beyond Earth orbit, focusing primarily on the destinations of the Moon, Near Earth Asteroids (NEAs), and Mars. In order to more effectively define future programs, key system level features impacting mission infrastructure performance and cost must be identified and assessed prior to concept down-selection. This paper presents a framework in which the in-space portion of beyond Earth orbit human exploration infrastructures is described by a set of function mappings to habitation and transportation hardware along with specification of advanced technology alternatives. Using system architecture techniques, a comprehensive set of possible mission design alternatives is enumerated and explored via assessment of cost proxy metrics. The scope of the tradespace includes multiple mission destinations (Moon, Mars, and NEAs), high-level mission mode decisions, and the selection of specific performance-improving technology options. The system modeling approach relies on the identification of a set of fundamental and invariant functions present in all exploration architectures. Use of the fundamental functions is unique to this model and allows for a broader exploration of design alternatives than previous studies. Presented in this paper are the formulation of the model, focusing on the scoping of the model, the decomposition of the system, and determination of the metrics of interest; validation against heritage systems and vetted design reference architectures; results generation, using both heuristics and full enumeration techniques; and a brief set of model outputs, focused on Mars and Moon missions and their top architecture formulations.

[1] Massachusetts Institute of Technology, Aero/Astro Department, USA, jabattat@mit.edu
[2] Massachusetts Institute of Technology, Aero/Astro Department, USA, golkar@mit.edu - Now Assistant Professor, Skolkovo Institute of Science and Technology (Russian Federation)
[3] Massachusetts Institute of Technology, Engineering Systems Division, USA, mdwyer@mit.edu
[4] Massachusetts Institute of Technology, Aero/Astro Department, USA, bcameron@alum.mit.edu
[5] Massachusetts Institute of Technology, Aero/Astro Department, USA, crawley@mit.edu

I. INTRODUCTION

With the end of the Space Shuttle era and the creation of a new Space Launch System (SLS), NASA and its global counterparts have an opportunity to renew studies of manned missions beyond the scope of the Earth-Moon system. NASA teams such as the Human Spaceflight Architecture Team (HAT) have been tasked with developing paper studies of these missions. International groups such as the International Space Exploration Coordination Group (ISECG) have similarly begun to collaborate on international studies and partnerships toward this same end [1]. Previous studies, both domestic and international, such as the Jet Propulsion Lab's (JPL) Austere Mission [2] and the Australian Mars-Oz project [3], have gone through many iterations for determining the "best" method for manned exploration. Extensive studies are necessary, as these missions would require many years for hardware and technology development as well as many billions of dollars both for development and operations. With the development of the SLS, the infrastructure for the launch and operation of complex space systems is being established. As exploration systems become closer to physical development, the understanding of the drivers of performance and cost for such systems becomes critical.

In order to facilitate high-fidelity point design studies and inform future decision-makers, the key drivers of system-level needs should be identified and quantified prior to concept down-selection. Systems architecting techniques allow for the analysis of the system tradespace while retaining the necessary system complexity to differentiate between good alternatives. These methods allow analysis of the broader perspective while retaining enough fidelity to distinguish between architectures. Unlike traditional point designs, systems architecture analysis allows engineers to more rigorously quantify architecture-level trades and assess the impact of design decisions within the space of possible architectures. Recent research conducted at M.I.T. has developed a set of tools for quantitative analysis of systems architectures, such as the Architectural Decision Graph (ADG) Framework [4] and the Algebra of Systems [5]. Building on these concepts, quantitative analysis of human exploration infrastructures was performed in order to inform future high-fidelity point designs.

We present a unique system model for the analysis of the in-space infrastructure for manned exploration missions. The modeling approach focuses on three components. The first component is comprised of a group of mission mode parameters that set the scientific value of the various exploration infrastructures. Second, the physical habitat and propulsion elements are established by allocating a set of fundamental invariant functions to hardware components. Lastly, a suite of technology options accompanies these system parameters, allowing for the analysis of technology impacts on the system. Although system architecture techniques have been applied to this type of system in the past, they have been limited by assumptions about the habitation and transportation tradespace. This also differs from general model-based engineering by exploring more alternatives in a lower-fidelity space. Model-based engineering is used internally in order to analyze architectures, but the focus is on high-level, architecturally distinguishing features rather than modeling of the complete system in detail. This also requires that the models are formulated for efficient analysis of many architectures, rather than efficient analysis of many system details using extensive domain knowledge. Effective exploration of this space is accomplished by utilizing vetted techniques on a unique decomposition of the manned exploration system.

The scoping of this study in the broader context is first addressed, followed by a determination of the architecturally distinguishing features present in manned exploration infrastructures. The current model's development from previous models is also discussed, along with the past and present capabilities and limitations. Validation cases are presented, in detail for Mars and briefly for the Moon and Near Earth Asteroids (NEAs). Methods for the generation of results are discussed, followed by a small set of key results.

Specifically, the paper is structured as follows. Section II will discuss the study formulation, including scoping and problem formulation. Section III will describe the validation of the model and key differences between the model outputs and particular point designs used for validation. Section IV addresses the generation of results, and section V presents key results. A conclusion is given in section VI, along with acknowledgements in section VII and references in section VIII.

II. STUDY FORMULATION

Beyond the Point Design

Traditional analysis of complex space systems, particularly manned exploration systems, has been accomplished through the use of highly detailed point design studies. These studies have been conducted by teams of highly skilled and experienced individuals, often requiring months or years to accomplish. Although a great deal of unpublished background research is performed prior to these design studies, without detailed analysis of many possible systems, these types of analysis may fail to capture the best possible system-level design. The interactions between the components of these complex systems result in emergent properties that are not always readily predictable. This means that some systems with particular combinations of elements may have unexpected properties or behaviors not readily predictable prior to modeling. Therefore, without modeling the variety of system element combinations, engineers cannot be certain that the chosen set of elements for detailed point designs have the best performance characteristics. This is especially prevalent in systems without significant heritage and/or highly complex interactions between system elements. Both of these characteristics are present in large-scale human exploration missions.

To circumvent this problem and feed information to engineers about elements of point designs to be considered, architecture-level systems models are integrated into the early-stage design process. These models decompose the system of interest into architecturally distinguishing features, allowing analysis of the various combinations of these features without the need for detailed models. Due to the reduced fidelity of architecture-level models, the results are not taken to be exact, but the features are selected such that the impact of changing one or more feature results in a correct ordinal ranking along the metric or metrics of interest. In other words, the ranking of architectures resulting from the analysis remains correct regardless of the lack of detail in the model, if the simplifying assumptions posed by the model are valid. Feeding these top choices from a lower fidelity study to one of higher fidelity builds confidence that the point design reflects the best architecture in the broader tradespace.

Scoping

Downscoping of the Tradespace

As a description of the full tradespace of human exploration systems infrastructures is far too broad, downscoping of the problem at hand is necessary in order to create a comprehensible set of analysis. For this model, the overall tradespace was downscoped by decoupling the system from parts of the mission infrastructure. The final designs for the future U.S. launch system are being established, and so an analysis of the space including a launch system is both unnecessary and overly complex. Combined with the knowledge that many case studies focus on surface operations, where the greatest design uncertainties are present, along with the fact that science payload to the surface and surface operations set the scientific value of a mission, we determined that it was most appropriate to model the in-space transportation infrastructure of the system. This allows for a deeper understanding of a portion of the system that drives both material and development costs as well as launch capability needs.

Destination Selection

NASA and its international counterparts have debated the set of destinations appropriate for human exploration missions for many years, culminating in deliberations revolving around the concept of the "Flexible Path" [1][6] within the last ten years. The three destinations described in such a path, regardless of the order in which they are explored or not explored, are generally the Moon, Mars, and Near Earth Asteroids (NEAs). In theory, these destinations allow for incremental technology development, building to systems for more and more difficult destinations to reach. NASA HAT has taken it upon itself to explore detailed point designs for missions to these destinations. The model reflects these destination choices.

As there are many asteroids with a wide range of characteristics that may be classified as NEAs, two asteroids that represent the widest breadth of possibilities, one "High Energy NEA" and one "Low Energy NEA", the difference being the energy needed to reach the destination from Earth from as astrodynamics point of view, were included. Although these in theory represent a set of possible destinations, the specific numerical values used to model these generic destinations come from specific asteroids also being studied by NASA HAT, namely 2000SG344 and 2008EV5 [7].

The “Flexible Path” often also includes other “non-solid” destinations, usually Earth-Moon or Earth-Sun Lagrange points. As the scientific value of these destinations is still hotly debated, as well as the fact that these are used almost exclusively as mid-points in larger, multi-mission exploration schemes, these destinations were excluded from the tradespace as final destinations. However, they were included as possible rendezvous points, meaning that exploration elements could be pre-deployed to these locations and “picked up” on the way to a further final destination.

In summary, this study includes the evaluation of round-trip missions to the Moon, Mars, and a set of two representative NEAs. Each of these missions may also use one of several rendezvous points, including the most favorable Earth-Moon and Earth-Sun Lagrange points [8].

Further Assumptions

For the purposes of limiting the tradespace and reducing computational complexity, further assumptions regarding the nature of the systems in question were necessary. These include both broader scoping assumptions as well as restrictive technical assumptions. Broad scoping assumptions reduce the overall design space and complexity by simplifying the system being modeled. Technical assumptions reduce the complexity of interactions between elements or simply reduce the number of possibilities for technology implementation.

For computational simplicity and in line with previous architectural studies, the analysis was limited to sortie-like missions only. This means that each mission is designed to deliver crew and payload to the surface of a rocky body, allow a stay on that surface for a given amount of time, and return the crew to Earth. This eliminates two other possible modes. The broader campaign of missions, where multiple crews explore a given destination or set of destinations, is not considered. Elements within an architecture are assumed to be designed for the single mission only, and no effort is given to leaving any system elements for future missions at any point. All elements are discarded after their functional requirements are fulfilled. The second possible mission mode this also eliminates is the highly controversial one-way mission, where astronauts are sent on a mission to a destination but are explicitly not returned.

Technical assumptions limited the elements of the various architectures to those that have already been developed or have development history. Technical elements were also limited to those that have sufficient data to be modeled accurately. System elements with absolutely no development history or insufficient available data were not considered. For example, LOX/LCH_4 propulsion elements were included in the system while solar sails were not. This is due to the testing of large-scale methane engines, such as the XCOR XR-5M15, while solar sails are still conceptual [9].

Evolution of the Architecture Formulation

System Architecting Problems (SAP) and the Assignment Problem

In Dr. Selva’s Ph.D. work [10], a set of System Architecting Problems is proposed. These cover the majority of different formulations for the modeling of complex systems at low fidelity. Prior modeling of manned exploration missions at the architectural level have focused on the “Generalized Assignment Problem”. This is accomplished by breaking down a system into a series of decisions, each decision corresponding to a particular architecturally distinguishing feature. The choices available for each decision represent the different possibilities for the distinguishing features. For example, a decision for a Mars architecture may be “How will the crewed stack circularize the orbit around Mars?” Possible choices might include: 1) propulsive breaking, 2) aerocapture, or 3) a combination of propulsive breaking and aerocapture. Each decision may have one and only one choice selected, and these decisions and choices aggregate to define an architecture. The general assignment problem has been mathematically shown to be in the category of non-polynomial-complete (NP-complete) problems [11]. Therefore, in theory, any formulation of an architecture modeled in any NP-hard method can be formulated as an assignment problem [12]. This allows for generality in properties across many architecture-level formulations. Furthermore, the use of decisions and choices matches well with the needs of real-world decision-makers. Decision-makers must choose among the various good options in much the same way as an architecture is formulated using the assignment problem. Not only does this match well from a modeling standpoint, but this also increases acceptance of the model formulation and results by the recipients of the data.

Past Models, Their Capabilities, and Their Limitations

Hofstetter formulated manned exploration systems, specifically lunar missions, in the arrangement of an assignment problem in his master's thesis [13]. His approach focused on the "events" occurring at five nodes in the system, being most concerned with the movement of the crew from element to element. In total, Hofstetter used seven decisions to define an architecture, allowing for the full enumeration and analysis of approximately 30 architectures. Not all combinations of decisions and choices were possible, as some tightly coupled decisions were incompatible in certain combinations. For example, the choice of three crew transfers without choosing a set of three transfers in the five nodes is both illogical and impossible. These additional constraints, which required hard encoding of logical impossibilities, were acceptable due to the small number of possible architectural combinations with the set of seven decisions. However, this limited formulation prevented the analysis of anything beyond basic groupings of habitat elements. This system formulation was adequate to evaluate a set of basic habitat groupings, but any architectural features beyond that cannot be modeled by the system.

Extending that concept, Simmons developed the Architectural Decision Graph (ADG) Framework [4]. This framework allows the architect to formulate complex systems in the form of an assignment problem, allowing both the modeling of the system architecture and the evaluation of that system in the same structure. Simmons went on to demonstrate the use of this framework by formulating the Apollo Program. Once again only key groupings of habitats and propulsion stages are considered, although this reflects the discussion during the Apollo era more closely than a full set of all possible habitat and propulsion element combinations. Simmons went a step further than Hofstetter by including propellant technology options for the major stages as well as options for the number of crew. Simmons' formulation included nine distinct decisions, most pertaining to the formulation of the propulsion and habitat elements. This formulation encoded more of the constraints by cleaving the decisions along more natural lines, meaning that multiple decisions did not clearly pertain to the same logical constraint. This was not fully generalized, however, and thus still required additional post-formulation constraints.

These studies lack several major aspects that allow for the broader exploration of the manned mission tradespace. First and foremost, they each make the assumption that only a small set of possible transportation and habitation strategies can be used. The Apollo Program has been said to have used over 1,000,000 man-hours in the decision to employ the Lunar-orbit rendezvous strategy [14]. While this may have worked well for Apollo, this strategy requires resources that many missions will not have available. As such, it is important to have a consistent way to analyze the many different transportation strategies possible for these extended systems without huge resource expenditure. Habitation strategies, meaning the choice of how many habitats should accomplish the various necessary functions and which functions belong to which habitat, require a similar exploration across the design tradespce.

Within these design tradespaces, an important concept that has not gotten sufficient treatment in the past is the concept of re-use. More specifically, this is the re-use of a system element for a second or greater function as temporally displaced from the first function but still within the same mission. An example would be a propulsive stage that is used for Earth departure, sits in Mars orbit, and then is re-used for the return to Earth. The re-use requirements of all physical elements are critical to the design of the system. A method for formulating all propulsive stage and habitat element allocations allows for any kind of re-use to be present in the modeled system. Therefore, by not limiting to a small set of assumed allocations, full re-use analysis can be performed.

System Model Formulation: Capabilities and Limitations

Results from an architecture level study are highly dependent on the method of decomposition of the system of interest. As such, it is critical to determine *a priori* the key drivers of the system level metrics of interest. The drivers for the metrics of interest determine the architecturally distinguishing features of a system, therefore driving the method of decomposition. This section describes the decomposition approach, key architecture level parameters, and the metrics upon which each architecture was evaluated.

How Metrics Drive Formulation

How does one determine the architecturally distinguishing features of a system? The answer is highly dependent on what metrics of evaluation the modeler will use to determine the quality of the architectures. If performance or a proxy for performance is the key metric, system elements that heavily impact performance should be considered first when determining the distinguishing features. These may differ greatly from a model focusing on risk as the primary metric. Therefore it is critical to formulate the metrics prior to the development of the model.

For this model, overall lifecycle cost was chosen as the primary metric. Risk and architecture extensibility metrics were also considered but found to be either too coarse or overly complex. Cost was not directly calculated, as it has been repeatedly shown that cost estimates are unreliable at best [15]. Therefore proxy metrics embodying the principal drivers of cost were selected.

Initial Mass in Low-Earth Orbit (IMLEO)

Mass drives cost. This has been well established in the aerospace community over the past decades [16]. Mass on the destination surface drives the mass of the EDL system, which drives the mass of the in-space system, which drives the launch cost from Earth's surface. Although mass clearly does not encompass all of the aspects of system lifecycle cost, it is a primary component of it. An IMLEO metric perpetuates the concept of decoupling from the launch infrastructure while measuring an established cost driver. IMLEO is calculated using the basic formulation in Equation 1.

$$IMLEO = m_{pl} + m_{crew} + m_{log} + m_{prop} + m_{hab} \quad [1]$$

where m_{pl} is the payload mass

m_{crew} is the crew mass

m_{log} is the logistics mass

m_{prop} is the propulsion stages' mass

m_{hab} is the total mass of all habitat elements

Number of Development Projects

Launch costs constitute only a portion of the total lifecycle cost. Another primary factor is the development cost of new large-scale systems and technologies. The Number of Development Projects metric is a rough integer metric designed to reflect the number of these new projects that require significant resources many years prior to the operation of the system. This metric was first introduced in a study of the heavy lift launch vehicles [17]. In this case, the total number of distinct habitat elements, the total number of distinct transportation elements, and a weighted sum of the technology options, as influenced by readiness level, constitute the number of development projects for a given architecture. Although this clearly does not account for all system development cost due to complexity, as evidenced by the fact that a habitat that must perform all required function still only counts as a single development project, it is the belief that this is adequate to distinguish among the architectures in an ordinal ranking.

Functional Decomposition Approach

The most natural approach to decompose a technical system is to cleave upon obvious physical system elements. This is most directly reflected in traditional systems engineering, where design teams are broken into the primary subsystems of a larger, more complex technical system. In the case of interplanetary human exploration systems, there is no real system heritage and only limited design reference heritage. This drives a more generalized functional decomposition, looking at what fundamental functions the system must perform. Typically, these functions are then matched with possible combinations of physical system elements that perform these various functions, referred to as functional allocations, ranging from multiple elements performing the same function (redundancy) to single elements performing multiple functions.

After many cycles of analysis on the decomposition of the system, we found that transportation systems of this level only have two fundamental functions that must be performed: 1) the "habitation" function, which provides the life support as well as living space for the astronauts, and 2) the "transportation" function, which is the physical movement of the people and cargo to the destination. However, these fundamental functions are not the only system aspects that drive the key metrics. That is to say that the metrics may be influenced by system properties that are independent of how the fundamental functions are allocated to physical elements.

Architectural Variables, Functional Invariance, and the Set Partitioning Problem

As described, the in-space transportation infrastructure performs a set of two primary functions. Both of these functions must always be performed, otherwise the system cannot successfully operate. This concept of functional invariance is critical to the successful modeling of this system. As all architectures are meant to accomplish the same goals with the same scientific value, the functions performed within the system must likewise be constant across all architectures. The in-space portion of the transportation infrastructure must fundamentally accomplish the functions of transporting the astronauts and providing habitation for those astronauts. Although this seems like a trivial statement, these fundamental functions are difficult to establish. For one, engineers tend to focus on physical aspects of a system rather than functional aspects, making disassociation from physical forms non-intuitive. Such complex systems are also perceived as being exactly that: complex. Driving down into the fundamental nature of these complex systems means working beyond the daunting complexity.

These primary functions are further broken down into sets of seven and ten functions, respectively, which must also always be performed. The seven "habitation" functions include: 1) Earth launch (which may be accomplished by an element in the in-space transportation infrastructure), 2) deep space outbound habitation, 3) descent habitation, 4) surface habitation, 5) ascent habitation, 6) deep space inbound habitation, and 7) Earth re-entry. It is possible to imagine that these may be further broken into a continuum of infinite possible functions, but these segments comprise the set of logical functional subdivisions of the space. This also coincides with the common physical separations, as each function requires different performance characteristics from physical systems. This reinforces the concept that these cleavage points are correct, as they correspond to established systems and models. Similarly, the ten "transportation" functions include: 1) Earth departure, 2) outbound staging location breaking, 3) outbound staging location departure, 4) destination orbit arrival breaking, 5) descent, 6) ascent, 7) destination orbit departure, 8) inbound staging location breaking, 9) inbound staging location departure, and 10) Earth orbit breaking. No Earth surface launch segment is included due to the de-coupling from the launch infrastructure. No Earth entry transportation leg is included as it is assumed that aerobreaking in the atmosphere is always used. This last assumption indicates that any habitat element used for Earth entry must have an associated aeroshield. This leads to a trade between the desire for a single habitat to reduce additional dry mass and the need for a large aeroshield. Unlike the habitation functions that truly must always exist, it is clear that without staging locations for a given set of architectures, there is no need for transportation functions 2, 3, 8, and 9. These "points" must always be passed, in a sense, within the mission, but the cost of going through them becomes zero when there is no need to stop at any staging location. A pictorial representation of these functions in the context of a Mars mission, which is the most comprehensive mission mode, can be seen in Figure 1. The physical elements shown on this chart are not meant to restrict the possible forms of the various habitat elements. Rather, they are meant to reinforce the concept of the habitation functions. This method of segmenting the space is reinforced by the fact that these elements of form are normally conceptualized in this fashion.

Each of the two primary functions is composed of a set of sub-functions that must always be present and must always be accomplished once. This clearly fits the criteria for a set partitioning problem [18], which, like the assignment problem, is one of the many NP-complete problems [19]. As the formulation and calculation of these problems is well known and well-studied, it is advantageous to encode this information as two set partitioning problems. It should be noted that this causes a distinct separation from the traditional assignment problem, as the model is now composed of a set of decisions and choices and two assignments to one of many set partitions schemes. After eliminating logical inconsistencies, each of these problems has a selection group of 120 and 776 options, respectively. Unlike many other problems formulated under an assignment problem decomposition, the use of set partitions with many alternative schemes creates a challenge to the engineer as well as the various outside stakeholders regarding the intelligibility of the system. However, this is more intelligible than the traditional assignment problem formulation, which would require a set of 10!+7! binary decisions (yes/no) regarding the matching of functions to formal elements in the worst case.

Along with these two set partitioning problems, a third concept of pre-deployment creates the group of so-called architectural features for a given architecture. This name is given because these three decisions determine the majority of the large physical elements of the system, thus setting much of the hardware development cost as well as the majority of the system extensibility capabilities. In this case, pre-deployment describes the use of a low thrust system to pre-deploy cargo to the destination surface or orbit. This decision has been simplified to a binary switch (yes/no). "Yes" means that every element that can be logically pre-deployed to the destination surface or orbit will be, such as propulsion stages used after arrival at the destination orbit by the crewed stage that are not required prior to that stage. For instance, if a system consisted of three propulsion stages, being an Earth departure stage (EDS), a descent and ascent stage, and an Earth return stage, the second and third stages could be pre-deployed while the EDS could not. However, if the system was only two propulsion stages, where the EDS and descent and ascent stage were combined, only the Earth return stage could be pre-deployed.

In this system decomposition, any element pre-deployed from LEO would use a solar electric propulsion (SEP) stage, which is the low thrust system of interest to NASA [7]. Traditional cryogenic and nuclear propulsion stages that are typically used to pre-deploy elements in many of the reference designs were not considered for the pre-deployed elements. This is due to the fact that the rocket equation is linear with payload [20]. This means that, from a mass standpoint, having all the mass on one stack versus the mass being on two different stacks makes no difference with regard to propellant mass. This only works under the assumption that propulsion elements have a constant dry mass ratio. Although reality differs slightly from this case, it is not worth the additional computational complexity to model this relatively minor variation. Given that the same traditional propellant is used for both the crewed stack and the pre-deployment stack, in the metrics of interest, the architectures having pre-deployment or not having pre-deployment are identical. Low-thrust propellants, on the other hand, have very different energetic requirements and thus make a significant difference in the metrics of interest. The use of such a system also significantly impacts the development of system elements, and so

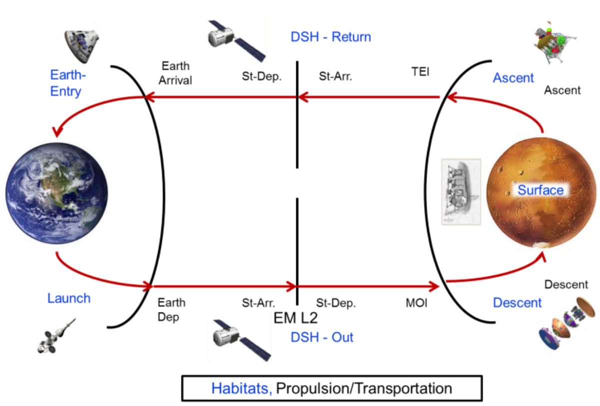

Figure 1: Habitat and Transportation Functions in a Mars Mission Context

it is grouped with the set partitioning problems as a member of the architectural decisions.

Advanced Technology Options

The third and final set of system decisions that define a given architecture are the advanced technology options. These fall into two groups: 1) decisions between technology options where one must always be chosen, and 2) decisions about whether or not to include certain advanced technologies. The propellant choices fall into the first category, and it was decided based on the current interests of NASA to include LOX/LH_2 and LOX/CH_4 as options for all stages, as well as nuclear thermal rockets (NTR) for the in-space stages and hypergolic propellants for the descent and ascent stages. Hypergolic propellants are too low performance to be reasonable for the in-space stages, while NTR options are unreasonable for descent and ascent due to the large dry mass necessary as well as the continuous thrust characteristics of these systems.

The set of binary decisions included in the study cover several technologies that have large impacts at an architectural level. Aerocapture for Mars missions allows the trade between a large propulsion stage for breaking into Mars orbit and a large dry mass for using atmospheric drag. *In-situ* resource utilization (ISRU) similarly trades the need for carrying large amounts of ascent propellants from Earth for the dry mass necessary to extract them from the destination surface and possibly atmosphere. Boil-off control, which limits the amount of propellant burn-off due to heating and containment issues, trades larger amounts of propellant for extra dry mass. In this case, the various boil-off control capabilities are not individually analyzed. Rather, an assumption about the reasonable capabilities for future systems is made, which matches with the assumptions made by the NASA HAT group [21].

Iso-performance Analysis

In addition, each architecture also includes a set of mission mode parameters. These set the scientific performance level of the mission architecture, characterized by five mission aspects. These include 1) the destination choice, 2) the number of days on the surface, 3) the number of crew in the mission, 4) the possible outbound staging location (rendezvous point), and 5) the possible inbound staging location, which may differ from the outbound location. Although some science value can be altered by operations in the in-space portion, the vast majority

Mission Mode Parameters										
Destination	**Mars**	**Moon**	NEA (High Energy)	NEA (Low Energy)						
Number of Crew	3	**4**	5	6						(any user input)
Surface Mission Duration	**7**	14	21	30	90	180	360	**500**		(any user input)
Possible Outbound Staging Locations	EML1	**EML2**	SEL1	SEL2	Phobos	**None**				
Possible Inbound Staging Locations	EML1	**EML2**	SEL1	SEL2	Phobos	**None**				
Enumerated Features										
Architectural Features										
Habitation Distribution	120 Combinations									
Propulsion Element Distribution	776 Combinations									
Pre-Deployment (via SEP)	Yes	No								
Technologies										
Boil-off control for propellants	Yes	No								
ISRU	Yes	No								
Aerocapture	Yes	No								
TDI Propellant Type	LOX/LH2	LOX/LCH4	NTR							
Descent Propellant	LOX/LH2	LOX/LCH4	N2O4/MMH							
Ascent Propellant	LOX/LH2	LOX/LCH4	N2O4/MMH							
TEI Propellant	LOX/LH2	LOX/LCH4	NTR							
(Pre-Deployment Propellant)	SEP									

Figure 2: Architecture Morphological Matrix, with mission mode parameters presented in the results in bold

of scientific value is set by the combination of surface operations and these mission mode parameters. Therefore it is adequate to assume that all architectures under these parameters have equal science value. For the analysis presented here, each set of architectures for a given destination has the same set of these variables.

Altogether, an architecture is defined by a set of 16 decisions, including five science value decisions, three architectural decisions, and eight technology options. These are most easily viewed in the form of a morphological matrix, seen in Figure 2.

Validation

The evaluation of architectures for each of the destinations was validated against the most relevant design case. Two clear methods for validation of such a model exist. The first method, used in this study, is to model an reference mission and check the output on the metrics of interest, determining the match with the detailed point designs. The second method would be to analyze the full design space and determine the placement of the point designs in the overall space. This method is difficult to apply in this situation, as it requires knowledge of how the point design fits in the overall space in order to verify its placement in the model. As full analysis of the overall space has not been rigorously conducted to a sufficient level of fidelity, this method would not adequately validate the model.

Mars Case – Design Reference Architecture (DRA) 5.0

As there is no flight history for manned Mars missions, the most detailed and vetted design study was taken to be the baseline for comparison in the validation of the Mars case. Design Reference Architecture 5.0 is NASA's latest Mars reference mission, published in 2009 [22]. Once the reference mission was encoded in the formulation required for the established model, it was found that there were fundamental assumption differences between DRA 5.0 and the model baseline. When unaccounted for, these assumption differences caused a significant shift in the model results away from those established by the design study. However, when adjusted in the model to match that of DRA 5.0, it can be seen that the mass results from the model match well with those of the design reference architecture. The assumption differences found to have the greatest impact include the ΔV requirements, propellant boil-off rate, consumable usage rate, and the size of the deep space habitat. These were all dependent on input assumptions to the model. Had the model been formulated in the same fashion as DRA 5.0, there would not have been a discrepancy between the model output and the design reference mission.

In the detailed reference study, NASA assumed that the mission would launch on a highly favorable launch date, requiring significantly less ΔV capability from the system. The model, on the other hand, assumes a more average energetic requirement, although it still assumes that the mission will be launched within the favorable portion of the Mars launch windows [23]. DRA 5.0 also assumes that the propulsion stages will have zero boil-off capability, meaning that no propellant will be lost in the system during the long in-space segments. NASA's more recent HAT studies as well as outside group studies [24] have shown that this is in fact not a reasonable assumption, and that boil-off rates should be expected to be closer to those reflected in the model. NASA JPL produces a consumable intake estimation tool, and this tool is integrated into the framework of the model calculations. However, DRA 5.0 makes the assumption that consumables will be used at a lower rate. As these consumables are present during the entirety of the mission (in some portion, obviously in relation to how they are consumed), the ripple effect of even a relatively small change in consumable mass creates a much bigger impact on the overall system IMLEO. The final difference to be accounted in the validation study was the sizing of the deep space habitat. The most recent sizing estimates coming out of the global space community were used in the model [25]. DRA 5.0, on the other hand, used ground-up creation of custom habitats. The model incorporated a parametric relationship to establish the total habitable volume required based on the duration of the mission and number of crew present [26], as well as the type of functions that each given habitat must perform. This is then matched with historical data regarding the mass of rigid habitats on a mass per volume basis. This establishes the overall mass of the habitat systems. The design reference mission, on the other hand, was able to design the habitats to a component level. It should be noted that this study makes the assumption that all habitats will be rigid,

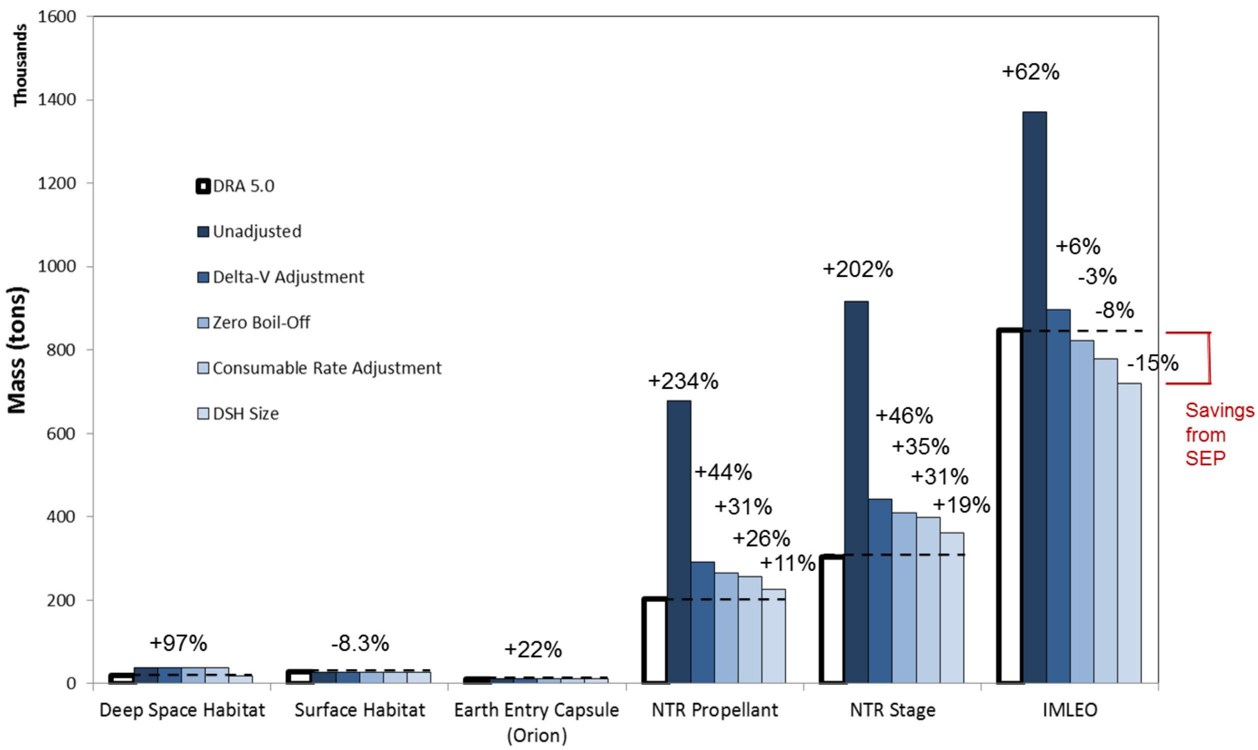

Figure 3: DRA 5.0 Validation

as inflatable habitats are in many respects still in their infancy, and no reliable parametrics for sizing of inflatable habitats can be established without significantly more flight history.

The aggregate information on the validation of the model versus DRA 5.0 can be found in Figure 3. This figure shows a series of bars for various system elements that drive IMLEO. The left-most white bar represents the reference mission. The blue bars represent the progression of assumption adjustments from no adjustment to the final validation case. The overall IMLEO difference of 15% is believed to be associated with the savings from pre-deployment with SEP instead of NTR as present in DRA 5.0. This, even without this knowledge, is within reasonable error bounds for both the study and DRA 5.0.

Moon and NEAs

The validation for the Moon case was performed against the obvious candidate of the Apollo missions. Once encoded, the original validation without any adjustment came within 16% of the overall IMLEO. However, it was again found that there were fundamental assumption differences between the model and the actual Apollo missions. The principal difference was, once again, the energetic requirements. The model assumes that a Moon mission would require full lunar access, which innately requires a greater ΔV capability than the Apollo equatorial access requirements [27]. Once adjusted, the total IMLEO from the model estimate came down to +3.7%. However, a 20% difference in mass in the command module also affected the system, for similar reasons as the Mars deep space habitat. This resulted in a final IMLEO difference between the model and Apollo data of -3.4%.

Validation studies were also performed for the NEAs against preliminary data from NASA HAT. This validation study showed adequate agreement with the point designs. The generality of the assumptions in the model is more pronounced in the NEA designs due to the variety of possible destinations in the two primary categories. It is therefore re-emphasized that these are meant to be representative of the broader class of asteroids.

III. GENERATING RESULTS

Tradespace Issues and General Applicability

One of the greatest challenges of working in the field of systems architecture, compounded by trying to deal with large quantities of data, is the fact that the encoding of the systems usually requires the use of categorical variables. The assignment problem, when the choices are not numerical, is always encoded in this fashion. For example, the propellant options for the EDS are encoded as "LOX/LH_2", "LOX/CH_4", and "NTR". Therefore, when we look at the option space mathematically, there are no established mathematical differences between options, therefore there are no gradients and no contours of the overall space. This means that traditional methods of optimizing systems within the space are unlikely to be successful. Gradient-based methods, the more mathematically efficient and rigorous methods among optimization [28], clearly do not function. Gradient-free numerical methods require special forms of encoding of the space and have been shown to produce unreliable results for this model formulation. The final major group of optimization methods, heuristics, remains. By their nature, heuristics do not guarantee optimality, nor do they guarantee efficiency or convergence. It is non-trivial to attempt to apply them to a purely categorical model.

Design of Experiments (DOE) methods can be used to determine the major areas of the tradespace, both prior to an analysis to inform that analysis or post analysis to determine coverage of the space. DOE methods that reduce computational space often require that system elements are not coupled, meaning that there is no complex relationship between combinations of system elements. This allows the order in which decision choices are selected to be inconsequential to the outcome. This is never guaranteed with complex space systems or complex systems in general, which have many couplings between elements. *A priori* characterization of these couplings could lead to appropriate use of DOE methods for full analysis, but this requires a great deal of knowledge of the system prior to modeling. If such knowledge exists, the need for architecture-level studies probably does not exist.

With these constraints in mind and in an attempt to reduce the computational requirements for the analysis of this model, a Genetic Algorithm (GA) was developed and tested for the analysis along the IMLEO metric. Once the analysis code was optimized to a point where the computational expense was acceptable, full enumeration analysis was also employed.

Genetic Algorithm Optimization

For a given set of mission mode parameters, without logical reduction of the allocation tradespace, the total number of possible architectures exceeds 120 million. Even with highly simplified and computationally efficient analysis, calculation of metrics for this number of architectures is extremely expensive. One method of avoiding full tradespace searching is the use of heuristic algorithms, which search the space in an intelligent manner. When a space is well-structured, these algorithms work well at finding optimal or close to optimal solutions, although optimality is never guaranteed. However, as the categorical design space is not well structured, it becomes non-trivial to develop a heuristic algorithm that effectively covers the space of architectures.

Genetic algorithms (GA) use mutation of the encoded "alleles", which are simply pieces of system information, to search the overall space quasi-randomly. By encoding the choices that set the architectural definition as integer alleles and restricting the values possible in any given mutation, a categorical design space can be searched. To ensure adequate coverage, a large seed population is used, in this case 2000 architectures, all pre-selected and analyzed as good architectures. For each generation in the GA, a large population is retained to keep space coverage, and a high total generation count for convergence is needed. These properties make the GA less efficient than most implementations of heuristic methods, but it is one of the few ways to ensure adequate coverage in a poorly structured design space.

It should be noted that part of the overall design space issue stems from the formulation of the habitat and transportation allocations, which require 7 and 10 alleles to encode, respectively. Pure mutation within these variables should not be allowed, as many mutated combinations are not physically possible. This also means that the total tradespace that the GA theoretically looks at, which would include the mutation of these alleles, has vast regions of infeasibility. Furthermore, many of the feasible regions are sharply structured. This means that IMLEO shifts quickly from an infeasible region to "good" values, and the region of interest is littered with deep wells of good values surrounded by areas of infeasibility. This creates a tradespace similar to an extreme egg-crate problem, which is known to be one of the most difficult problems to optimize [29].

Through the reduction of the allowed tradespace, which eliminated areas of infeasibility by elimination of allowed mutations, such as the reduction of set partitioning schemes to the most prevalent ones in previous top architectures, and tweaking of the GA properties, a heuristic wrapper around the formulation and analysis of architectures was created. To analyze the tradespace, the GA typically required roughly 60 generations, each with 1000 architectures, initialized with randomly drawn architectures from a pre-analyzed seed population of 2000 architectures. It was also shown post-analysis that this method adequately covered the design space, as the major architectural groups of interest were represented in the various generations. It should also be noted here that this was performed only on one metric, IMLEO. It is possible to implement this method using a multi-objective genetic algorithm (MOGA). However, MOGAs require a great deal more computational power and resource consumption. The calculation of Number of Development Projects is relatively trivial, meaning that, although the optimization does not occur along both metrics, the architectures analyzed during the GA run can easily be analyzed in terms of the second metric. As all 60,000 architectures are also stored during this operation, a more complete Pareto front analysis can be performed using the full data set, although it will be skewed toward the group of architectures with lower IMLEO.

Full Enumeration Analysis

Through the use of CPU parallelization techniques and various efficiency increases within the code, the analysis time in real world hours was reduced 100-fold from the original method. This recent advance allowed for the reasonable analysis of all possible architectures within the tradespace. Following the GA analysis method, it was found that only approximately 1% of the tradespace of 120 million architectures were actually feasible due to mixing of propellant types within the propulsion stage allocations. Furthermore, only about 1/3 of those architectures were solvable with the Tsiolkovsky rocket equation. This means that 2/3 of the architectures included segments where the total energetic requirements on a propulsion stage were impossible to achieve, hitting the asymptote of the rocket equation. Overall, only 0.3% of the space is analyzed in full along the metrics of interest. This means that 120 million architectures are formulated, roughly 1 million pass through the logical filter, and 350,000 are fully analyzed. Even given this reduction, a full enumeration run for Mars missions requires 7 hours of computation time on 4 processor cores at 3.2GHz.

Full enumeration analysis does bring significant advantages, however. First and foremost, this allows for a guarantee of optimality, given that the model assumptions are taken to be correct. This also allows for analysis along the entire Pareto frontier and Pareto space. So long as adequate numbers of architectures with all various properties pass through the logical filters and can be solved in the fundamental physics, interactive effects between architecture properties can also be assessed with a more rigorous treatment. In short, this allows more rigorous treatment of all analysis due to the collection of all possible data from the model.

IV. FINDINGS

Top Architectures and Their Features

Full analysis results are not covered by this paper, as they are presented in a companion paper. Presented here are the top architectures in terms of IMLEO for the Moon and Mars cases. Key technology choices of interest are also shown in Figure 4. For clarification, these are the best overall architectures in each of the categories, not the top overall architecture analyzed without the various technology options. Some key features emerge. For example, ISRU is not as critically important as previously thought, as overall it saves approximately 21 metric tons from the total architecture, or 5%. Architectures without nuclear thermal rockets are also feasible for Mars missions, although they require close to twice the total mass. The total IMLEO for such a mission is approximately 131% of the International Space Station (ISS) mass [30]. Aerocapture at Mars also plays a big role, influencing the architectures even more than NTR. Boil-off control remains the most critical aspect for Mars missions, as even the best architecture without this technology is over 3 times the mass of the best architecture with boil-off control.

Similarly, trends emerge for architectures for Moon missions. Aerocapture difference appear due to the use of aerocapture on return to Earth's atmosphere, not a false use of aerocapture at the Moon. SEP pre-deployment aids the architectures by about 10 metric tons. Boil-off control has much less impact due to the short duration of mission segments to and from the Moon. However, it does make some difference, as return propulsion stages must sit in orbit for the duration of the surface mission.

Analysis is ongoing for both NEA cases, and so results are not presented here.

	MOON	MARS
Best Architecture	IMLEO (mt)	IMLEO (mt)
Overall	65	383
No ISRU	-	404
No NTR	-	588
No Aerocapture	84	653
No SEP	75	-
No Boil-Off control	73	1,238

Figure 4: Top Architectures for Moon and Mars

V. CONCLUSIONS

Design of a system-level model that captures the key aspects of manned exploration missions is not a trivial task. An understanding of the metrics of interest drives the decomposition of the system into architecturally distinguishing features. In this case, a core functional decomposition allowed for more rigorous studies of habitat and transportation allocations. When coupled with a suite of technology options, a broad range of possible architectures appears in the design space. A key property of the functional decomposition is the determination of a set of invariant functions present in all possible architectures. These functions are fundamental to exploration systems, and expanding upon these functions ensures a broader analysis of the tradespace. Also fundamental to the usefulness of this model is the iso-performance characteristic along scientific value. This eliminates the inherent trade between science performance and system cost. Furthermore, decoupling from surface operations and launch systems allows for an objective view of the "best" in-space infrastructure architectures.

It was shown that the use of this decomposition method allows for the creation of a modeling environment that matches well with both heritage systems and the most vetted paper models. It was further shown that heuristic methods can be successfully applied to the design space, although optimality is not guaranteed and this requires significant adjustments to baseline genetic algorithms. The advantages of full enumeration, focusing on the rigorous nature of analysis, were also discussed. Top architectures from this analysis were shown. It was noted that although NTR technology is hotly debated in terms of use in Mars missions, boil-off control is actually the main driver for IMLEO. Missions remain feasible, although at a great cost, without NTR. Aerocapture was also shown to be a significant cost driver, thus indicating that development of large-scale aeroshields should be prioritized in the development of Mars infrastructures. Lunar return missions were also analyzed, showing relatively small advantages with the inclusion of various technologies.

Complete analysis results will be presented in future work, but a great deal of analysis relating to technology investment is present in [31].

VI. ACKNOWLEDGEMENTS

This study was conducted under the NASA-MIT research grant "Comprehensive Analysis and Synthesis of Exploration Architectures", sponsored by NASA HEOMD.

VII. REFERENCES

[1] International Space Exploration Coordination Group (ISECG). *The Global Exploration Roadmap*. 2011.

[2] Price, H., Hawkins, A., and Radcliffe, T. *Austere Human Mission to Mars*. AIAA. 2009.

[3] Willson, D. and Clarke, J. *A Practical Architecture for Exploration-Focused Manned Mars Missions Using Chemical Propulsion, Solar Power Generation, and In-Situ Resource Utilization.* 6th Australian Space Conference. 2006.

[4] Simmons, W. *A Framework for Decision Support in Systems Architecting.* MIT Ph.D. Thesis, Cambridge MA. 2008.

[5] Koo, B., Simmons, W., and Crawley, E. *Algebra of Systems: An Executable Framework for Model Synthesis and Evaluation*. International Conference of Systems Engineering and Modeling. 2007.

[6] NASA Review of Human Spaceflight Plans Committee. *Final Report: Seeking a Human Spaceflight Program Worthy of a Great Nation*. 2009.

[7] NASA HAT. *HAT Cycle-C Draft Briefing for MIT*. Slide 14. 2011.

[8] Jones, T. *et al*. *The Next Giant Leap: Human Exploration and Utilization of Near-Earth Objects*. The Future of Solar System Exploration. 2002.

[9] XCOR Aerospace Inc. *XR-5M15, LOX/Methane Main Engine*. Online. Available <http://www.xcor.com/products/engines/5M15_LOX-Methane_rocket_engine.html>

[10] Selva, D. *Rule-Based System Architecting of Earth Observation Satellite Systems*. MIT Ph.D. Thesis, Cambridge, MA. 2012.

[11] Cohen, R., Katzir, L., and Raz, D. *An Efficient Approximation for the Generalized Assignment Problem*. Information Processing Letters, Vol. 1000, Issue 4. 2006.

[12] Garey, M., and Johnsen, D. *Computer and Intractability: A Guide to the Theory of NP-Completeness*. A Series of Books in the Mathematical Sciences. W.H. Freeman and Co. 1979.

[13] Hofstetter, W. *Extensible Modular Landing Systems for Human Moon and Mars Exploration*. MIT Masters Thesis, Cambridge, MA. 2004.

[14] Dyson, M. *Space and Astronomy: Decade by Decade*. Twentieth-Century Science. P.142. 2007.

[15] Jorgensen, M. *A Systematic Review of Software Development Cost Estimation Studies.* IEEE Software Engineering. 2007.

[16] Wertz, J. and Larson, W. *Space Mission Analysis and Design, Third Ed.* Pp. 790. 1999.

[17] Aliakbargolkar, A. *et al. Heavy Lift Launch Vehicle Systems Architecting.* 62nd International Astronautical Congress, Cape Town, S.A. 2011.

[18] Garfinkel, R., and Nemhauser, G. *The Set-Partitioning Problem: Set Covering with Equality Constraints*. JSTOR Operations Research. 1969.

[19] Garey, M., Johnson, D., and Stockmeyer, L. *Some Simplified NP-Complete Graph Problems.* Theoretical Computer Science. 1976.

[20] К. Э. Циолковский, *Исследование мировых пространств реактивными приборами*. 1903.

[21] NASA HAT. *Cycle B Concept Description Document.* 2011.

[22] NASA. *Human Exploration of Mars Design Reference Architecture 5.0*. NASA-SP-2009-566. 2009.

[23] Williams, D. *A Crewed Mission to Mars…* Online. Available <http://nssdc.gsfc.nasa.gov/planetary/mars/marslaun.html>.

[24] Schaffer, M. *A Study of CPS Stages for Missions Beyond LEO*. Future In-Space Operations Working Group. SpaceWorks Inc. 2012.

[25] Simon, M., *et al. Factors Impacting Habitable Volume Requirements: Results from the 2011 Habitable Volume Workshop*. NASA/TM-2011-217352. 2011.

[26] NASA. *Human Integration Design Handbook*. NASA/SP-2010-3407. 2010.

[27] NASA. *NASA's Exploration Systems Architecture Study: Final Report*. NASA-TM-2005-214062. 2005. Pg. 75-216.

[28] Jameson, A. *Gradient Based Optimization Methods*. Princeton University. MAE Technical Report No 2057. 1995.

[29] Hassan, R., Cohanim, B., and de Weck, O. *A Comparison of Particle Swarm Optimization and the Genetic Algorithm*. M.I.T., Cambridge, MA. AIAA. 2004.

[30] NASA. *ISS Facts and Figures.* Online. Available: <http://www.nasa.gov/mission_pages/station/main/onthestation/facts_and_figures.html>

[31] Battat, J. *Technology and Architecture: Informing Investment Decisions for the Future of Human Space Exploration.* MIT Master's Thesis, Cambridge, MA. 2012.