Using Technical Performance Measures

Christopher J. Garrett[1]
Pratt & Whitney Rocketdyne, Canoga Park, California, 91309

Daniel J. H. Levack[2]
Pratt & Whitney Rocketdyne, Canoga Park, California, 91309

Russel E. Rhodes[3]
NASA, Kennedy Space Center, Florida, 32899

All programs have requirements. For these requirements to be met, there must be a means of measurement. A Technical Performance Measure (TPM) is defined to produce a measured quantity that can be compared to the requirement. In practice, the TPM is often expressed as a maximum or minimum and a goal. Example TPMs for a rocket program are: vacuum or sea level specific impulse (Isp), weight, reliability (often expressed as a failure rate), schedule, operability (turn-around time), design and development cost, production cost, and operating cost. Program status is evaluated by comparing the TPMs against specified values of the requirements. During the program many design decisions are made and most of them affect some or all of the TPMs. Often, the same design decision changes some TPMs favorably while affecting other TPMs unfavorably. The problem then becomes how to compare the effects of a design decision on different TPMs. How much failure rate is one second of specific impulse worth? How many days of schedule is one pound of weight worth? In other words, how to compare dissimilar quantities in order to trade and manage the TPMs to meet all requirements. One method that has been used successfully and has a mathematical basis is Utility Analysis.

Utility Analysis enables quantitative comparison among dissimilar attributes. It uses a mathematical model that maps decision maker preferences over the tradeable range of each attribute. It is capable of modeling both independent and dependent attributes. Utility Analysis is well supported in the literature on Decision Theory. It has been used at Pratt & Whitney Rocketdyne for internal programs and for contracted work such as the J-2X rocket engine program.

This paper describes the construction of TPMs and describes Utility Analysis. It then discusses the use of TPMs in design trades and to manage margin during a program using Utility Analysis.

Nomenclature

ATP	= Authority to Proceed
IOC	= Initial Operating Capability
IR&D	= Internal Research and Development
NASA	= National Aeronautics and Space Administration
O/S	= Operations and Support
PWR	= Pratt & Whitney Rocketdyne
RS	= Rocket System
T	= Time
TPM	= Technical Performance Measure

[1] Systems Architecture, J-2X Program, Pratt & Whitney Rocketdyne, P.O. Box 7922, 6633 Canoga Ave, Canoga Park, CA 91309/ MS RFB19.

[2] Program Manager, Advanced Programs, Pratt & Whitney Rocketdyne, P.O. Box 7922, 6633 Canoga Ave, Canoga Park, CA 91309/ MS RFB19, AIAA Member.

[3] AST, Technical Management, Engineering Directorate Design & Development Eng Div Sys Engineering & Integration Br, Kennedy Space Center, Florida/NE-D2, AIAA Senior Member.

Introduction

A LL programs have requirements. For these requirements to be met, there must be a means of measurement. A Technical Performance Measure – a TPM – is defined to produce a measured quantity that can be compared to the requirement. In practice, the TPM is often expressed as a maximum or minimum and a goal. Examples of TPMs for a rocket program are: vacuum or sea level specific impulse (Isp), weight, reliability (often expressed as a failure rate), schedule, operability (turn-around time), design and development cost, production cost, and operating cost. Program status is evaluated by comparing the TPMs against specified values of the requirements. During the program many design decisions are made and most of them affect some or all of the TPMs. Often, the same design decision changes some TPMs favorably while affecting other TPMs unfavorably. The problem then becomes how to compare the effects of a design decision on different TPMs. How much failure rate is one second of specific impulse worth? How many days of schedule is one pound of weight worth? In other words, how to compare dissimilar quantities in order to trade and manage the TPMs to meet all requirements.

One Possible Approach – Utility Analysis

One method that has been used successfully at Pratt & Whitney Rocketdyne (PWR) is Utility Analysis. Utility Analysis enables quantitative comparisons among dissimilar attributes. In other words, it takes apples and oranges and converts them all into peaches. For programs that have competing objectives, a multi-attribute (multi-TPM) utility function can be used to resolve the tradeoffs among the different TPMs.

Utility is a mathematical representation of how much a decision maker values different amounts of one or more attributes. The mathematical model maps decision maker preferences over a trade space. A utility function, $u_i(x_i)$, is constructed for each attribute, modeling the strength of preference over the feasible range of the attribute. The utility functions are developed through a set of exercises with the decision maker evaluating hypothetical trades and degrees of risk aversion. Then the overall utility is (usually) a weighted sum of the values of the individual attribute utility functions.

The focus on the decision maker directly captures the sensitivity of the decision maker's expectations, thus the process captures the decision maker's subjective influences.

Utility Analysis can be used for non-linear and for dependent preferences. It is also very efficient for performing large numbers of trades. It requires some up-front effort to develop the model, but there is very little recurring effort needed during the program. The lack of recurring effort is in contrast to tools such as AHP or Accord which develop a new model for each individual trade.

Utility Analysis is well-supported in academic literature on Decision Theory, and decisions based on utility satisfy "Axioms of Rational Behavior".[1-5]

Overall, Utility Analysis provides a formal, traceable, repeatable, and consistent basis for making decisions within a program.

Using Utility Analysis

The first step in using Utility Analysis is to determine what TPMs are going to be used to manage the trades or the program. For a rocket engine, a common set of TPMs is: performance, weight, reliability, operations, cost, and schedule. Others could be used, but this set is typical. Each of these TPMs needs to be broken down to one or more measurable attributes. Figure 1 shows a set that could be used. For example: cost is broken down into non-recurring cost, recurring cost, and operations and support cost. For each of the eight attributes shown in the figure a utility curve is then constructed. Then the overall utility will be a function of the eight attributes that together represent the overall strength of the customer preference:

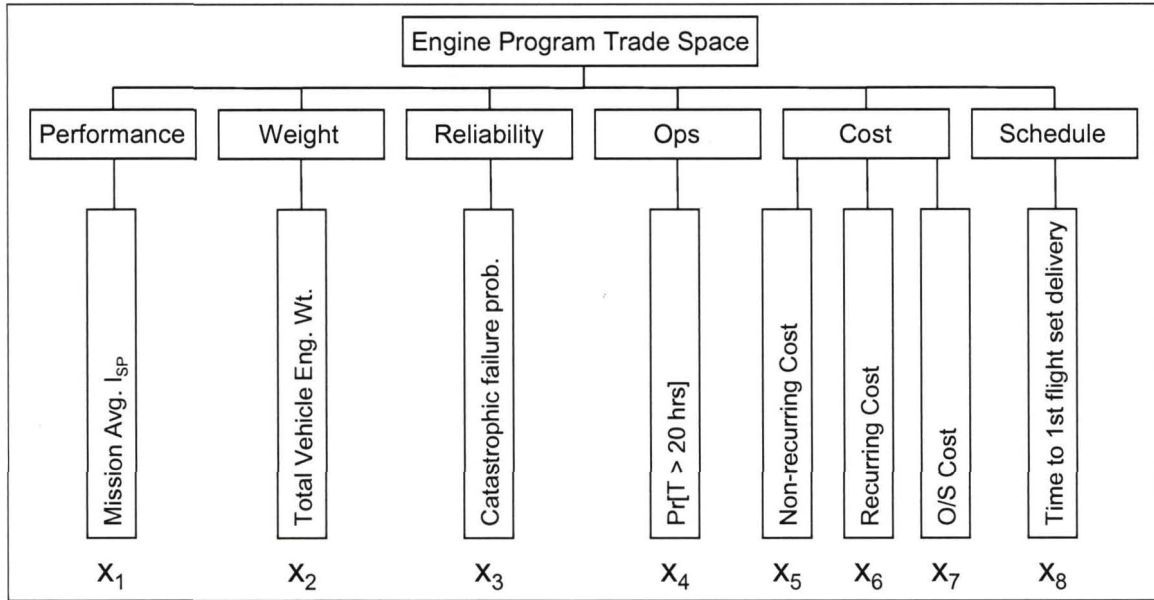$$U(x_1, x_2, \ldots x_n) = k_1 u_1(x_1) + k_2 u_2(x_2) + \ldots + k_n u_n(x_n)$$

Figure 1. Typical set of attributes for a set of TPMs.

Generating the Individual Utility Curves

The goal is to determine the customer preferences, so the involvement of the decision maker is imperative. The decision maker may defer to judgments of a number of individuals with specific domain knowledge, but in that case, the number should be kept small because the time to perform the process described below increases quickly as the number of people involved increases.

The process is: considering one attribute at a time, the decision maker is asked to trade outcomes that are certain against outcomes that are uncertain. This will establish risk aversion versus risk seeking attitudes over the attribute range. It identifies "pain thresholds" and points of diminishing return.

For each attribute, a minimum and a maximum value that bounds the trade space is selected. These should be extreme enough to serve as a boundary, but not so extreme as to be unrealistic. The maximum value of the attribute is assigned a utility of 1 and the minimum value is assigned a utility of 0. For all of the points between these two, the utility is scaled such that the expected utility of a lottery of outcomes is equal to the utility of the certain outcome that the decision maker feels is of equal preference.

In more process oriented terms, the utility curve can be generated by using the following process.

Step 1. Agree on the minimum and maximum possible values of the attribute in question. This agreed-upon minimum value is referred to in the questions that follow as x_0 and the maximum value is referred to as x_1.

Step 2. Arbitrarily select a value between x_0 and x_1. This value is referred to below as x^*. x^* should be close enough to either x_0 or x_1 that the respondent should have a clear preference in answering the following question:

Question 1: "You have two alternatives, one with an outcome that is certain, and one with an outcome that is uncertain. If you select the first alternative, you are guaranteed an outcome of x^*. If you select the second alternative, there is a 50% chance the outcome will be x_1, and a 50% chance the outcome will be x_0. Which of the two alternatives would you prefer?"

Step 3. Select a new x^*, this time choosing a value toward the opposite end of the spectrum, so that the respondent, in answering Question 1, would have an equally clear preference for the opposite alternative. Repeat Question 1 with the new value for x^*.

Step 4. Continue to repeat Question 1, each time selecting a new value of x^* from alternating sides of the X spectrum, such that each time the respondent's preference will become less clear and the

decision becomes more difficult. When a value of x^* is reached for which the decision cannot be made and there is no clear preference between the alternatives of Question 1, we'll say we have reached an *indifference point* and the value of x^* will be denoted by $x_{0.5}$.

Step 5. Repeat Steps 2-4, replacing x_0 with $x_{0.5}$. The resulting indifference point is denoted by $x_{0.75}$.

Step 6. Repeat Steps 2-4, replacing x_1 with $x_{0.5}$. The resulting indifference point is denoted by $x_{0.25}$.

Following this procedure will result in the determination of three interior points on the utility curve: $x_{0.5}$, $x_{0.75}$, *and* $x_{0.25}$. As a sanity check, these three points may correspond roughly to the following quantities from the decision maker's perspective:

$x_{0.5}$ just enough (of this attribute) to keep the program viable;
$x_{0.75}$ point with enough margin that it is a "slam-dunk";
$x_{0.25}$ probably untenable, but enough to "prolong the pain".

The complete utility curve for this particular attribute is the best fit through the five points x_0, $x_{0.25}$, $x_{0.5}$, $x_{0.75}$, and x_1. An example of the type of curve to be expected is shown in Figure 2.
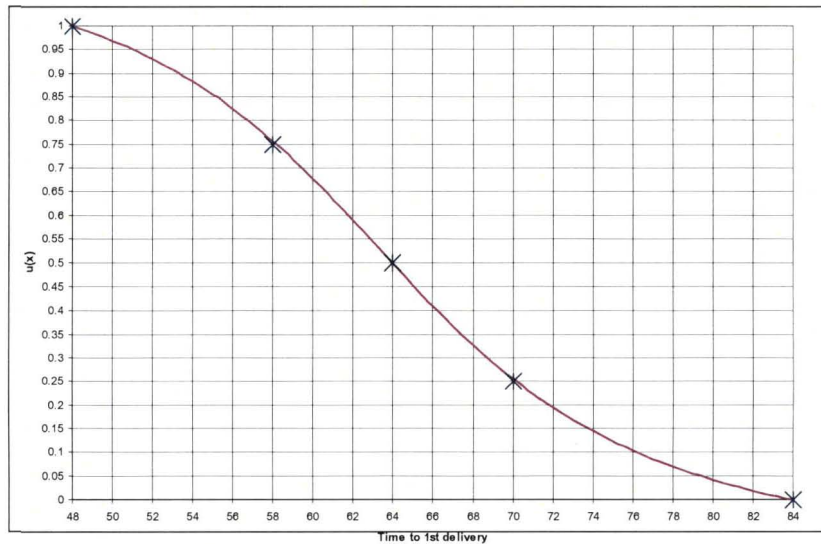

Figure 2. Typical single attribute utility curve.

Utility Curve Characteristics

If assessed over a broad enough range, most utility curves will typically take an "S" shape with three distinct regions as illustrated in Figure 3.
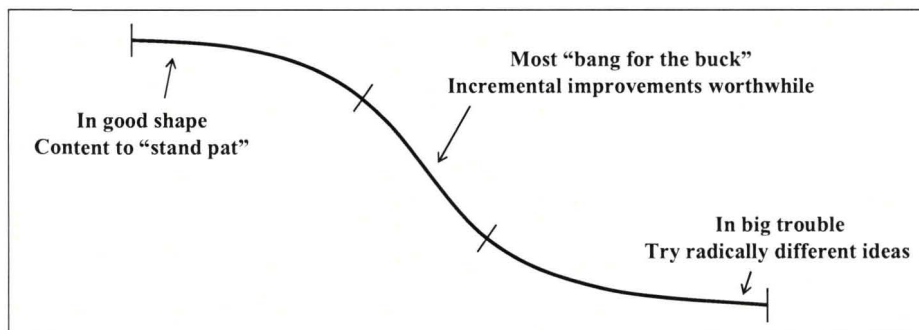

Figure 3. Interpretation of typical utility curve.

Examining Figure 3, a number of observations can be made. Portions of the curve that are concave down indicate regions where the decision maker is risk averse. Portions that are concave up indicate regions where the decision maker is risk seeking. And linear segments of the curve indicate regions in which the decision maker is risk neutral. The slope of the utility curve indicates the relative worth of marginal improvement about any point on the curve.

Risk aversion means that the benefit of improving attribute x by an amount δ is less than the harm in changing it by $-\delta$. On the risk averse portion of the curve the decision maker would rather not take a 50-50 chance of changing the attribute value by $+/-\delta$. Risk seeking has a similar, but opposite, interpretation.

The risk averse portion of the curve approaching the utility of 1 is a reflection of the law of diminishing returns, while risk seeking is observed as the utility of 0 is approached.

Some decision makers claim to be risk neutral in all regions above some minimum threshold – in effect denying the law of diminishing returns. Whether this is a reasonable attitude is a subject of some debate and should be discussed carefully with the decision maker and treated accordingly in the construction of the curve.

The concepts of risk aversion, risk seeking, and risk neutrality are very useful in interpreting the shape of the utility curve and can be used to "troubleshoot" curves that do not come out looking the way the decision maker expected.

Depending on the range of the attribute chosen, the utility curve can be any subsegment of the general shape shown in Figure 3.

Scaling Factor Determination

Each utility curve in the overall utility function has a scaling (or weighting) factor. The utility curve is of the form $U_i = k_i u_i(x_i)$. The scaling factor, k_i, for the utility curve for each attribute is set by pairwise comparison with other attributes. Using Figure 4, x_1^+ represents the value of x_1 with the greatest utility. If the decision maker prefers point A to point B, then $k_1 > k_2$. Then finding point A', which has equal value to the decision maker as point B, sets the value of k_1 relative to k_2. Scaling n utility curves requires n-1 pairwise comparisons of attributes for n equations in n unknowns. The sum of the weighting factors equals 1, i.e., $\Sigma k_i = 1$. As an example, assuming a case with four attributes, then there are seven possible equations available plus the fact that the sum of the k's equals one:

$$k_1 = k_2 u(x_1') \qquad k_2 = k_3 u(x_2')$$
$$k_1 = k_3 u(x_1') \qquad k_2 = k_4 u(x_2')$$
$$k_1 = k_4 u(x_1') \qquad k_3 = k_4 u(x_3')$$
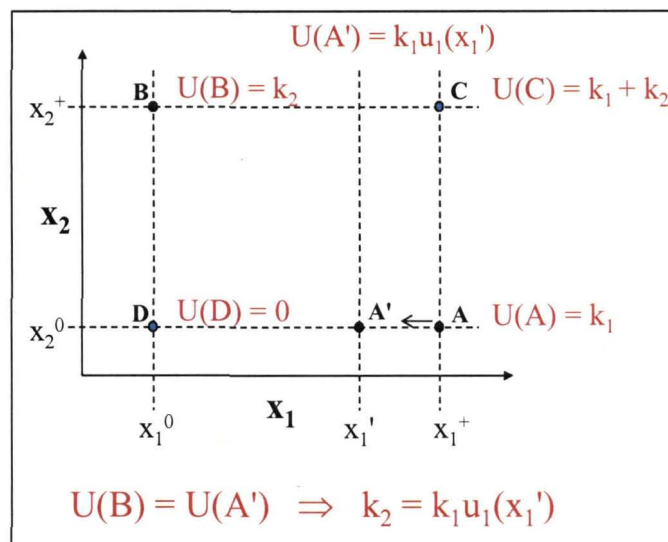$$k_2 = k_2 u(x_2') \qquad \Sigma k_i = 1.$$

Figure 4. Setting the utility curve weighting factors.

Only four equations are needed to determine the k's. Pick the combinations that have attributes that are reasonably comparable, e.g., schedule and cost, not cost and center of gravity. Then since the problem is over determined, choose a couple of others as cross-checks.

Dual-Attribute Utility Curve Construction

What happens if the decision maker's preferences are not independent? In other words, what if the utility of x_1 depends on the value of x_2, and vice versa? Assume that the utility of x_2 given a particular value of x_1, $u(x_2|x_1)$, can be described by a family of functions over x_2 whose parameters are a function of x_1. The entire dual-attribute utility surface can be generated by calculating three single-attribute utility curves.

Figure 5 shows an example of a dual attribute utility where the weight (wt), denoted as x_2, and the center of gravity (cg), denoted as x_1, are preferentially dependent. Looking at Figure 5, the upper and lower utility curves are calculated using the procedure described earlier in the section "Generating the Individual Utility Curves," with one difference. In this case, the utility at points f and g of Figure 5 are known (1 and 0). However, the utility at points e and h are not known. It is necessary to independently assess the utilities of points e and h relative to point f. This is accomplished by again asking the decision maker to choose between an outcome that is certain and an outcome that is uncertain. In this case, the certain outcome is the one represented by point e (or h), and the uncertain outcome is one that results in point f with probability p, and point g with probability $1-p$. The probability p_e (or p_h) which leads to the decision maker being indifferent to the uncertain outcome relative to the certain outcome is then, by definition, equal to the expected utility of point e (or h)..

Having developed two consistently scaled conditional utility curves for x_1, a conditional utility curve for x_2 is then calculated using the same procedure, and then scaled such that its endpoints meet the boundary conditions established by the two x_1 curves.
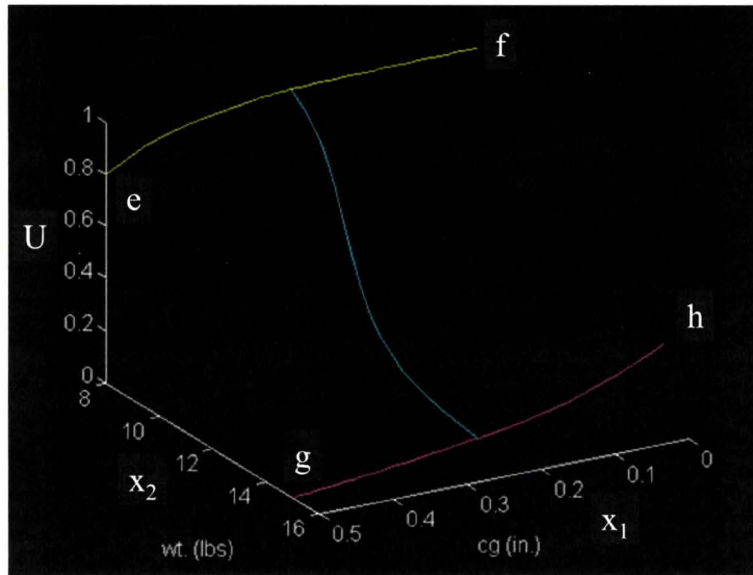

Figure 5. Conditional single-attribute utility curves.

Now assume that the same functional form can be used to generate a family of conditional utility curves for x2, given different values of x1. For the following discussion, an arctangent form will be assumed. Then the functional form is:

$$U(x_2|x_1) = C_1 \, atan[C_2(x_2 + A)] + B$$

Where C_1 and C_2 are constants and A and B are functions of x_1. A series of parallel slices are taken at various values of x_1 that span its range. Nine slices work well, but the number must be such that curve fits

of A and B can be performed. For each slice the boundary conditions (the two x_1 curves) are known, yielding two equations in two unknowns:

$$u_1(x_1,x_2^*) = C_1 \, atan[C_2(x_2^* + A(x_1))] + B(x_1)$$

and

$$u_1(x_1,x_2^0) = C_1 \, atan[C_2(x_2^0 + A(x_1))] + B(x_1)$$

where
x_2^* is the x_2 for best utility and x_2^0 is the x_2 for worst utility.

This process generates a set of points for $A(x_1)$ and $B(x_1)$. These sets of points are then curve fit to yield the functional forms of A and B. Figure 6 shows the surface build up.
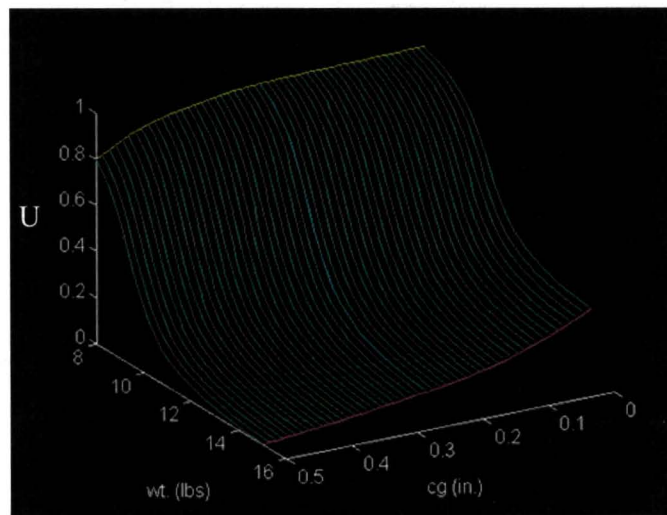

Figure 6. Utility surface construction.

Figure 7 shows the resultant utility surface for this example. If another set of parallel slices are taken at values of constant utility, Figure 8 results. Figure 8 is an example of a set of indifference curves. For any given utility plotted on the figure, there is no change in utility (preference) between weight and center of gravity along that line.
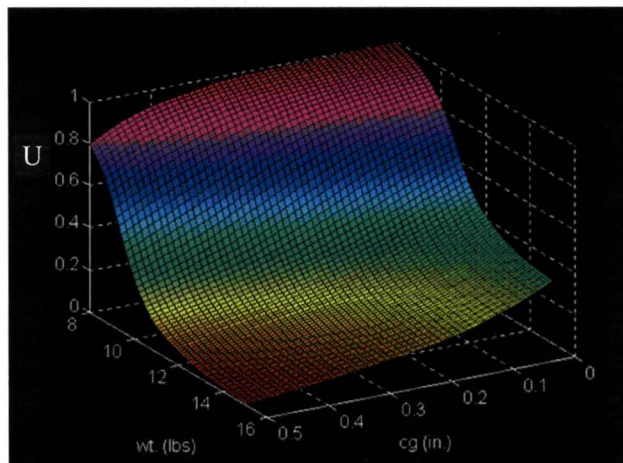

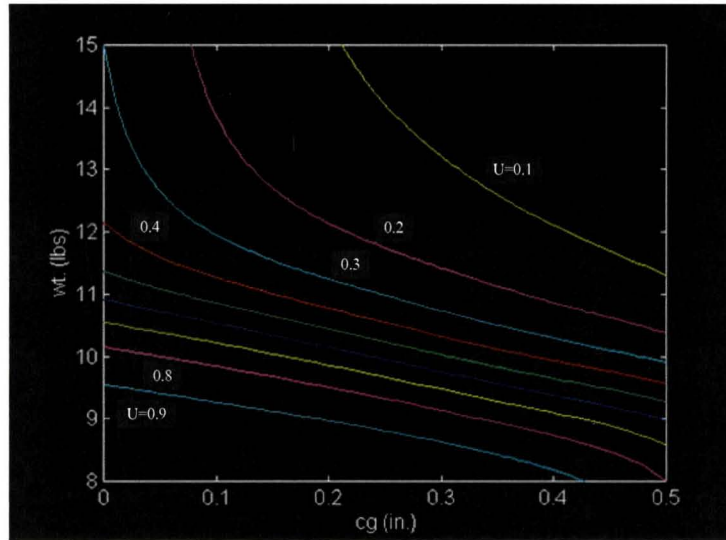**Figure 7. Dual-attribute utility function $u_{12}(x_1,x_2)$.**

Figure 8. Example of indifference curves.

Figure 9, another indifference curve, illustrates that the exchange factor between any two attributes, whether or not preferentially dependent, can, and probably will, be different depending on the current value of each attribute. Consequently, during the course of a program, the optimum trades between them will vary. On any constant utility curve the direction of maximum increase of utility is perpendicular to the line of constant utility, and at different points in the program, this will change. To illustrate, consider the lower right quadrant of Figure 9. If the IOC date is around June 2011 and delivered payload is in the low 50,000s, then increasing payload capability is more important to the decision maker than improving IOC date. Conversely, if the expected IOC date has slipped to January 2012, but payload has increased to the upper 50,000s, then shortening the schedule becomes more important to the decision maker than adding more payload. The indifference curves can be useful for quick visual information about the relative importance of various attributes.
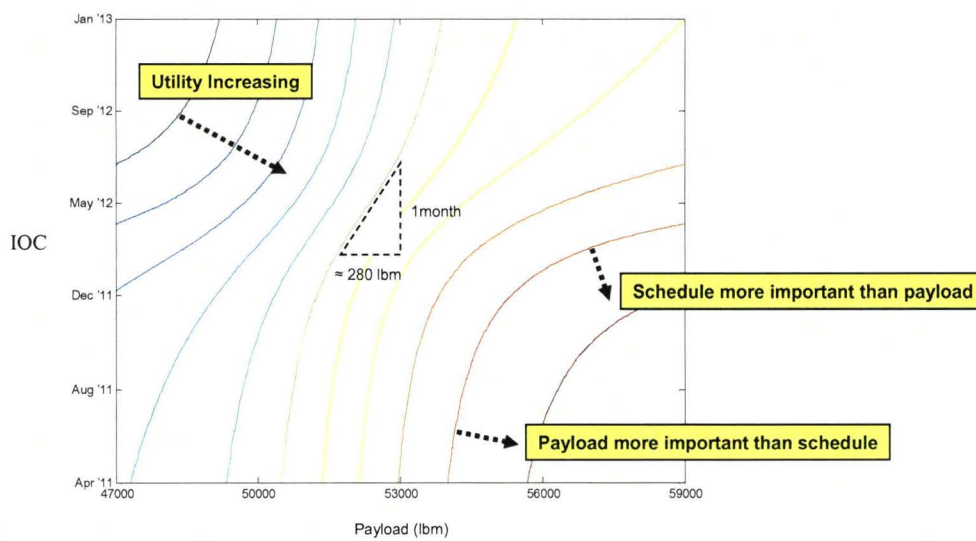

Figure 9. Direction of optimality can change as design matures.

Using Utility Functions for Trades

For each option in the trade space, the values of the affected attributes are calculated. Then the utility function is evaluated for each option. The trade can then be compared on the basis of utility to the decision maker. Figure 10 shows the steps in this methodology. Note that the utility curves do not need to be regenerated for additional trades even with entirely new options.
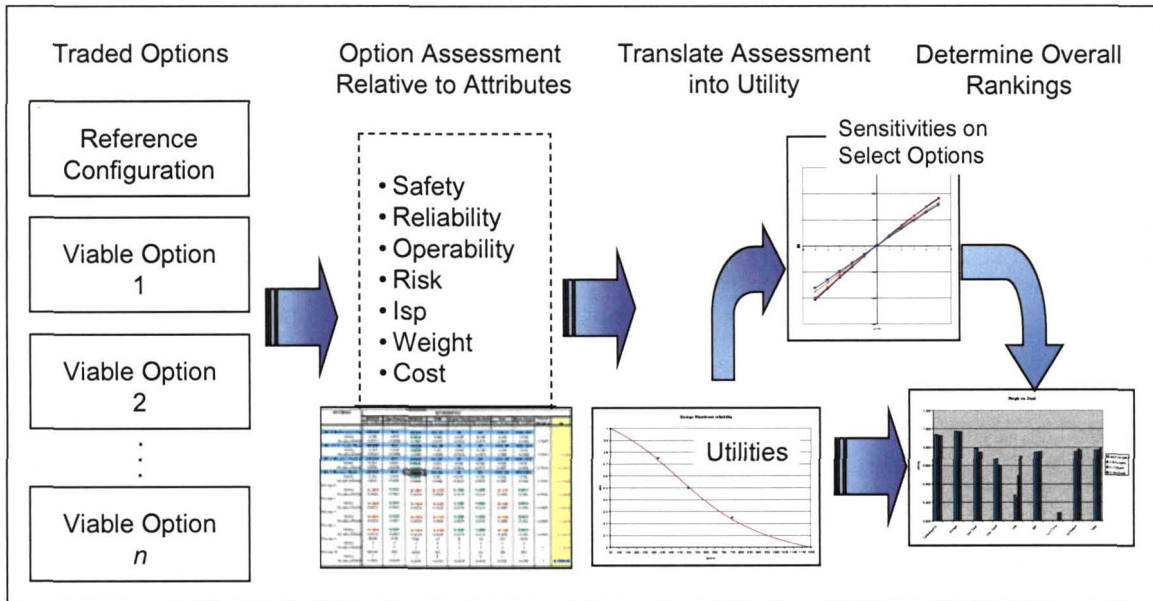

Figure 10. Utility-based trade methodology.

Figure 11 shows an example of a trade in a totally hypothetical program developing a vehicle architecture. The engine parameter values do not represent any real engines. The trade is between vehicles using seven and eight engines. The eight engine vehicle has better specific impulse, better non-recurring cost, and equal operation and support costs, whereas the seven engine vehicle is somewhat better in the other TPMs. Which configuration should be picked and how should the decision be justified? This could be a difficult decision. But by using Utility Analysis and applying it to this case, the results show that, for the preferences and degree of risk aversion exhibited by the decision makers for this example, there is a significant preference for the seven engine configuration.

	7 Engines	8 Engines
Performance		
Mission Avg. I_{SP}	435.8 sec	436.4 sec
Weight		
Total Vehicle Eng. Wt.	48,547 lb	49,732 lb
Reliability		
Catastrophic failure prob.	0.99984	0.99972
Ops		
Pr[T > 20 hrs]	0.45	0.50
Cost		
Non-Recurring	$859M	$849M
Recurring	$632M	$709M
O/S	$31.1M	$31.1M
Schedule		
ATP to 1st set delivery	46.5 mo	48 mo
Overall Utility	*0.846*	*0.753*

Figure 11. Example RLV system-level trade.

Utility Analysis provides a convenient method for developing trade factors for use during design. At any point in the program the values of the TPM attributes are known. In such a case, the partials for each attribute can be calculated. Figure 12 shows such a set for the previous example program. Now the impact of a change in any one attribute can be seen and comparisons, such as how much one second of specific impulse is worth in pounds of weight, can be calculated. For this example 0.1 seconds is worth 69.5 pounds.

$$
\begin{array}{lll}
I_{SP} & \partial U/\partial x_1 & = \ \ 0.025 \text{ / sec} \\
\text{Weight} & \partial U/\partial x_2 & = -0.036 \text{ / klb} \\
\text{Non-Recurring Cost} & \partial U/\partial x_3 & = -0.042 \text{ / \$B} \\
\text{Recurring Cost} & \partial U/\partial x_4 & = -0.028 \text{ / \$B} \\
\text{Reliability} & \partial U/\partial x_5 & = -0.00192 \text{ / unit prob.} \\
\text{Schedule} & \partial U/\partial x_6 & = -0.046 \text{ / mo} \\
\text{O/S Cost} & \partial U/\partial x_7 & = -0.673 \text{ / \$B} \\
\text{Ops Turnaround} & \partial U/\partial x_8 & = -0.093 \text{ / unit prob.}
\end{array}
$$

Figure 12. Use of utility surface factors for sensitivity analysis.

Summary/Conclusions

PWR has successfully used Utility Analysis on the RS-83, RS-84, and J-2X programs. It has been used on a number of NASA study contracts as well as on IR&D efforts.

Utility Analysis is well-supported in the academic literature on Decision Theory. And decisions based on utility satisfy "Axioms of Rational Behavior". Overall, Utility Analysis provides a formal, traceable, repeatable, and consistent basis for making decisions within a program.

References

[1] Keeney, R. L., Raiffa, H., *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, Graduate School of Business and Kennedy School of Government, Harvard University, 1976.

[2] Mosteller, F., Nogee, P., "An Experimental Measurement of Utility," *Journal of Political Economy*, 1951.

[3] Grayson, C. J., *Decisions Under Uncertainty: Drilling Decisions by Oil and Gas Operators*, Division of Research, Harvard Business School, Boston, 1960.

[4] Thurston, D. L., "Multiattribute Utility Analysis in Design Management," *IEEE Transactions on Engineering Management*, 1990.

[5] Stanney, Pet-Edwards, Swart, Safford, Barth, "The Design of a Systematic Methods Improvement Planning Methodology: Part II - A Multiattribute Utility Theory (MAUT) Approach," *Int'l Journal of Industrial Engineering*, 1994.