Risk Analysis for Non-Deterministic Mission Planning and Sequencing[1][2]

Kar-Ming Cheung[3], Adans Ko[4], Van Dang[5], David Heckman[2]

Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109
818-393-0662
Kar-Ming.Cheung@jpl.nasa.gov

Abstract—In this paper we address the dilemma of planning in the presence of uncertainty – the problem of scheduling events where some events might have non-deterministic durations.

Real world planning and scheduling problems are almost always difficult. Planning and scheduling of events with a mixture of deterministic and non-deterministic durations is particularly challenging. The idea of scheduling events into a conflict-free plan becomes obscure and intangible when event durations are not known in advance – there is no guarantee that when the plan is executed, the scheduled events would not violate any pre-defined rules and constraints, and the resource usages would not exceed their maximum allowable limits. This dilemma of not being able to a priori quantify the likelihood of achieving a conflict-free plan in the presence of uncertainty usually results in an overly conservative plan where resources are under utilized.

Making use of some standard communication link analysis techniques to characterize communication system performance, to support tradeoffs, and to manage the operational risks associate with the link usage, we instigate a probabilistic description of event durations and introduce the notion of risk in terms of probability that the plan fails to execute successfully, which we denote as P_F. We attempt to define a rational and systematic approach to weight risk against efficiency by iteratively applying constrained optimization algorithms and Monte Carlo simulations to the plan. We also derive a simple upper bound of P_F for a given plan, which is independent of the optimization algorithm. This risk management approach allows planners to quantify the risk and efficiency tradeoff in the presence of uncertainty, and help to make forward-looking choices in the development and execution of the plan.

Another emphasis of this paper is to demonstrate that the general criteria of optimality and rules and constraints for event planning can be described mathematically in terms of linear and non-linear functions and inequalities. This allows the use of customized and commercial off-the-shelf (COTS) constraint optimization algorithms to generate conflict-free plans.

The results described in this paper are applicable to many general planning and scheduling problems. However the emphasis of this work is on mission planning and sequencing of spacecraft events with a mixture of deterministic and non-deterministic durations.

Mission planning and sequencing is a critical component for mission operations. It provides a mechanism for scientists and engineers to operate the spacecraft remotely from the ground. It translates the science intents and spacecraft health and safety requests from the users into mission plans and sequences. After a rigorous process validating the plan, the plan will be transmitted to the spacecraft for its execution. Usually mission planning and sequencing and its validation are time consuming and costly operations.

We apply the aforementioned methodology for formulating and optimizing both deterministic and non-deterministic sequence events planning. We demonstrate this approach with examples of scheduling science and engineering activities for mission operations.

[1] 0-7803-8870-4/05/$20.00 © 2005 IEEE
[2] IEEEAC paper #1292, Version 11, Updated Decmeber 10, 2004
[3] Communications Systems and Research Section, JPL
[4] Mission System Engineering Section, JPL
[5] Avonic System Engineering Section, JPL

TABLE OF CONTENTS

1. INTRODUCTION

In this paper we outline the mathematical formation and the analysis approach for planning and scheduling of events of non-deterministic durations. We define a plan to consist of a number of events, each of which has a certain start time and duration. By planning, we mean the process of a priori scheduling the events within the planning horizon. When the events are executed according to plan, the occurrence of events might consume one or more resources that are bounded and are either replenishable or non-replenishable. Also we assume that there is a set of rules and constraints that governs the relationships and dependencies among the events. A plan is defined to be successful if 1) all events can be accommodated within the planning horizon, 2) there is no resource usage that exceeds the maximum allowable limit, and 3) there is no violation to the set of pre-defined rules and constraints.

Real world planning and scheduling problems are almost always difficult. Planning and scheduling of events with a mixture of deterministic and non-deterministic durations is particularly challenging. The idea of scheduling events into a conflict-free plan becomes obscure and intangible when event durations are not known in advance – there is no guarantee that when the plan is executed, the scheduled events would not violate any pre-defined rules and constraints, and the resource usages would not exceed their maximum allowable limits. This dilemma of not being able to a priori quantify the likelihood of achieving a conflict-free plan in the presence of uncertainty usually results in an overly conservative plan where resources are under utilized.

Inspired by the communication link analysis techniques to characterize communication system performance, to support tradeoffs, and to manage the operational risks associate with the link usage, we instigate a probabilistic description of event durations and introduce the notion of risk in terms of probability that the plan fails to execute successfully, which we denote as P_F. We attempt to define a rational and systematic approach to weight risk against efficiency of the plan in the presence of uncertainty. We also derive a simple upper bound of P_F for a given plan that is independent of the optimization algorithm.

Much work has been done in the area of constrained optimization algorithms to support planning and scheduling applications. The anatomy of these algorithms is not the emphasis of this paper. Many of these algorithms have been implemented in the commercial-off-the-shelf (COTS) planning and scheduling tools. For example, Matlab's FMINCON [1] and ILOG's ILOG Scheduler [2] are two popular schemes that are used in a wide range of commercial applications. Some algorithms are a good match to certain types of problems while others are not. Conceptually, if computational resources are available, one can apply multiple constrained optimization algorithms to a given problem to ensure that at least one algorithm would converge and deliver a good plan. In this paper we apply the Matlab's FMINCON optimization function to generate interim plans in our mathematical formulation that lends itself to systematic risk analysis.

Another major result of this paper is to propose a planning methodology that iteratively applies constrained optimization and Monte Carlo simulation to reach an optimal plan – a plan that is optimized with respect to certain criteria (usually related to resource usage, including time), free of constraint violations, and with an acceptable P_F. Though this approach is applicable to many general planning and scheduling problems, the emphasis of this work is on mission planning and sequencing of spacecraft events.

This paper is organized as follows: Section 2 describes spacecraft planning and sequencing as an important application of event planning and scheduling. Section 3 outlines an iterative approach on event planning and risk analysis. Section 4 describes the mathematical formation and detailed resource and constraint models that constitute the constraint optimization process. Section 5 provides the Monte Carlo results. Section 6 gives an upper bound of P_F. Section 7 discusses the concluding remarks and future work.

2. MISSION PLANNING AND SEQUENCING

Mission planning and sequencing is a critical component for mission operations, it provides a mechanism for scientists and engineers to operate the S/C remotely from the ground. It translates the science intents and spacecraft health and safety requests from the users into mission plans and sequences. After a rigorous process validating the plan, it will be to the spacecraft for its execution. Usually it is a time consuming and costly operation. This paper will provide a methodology for formulating and optimizing both deterministic and non-deterministic sequence of events planning. A deterministic sequence is a set of spacecraft events in time ordered sequence, all events are associated with a pre-scheduled time for on-board execution. Whereas, a non-deterministic sequence should support activities that are event driven in several ways: an activity could be contingent upon the completion of other activities, it could be contingent upon the state of the spacecraft and/or estimated resources, or be triggered by real-time events such as the observation of a supernova explosion or precision landing.

During a mission operation phase, traditional mission planning and sequencing tools with a set of spacecraft, resource, and constraint models at various levels of fidelity are being used. Each mission plan or sequence consists of activities that are modeled as to their effects on key resources, often power, energy, telecommunication data rates from the spacecraft to earth, and data volume of onboard storage and downlink. Modeling a plan against the models of resources can confirm or deny the viability of the plan. Unviable ones can be changed to become viable, yielding limits on durations or other aspects of activities. Usually, a typical mission plan or sequence design and validation is done manually throughout a number of iterations until it becomes conflict free. There is a natural progression to migrate the current practice to the aforementioned risk-based event planning and scheduling approach. This approach allows mission planning and sequence generation first to define the planning and scheduling optimization criteria in objective function. For example, scientists would like to over-subscribe a plan with a prioritized activity list; the objective is to schedule as many activities as possible based on a well-defined priority scheme, within a planning horizon. Second, this approach allows user to select different types of rules and constraints to be checked for a given mission plan. Examples of some of the rules and constraints are forbidden synchronic, exclusion, inclusion, precedent relation, and resources rate, etc. For example, a typical mission plan or sequence's power consumption will be validated against how much power is available to the spacecraft, which can constrain the way the spacecraft is operated. Another example is a temporal constraint, such as the CPU must be turned on prior to image capture activity.

Thus far, we have illustrated the new methodology is capable of formulating a complex mission planning and sequencing problem, hence resolve any scheduling and resource conflicts, and construct an optimized conflict free plan. This approach is applicable for both deterministic and non-deterministic event planning. However, a non-deterministic event plan cannot be validated as a single conflict free plan. A stochastic approach to provide a risk analysis of a non-deterministic plan would be a more viable approach. This approach will allow user to schedule any mission plan with a presence of uncertainties. The system is capable of iteratively applying optimization followed by simulation, it tests a plan for its effectiveness of scheduling, and based upon the simulation results, the uncertainties will be quantified in term of its likelihood of success.

3. Interactive Approach On Event Planning and Risk Analysis

This Section presents an analysis approach that iteratively applies constrained optimization and Monte Carlo simulation to reach an optimal plan that is free of constraint violations with an acceptable P_F. This approach was first suggested by William Gearhart in an internal JPL study titled "Nondeter-Deterministic Sequence Validation and Verification."

Let us begin with some definitions. Consider n events of interest and the planning horizon $[T_s, T_e]$, and T_e-$T_s \equiv T$. Let event E_i be characterized by the ordered pair (t_o^i, d_i) for $1 \le i \le N$, where t_o^i and d_i are the start-time and duration of event E_i respectively. Let t_o^i be bounded by $[T_{\min}^i, T_{\max}^i]$, and let d_i have a probability distribution $p_i(d_i)$. For the sake of simplicity, let the event durations d_i's be independent of each other, and let $p_i(d_i)$ have a unimodal probability distribution function characterized by m_i and σ_i, where m_i and σ_i are respectively the mean and standard deviation of d_i. To support constrained optimization, we choose fixed event durations $\Delta_i = m_i + \lambda_i\sigma_i$ such that $d_i \le \Delta_i$ with a reasonably high degree of confidence.

If the events are all independent, the problem can be as simple as populating the timeline(s) with as many events as possible. But in most real world scenarios, many events of interest are dependent on each other in one or more ways. These dependencies can be expressed in the form of rules and constraints. Also, the planning horizon T is usually not long enough to accommodate all events. Thus a criterion of optimality is to fit as many high-priority events as possible into the planning horizon without violating any constraints. The dependencies between events, or constraints, can be rather complicated. The following are some examples of dependencies among events:

1. Time order: E_i must occur before E_j for some i and j.
2. Inclusion: If E_i occurs, E_j must occur for some i and j.
3. Exclusion: If E_i occurs, E_j must not occur for some i and j.
4. Forbidden Synchronic: If E_i and E_j occur, they must not overlap.
5. Priority: E_i is given a priority score Θ_i for all i.
6. Finite Resources: The onset of any event triggers the consumption of resources. The amount of resources used at any time by all events must not exceed the maximum allowable amount.

We will show in Section 4 that the aforementioned dependencies among events and criterion of optimality can be expressed mathematically in the form of constraint models and objective functions.

Given the above mathematical formulation, we define the following procedure that iteratively applies constrained optimization and Monte Carlo simulation to reach an optimal plan:

Step I: Constrained Optimization of Events Planning

For a given set of event durations Δ_I's and the aforementioned linear and nonlinear criteria and constraint models, we use commercial optimization engines to generate the set of start time t_o^i's that defines an optimal plan that is free of constraint violation.

Step II: Monte Carlo Simulation of P_F

Now given the optimal plan, we run Monte Carlo simulations of d_i's using the given probability distribution functions $p_i(d_i)$'s to determine the overall probability of successful completion of the plan within the given time horizon. If the probability is acceptable, then stop. If not, modify the optimization problem, typically by increasing the task durations Δ_i's, then go to Step 1.

Note that this plan is intentionally "sub-optimal". Once a plan is generated, the start times t_o^i's are fixed, regardless of how the events are executed. That is, the start time t_o^i of event E_i is not dependent upon the completion time of any prior events. This guarantees successful execution of the plan as long as $d_i \leq \Delta_i$ for all i. We also observe the following in our simulations:

- Empirically we can show that given fixed start times t_o^i's the Monte Carlo simulations converge quickly to the same result. That is, the simulation always yields the same probability that the plan does not execute successfully, and is independent of the seed that initiated the simulation.
- P_F is always "well-behaved". That is, increasing the task durations Δ_i's will always yield lesser events to be accommodated but higher probability of completion or vice versa.

4. CONSTRAINED OPTIMIZATION

This section formulates the event planning and scheduling process into a constrained optimization problem. In this paper, the authors choose to use the Matlab Optimization Toolbox routine FMINCON as the main computational engine for solving the constrained optimization problem. FMINCON implements the Sequential Quadratic Programming (SQP) method [4], and it finds a minimum of an objective function, subject to linear and non-linear constraints. Therefore, the objective function in this problem must be set up so that "bad" schedules produce large objective function values. The algorithm will then minimize the badness in the schedule, producing good schedules that meet the desired constraints. Note that the SQP method for solving constrained optimization problem can usually only find locally-optimal solutions, and the starting point for the method (also called the initial guess) is often critically important for determining the quality of the final solution. In fact, for a problem of this type, poor initial points often can result in failure of the solver to find *any* feasible optimum point at all.

In the following subsections, the authors will provide the mathematical expressions of a number of objective functions and constraints for event planning and scheduling. The authors would like to point out that there are usually multiple ways to express the objective functions and constraints. In this paper we choose to express the inequalities to meet as many sufficient conditions as we can to assure the convergence of the SQP method.

A. Objective Function

In seeking an optimal schedule, it is first necessary to define the criterion of optimality and set up an objective function to measure the merit of any given choice of schedule. Often in practice, a "good" schedule is one that accomplishes as much as possible, as soon as possible. The objective function chosen for this optimization problem is given by

$$f(t_0^1, \ldots, t_0^n) = \sum_{i=1}^{n} t_o^i . \quad (1)$$

Other optimization criteria can be a weighted sum of t_o^i or $\max_i(t_o^i)$, depending on the desired outcome of the problem.

B. Constraints

In nearly all scheduling problems (and other optimization problems) it is not an unconstrained optimum that is sought; rather, there are specific needs that must be met unconditionally, and these determine which points are feasible in the problem. Then, within the context of meeting those needs, the optimization selects the set of state variable values that achieves the lowest (or possibly highest) objective function value. The state variables in this problem are the start-times of the events, $t_0^1, \ldots, t_0^n$. In the case of the SQP algorithm, a constraint will be a function (say, $c(t_0^1,...,t_0^n)$) that, when satisfied, is of the form $c(t_0^1,...,t_0^n) \le 0$.

For the rest of this section, we provide the mathematical description of a number of common constraint types that are found in the SEQGEN[6] Adaptation Guide [5].

1) Linear Constraints

The time window constraint requires that event i to fall within a specified time frame, $\left[T_{\min}^i, T_{\max}^i\right]$, regardless of whether it occurs within the planning horizon or not. This means that $T_{\min}^i \le t_o^i \le T_{\max}^i \;\forall i$.

So if we define $X = [t_o^1, \cdots, t_o^n]^T$ (vector of state variable), $L = [T_{\min}^1, T_{\min}^2, \cdots T_{\min}^n]^T$,and $U = [T_{\max}^1, T_{\max}^2, \cdots T_{\max}^n]^T$, we have the following linear constraints:

$$X \ge L, \tag{2}$$

$$X \le U \tag{3}$$

2) Nonlinear Constraints

<u>Forbidden Synchronic:</u>

This constraint dictates the simple criterion that when two given events are both scheduled, they must not occur simultaneously at any point in time; that is, they must not overlap. Note that when the two events in question are not *both* scheduled, there cannot be a violation of this type.
To develop the equivalent mathematical characterization of this constraint, assume first that both events i and j are scheduled, and that they are not to overlap. If one formed $\max(t_o^i + \Delta_i, t_0^j + \Delta_j) - \min(t_o^i, t_0^j)$, for events with any overlap, this value would be smaller than the sum of the durations of the events, $\Delta_i + \Delta_j$. Therefore, to ensure that no overlap of events i and j occurs when they are both scheduled, one needs to enforce the opposite; that is,

$$\max(t_o^i + \Delta_i, t_0^j + \Delta_j) - \min(t_o^i, t_0^j) \ge \Delta_i + \Delta_j. \tag{4}$$

Setting this expression in the form "$\bullet \le 0$" for use in numerical algorithms requiring constraints of that form, one gets

$$\Delta_i + \Delta_j - \max(t_o^i + \Delta_i, t_0^j + \Delta_j) + \min(t_o^i, t_0^j) \le 0 \tag{5}$$

Since this constraint should only apply when both event i and event j are scheduled (a scheduled event *could* overlap with a non-scheduled one without a violation occurring), it is required that the constraint function be multiplied by the two binary functions representing that both i and j are scheduled. If both events are not present together, then the constraint function will reduce to $0 \le 0$, which is vacuously true. Thus, the final form of the constraint is given by

$$\begin{aligned}(t_o^i + \Delta_i \le h)(t_o^j + \Delta_j \le h) \times [\Delta_i + \Delta_j \\ - \max(t_o^i + \Delta_i, t_0^j + \Delta_j) + \min(t_o^i, t_0^j) \le 0].\end{aligned} \tag{6}$$

While this equation is the technically correct form of the constraint for forbidden synchronic, implementing this in a SQP solver proved problematic. In some cases, due to the "flat terrace" nature of the function, the optimization algorithm could not adequately manipulate events into a satisfactory schedule and would halt, reporting a failure.

To modify this constraint into an expression that allows better performance of a SQP solver, instead of using the length of the overlap as a measure of the violation when a violation occurs, one could set up a monotonically increasing linear function of the distance between the midpoints of the two events in question. As the events' midpoints draw closer together, the measure of violation increases, to a maximum when that distance is zero. This alternate version of the constraint is given by

$$\begin{aligned}(t_o^i + \Delta_i \le h)(t_o^j + \Delta_j \le h) \times [\Delta_i + \Delta_j \\ - \left|2(t_o^i - t_0^j) + \Delta_i - \Delta_j\right|] \le 0.\end{aligned} \tag{7}$$

which is a *sufficient* condition for meeting the requirement of no overlap between events i and j. When implemented in an SQP solver, this variant of the forbidden-synchronic constraint proved to be easier for the algorithm to satisfy.

<u>Inclusion:</u>

The inclusion constraint used in this problem dictates that *if* event i is scheduled, then event j must be initiated in some

[6] SEQGEN generates and validates command sequence. The commands sequence comprises of a time ordered commands that will be executed to achieve science objectives, maintain spacecraft health and safety, and establish communications between spacecraft and DSN stations.

chosen time interval, here denoted $[w_o^j, w_f^j]$. To develop the mathematical expression corresponding to this criterion, consider that when event j is scheduled, the relationship that expresses that the start time of the event occurs in the time interval $[w_o^j, w_f^j]$ is

$$\left| t_o^i - \frac{w_o^j + w_f^j}{2} \right| \le \frac{w_f^j - w_o^j}{2}. \qquad (8)$$

Multiplying through by 2 and rearranging terms yields

$$\left| 2t_o^j - w_o^j - w_f^j \right| + w_o^j - w_f^j \le 0. \qquad (9)$$

Since this must hold if event i is scheduled, but does not have to hold if event j is *not* scheduled, multiplying the constraint function by the binary value associated with whether or not event i is scheduled will produce the full inclusion constraint. Therefore, the final form of the inclusion constraint is given by

$$(t_o^i + \Delta_i \le h)\left(\left| 2t_o^j - w_o^j - w_f^j \right| + w_o^j - w_f^j\right) \le 0. \qquad (10)$$

Note that it is fairly simple to modify this development so that the inclusion constraint *includes* the entire duration of event j in some chosen time interval (provided it is at least as long as the event's duration.) That modification is not shown in this work, but simply involves adding an appropriate duration to the time interval of intended inclusion.

Exclusion:

The exclusion constraint used in this problem is similar to the inclusion constraint, but expresses the opposite criterion; it dictates that *if* event i is scheduled, then event j must *not* be initiated in some chosen time interval. Note that this implies that there exist two different modes by which such a constraint can be satisfied when event i is scheduled. Either both events i and j are scheduled and the appropriate exclusion-inequality holds true, or event i is scheduled and event j is not. As is the case with the forbidden synchronic constraint, this constraint will require two binary functions for relaxing the inequality expression when events i and j are not both scheduled.

To develop the mathematical formulation of this constraint, consider the simple inequality $|x| \ge a$ for a positive. This expression is equivalent to $x \ge a$ or $x \le -a$, which is to say that x is at least a units from 0. Using the same logic, to exclude the start-time of event j, t_o^j, from the time interval $[w_o^j, w_f^j]$, one must keep t_o^j at least half of the length of $[w_o^j, w_f^j]$ away from that interval's midpoint. Therefore, when events i and j are both scheduled,

$$\left| t_o^i - \frac{w_o^j + w_f^j}{2} \right| \ge \frac{w_f^j - w_o^j}{2}. \qquad (11)$$

must also hold. Multiplying through by 2 and rearranging terms, one gets

$$w_f^j - w_o^j - \left| 2t_o^j - w_o^j - w_f^j \right| \le 0. \qquad (12)$$

This condition needs only hold when both event i and event j are scheduled, so the inequality should be multiplied by the two binary functions associated with the existence of event i and event j within the planning horizon. In final form, this exclusion constraint is given by

$$(t_o^i + \Delta_i \le h)(t_o^j + \Delta_j \le h) \times \left(w_f^j - w_o^j - \left| 2t_o^j - w_o^j - w_f^j \right| \right) \le 0. \qquad (13)$$

Precedence Relations:

Precedence relations enforce ordering of events in the schedule and also have an underlying logical implication between events' existence in the schedule. To make this statement about logical implication clearer, consider that when one speaks of an event *preceding* another, it is assumed that the latter event actually does occur. In this scheduling problem, though, there is a difference between "locating an event in time" and "scheduling" it. As such, in setting up a precedence relation, one must carefully examine not just the start- and/or end-times of events, but also whether or not dependent events are actually completed within the planning horizon, and the exact logical nature of the desired dependence affects how that examination must be done.

The authors chose to enforce two different types of precedence relations in this problem, where each constraint expression acts on a single pair of events. Both of those relations will be derived in the next two subsections.

"Event i starts before event j finishes."

In the absence of a planning horizon, when all events are scheduled, this precedence relation would simply be expressed as $t_o^i \le t_o^j + \Delta_j$. To include the planning horizon as a driving factor in this problem, the It was chosen for this precedence relation that events i and j should only be *allied*, in the sense that only if both events are scheduled should the precedence hold. If they are not both scheduled, then there is nothing to enforce, and the relation should be vacuously true. In this case, neither event's occurrence in the schedule is dependent on the other event's occurrence.

Deriving this constraint is very simple; one only needs to multiplicatively combine the three elements just mentioned. The precedence relation to be enforced, rewritten in the "$\bullet \le 0$" form, is

$$t_o^i - t_o^j - \Delta_j \le 0. \tag{14}$$

Since it is required that the expression relax to a vacuously true statement when either event i or event j (or both) are not scheduled, multiplying the inequality by the two binary values associated with those events being scheduled or not will produce the needed expression. Therefore, the final form of this precedence relation is given by

$$(t_o^i + \Delta_i \le h)(t_o^j + \Delta_j \le h)\left(t_o^i - t_o^j - \Delta_j\right) \le 0. \tag{15}$$

"Event i finishes before event j starts."

This precedence relation, in the absence of a planning horizon, asserts that $t_o^i \le t_o^j + \Delta_i$. It was chosen (arbitrarily) to set up this relation to enforce that the events are *coupled*, so that if either event occurs in the schedule, the other must also. Note that this exact precedence relation would be the one required if event i were some mandatory precursor event for event j, and if event i had no purpose operating without event j in the schedule.

To derive the mathematical expression for this relation as a constraint, first imagine that both events are scheduled and the desired precedence holds. Then one has

$$t_o^i + \Delta_i \le h,\ t_o^j + \Delta_j \le h, \\ \text{and } t_o^i + \Delta_i \le t_o^j \tag{16}$$

all true. Rewriting the third inequality in the form expression must be modified.

"$\bullet \le h$" will be helpful, as then one has

$$t_o^i + \Delta_i \le h,\ \ t_o^j + \Delta_j \le h, \\ \text{and } t_o^i - t_o^j + \Delta_i + h \le h \tag{17}$$

Consider that, in a general setting, with numbers $a, b, \text{and } c,$ if one has

$$a \le h, b \le h, \text{and } c \le h \tag{18}$$

then it is the exactly equivalent statement to write

$$\max(a,b,c) \le h \tag{19}$$

Using this fact, and comparing it with the logical expression in (17), one can see that if either event i or event j is scheduled, then

$$\max(\ t_o^i + \Delta_i, t_o^j + \Delta_j, t_o^i - t_o^j + \Delta_i + h\ ) \le h$$
(20)

must hold, which is equivalent to

$$\max(t_o^i + \Delta_i, t_o^j + \Delta_j, \\ t_o^i - t_o^j + \Delta_i + h) - h \le 0 \tag{21}$$

The last step in the derivation of this constraint is to multiply by an appropriate binary-valued function so that the previous equation becomes vacuously true (i.e., $0 \le 0$) if and only if neither event i nor event j is scheduled; otherwise the binary function should return 1, so that the previous inequality must be held true non-trivially.

With a moment of reflection one can see that, although the following is not the only choice, the binary-valued expression given by

$$\max\left((t_o^i + \Delta_i \le h), (t_o^j + \Delta_j \le h)\right) \tag{22}$$

meets the above requirement. Multiplying the previous inequality by this expression, the final form of this constraint becomes

$$\max\left((t_o^i+\Delta_i \le h),(t_o^j+\Delta_j \le h)\right)$$
$$\times[\max(t_o^i+\Delta_i, t_o^j+\Delta_j,$$
$$t_o^i - t_o^j + \Delta_i + h) - h] \le 0 \tag{23}$$

<u>Resource Rate Constraint</u>

In nearly any scheduling problem involving events that consume resource, either the maximum instantaneous rate of resource consumption or the total amount of resource consumed, or both, will necessary to consider. In this problem, it was chosen to constrain the maximum allowable rate of resource consumption for any feasible schedule.

To set this criterion as a constraint, let $r_{\max}$ be the maximum rate of consumption allowed, and assume that event i consumes the resource at a constant rate, r_i, when operating. The total rate at which the schedule is consuming the resource at any time t is given by

$$R(t)=\sum_{i=1}^{n} r_i(t), \tag{24}$$

where

$$r_i(t)=\begin{cases} r_i & \text{if } \max(t, t_o^i - t + h, \; t - t_o^i - \Delta_i + h) \le h; \\ 0 & \textit{otherwise} \end{cases} \tag{25}$$

Note that the expression $\max(t, t_o^i - t + h, t - t_o^i - \Delta_i + h) < h$ is equivalent to

$$t_o^i \le t, t \le t_o^i + \Delta_i, \text{and } t \le h \tag{26}$$

which must hold if at time t the rate r_i is nonzero.

If violations of the resource rate constraint are to correspond to positive numbers, then a very sensible (and easy-to-code) way to define the constraint when only numerical computations will be performed is

$$\int_o^h (R(t)-r_{\max} > 0)[R(t)-r_{\max}]\Delta t = 0 \tag{27}$$

The expression $(R(t)-r_{\max} > 0)$ refers to the logical (binary) value of whether or not $R(t)-r_{\max} > 0$.

It should be clear, upon careful examination, that equation (27) sets violations of the resource rate constraint to be the total *amount* of resource consumed at the rate beyond the allowed maximum, over the duration for which that rate is positive.

For the problem at hand, the authors used a different method to handle the computation of the resource rate constraint. Explicit integration of a rate function was not used. Instead, a sorted vector of time-endpoints of all events was processed to identify the total rate of consumption at each subinterval, and comparisons of those totals to the allowed maximum enabled rate-violating subintervals to be identified. Multiplying the length of each rate-violating subinterval by the difference between the operating rate and the allowed maximum rate and adding those values produced the violation amount. This method is faster than most numerical quadrature schemes, and it gives the exact value of the integral in equation (27) when the consumption rate for each event is a constant.

C. Problem Summary

In summary, the constrained optimization problem that this paper is trying to solve is to find the state vector $X=[t_o^1, t_0^2, \cdots, t_o^n]^T$ that minimizes $\sum_{i=1}^{n} t_o^i$ and satisfies the following conditions:

$$\text{for all } i, -t_o^i \le 0,$$
$$(t_o^i+\Delta_i \le h)(t_o^j+\Delta_j \le h)\times[\Delta_i+\Delta_j - \left|2(t_o^i - t_0^j)+\Delta_i-\Delta_j\right|] \le 0,$$
$$(t_o^i+\Delta_i \le h)\left(\left|2t_o^j - w_o^j - w_f^j\right| + w_o^j - w_f^j\right) \le 0,$$
$$(t_o^i+\Delta_i \le h)(t_o^j+\Delta_j \le h) \times\left(w_f^j - w_o^j - \left|2t_o^j - w_o^j - w_f^j\right|\right) \le 0,$$
$$(t_o^i+\Delta_i \le h)(t_o^j+\Delta_j \le h)\left(t_o^i - t_o^j - \Delta_j\right) \le 0,$$
$$\max\left((t_o^i+\Delta_i \le h),(t_o^j+\Delta_j \le h)\right)[\max(t_o^i+\Delta_i, t_o^j+\Delta_j, t_o^i - t_o^j + \Delta_i + h) - h] \le 0,$$
$$\int_o^h (R(t)-r_{\max} > 0)[R(t)-r_{\max}]dt = 0$$

D. Example of Problem and Results

We construct an example of 30 events scheduled with two precedence relations, an exclusion relation, an inclusion relation, and a maximum rate resource constraint of 3 units. Each event consumes the resource at one unit per unit time. Figures 1 and 2 show the optimized schedule and the resource usage profile respectively.

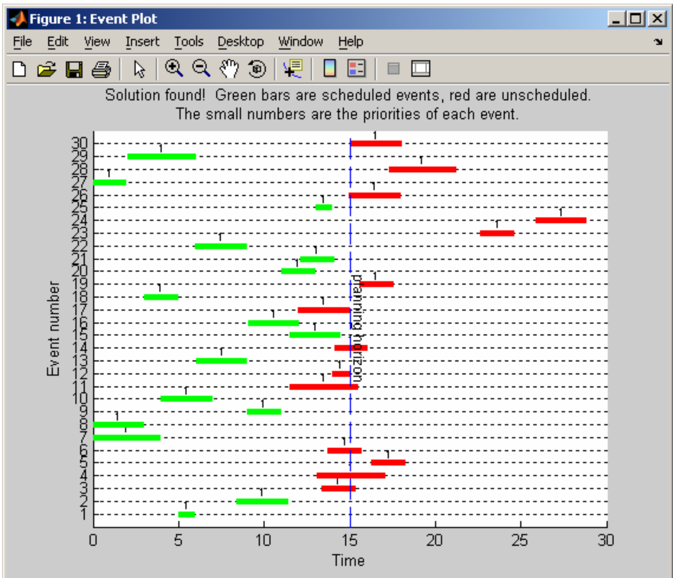

Figure 1

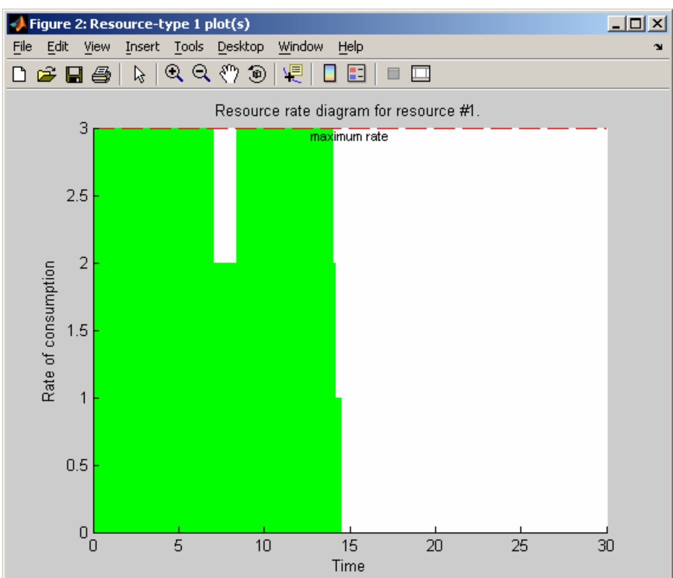

Figure 2

5. SIMULATION

When a schedule is optimized with the duration of each event fixed to a specified level of confidence, one can simulate the durations to obtain the probability that the given schedule failing due to constraint violations and events extending beyond the planning horizon. The distributions of the durations of events that were used to demonstrate the simulation techniques were the uniform, normal, log normal, beta, and triangular distributions. It is assumed that the random durations are independent of each other. We describe the simulations in details using an example of 5 events and an example of 10 events.

A. Example of a Simulation Problem of 5 Events

This example consists of a plan with five events that are scheduled optimally with the following constraints:

- Events 1 and 3 must not overlap
- Event 1 must finish before Event 4 begins,
- There is only one type of resource consumed, and all events consume that resource at a rate of one resource unit per time unit, with the maximum allowable consumption at any time to be 2 units at any time.
- The planning horizon must end at time unit equal to 26

Table 1 summarizes the probability distribution and its parameters of each of the 5 event durations.

Event ID	Type of Dist.	Parameters	Min. Value	Max Value
1	Uni.	NA	5	7
2	Beta	α=4, β=4	1	3
3	Norm	μ=10, σ=.5	NA	NA
4	Tri.	Peak=4	3	5
5	LogN	μ=2, σ=.5	NA	NA

Table 1

In this example each event has its duration fixed at the value such that each event has a 99% chance of taking that long or less to complete. Let $P_{S,i}$, which in this case is 0.99 for all i, denote that probability that event i would successfully end with the duration being a certain value or less. We apply FMINCON to the above problem to compute the optimal start times $t_0^1,\ldots,t_0^5$, and the results are given in Table 2. Figure 3 provides the timeline depiction of the optimized plan, and we can visually verify that there is no constraint violation.

Event ID	Initial Time
1	0
2	0
3	6.98
4	19.14
5	1.86

Table 2

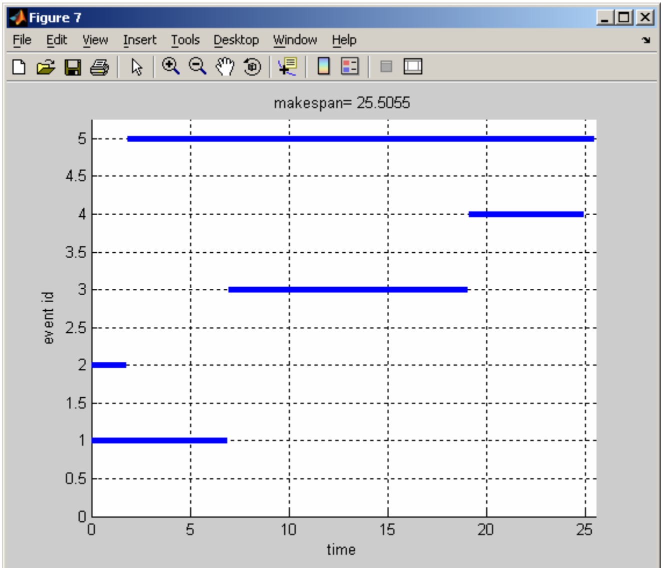

Figure 3

Based on the optimized plan of 5 events as described above, we perform 10 Monte Carlo simulations of 5000 runs each to estimate P_F. The results are tabulated in Table 4. In this simple case of 5 events, we can enumerate all possible scenarios of event violation and analytically compute P_F based on the individual $P_{S,i}$, $1 \le i \le 5$. As shown in Table 4, the simulation result matches closely with the enumerated result. In a typical scheduling problem for mission operation, there are typically many more events. To enumerate all possible scenarios of event violation to analytically compute P_F would be impractical, and it will need to resort to simulation to compute P_F.

Simulation ID	P_F of the 5 Event Plan
1	0.0292
2	0.0304
3	0.0266
4	0.0272
5	0.0278
6	0.0296
7	0.0256
8	0.0282
9	0.0296
10	0.0292
Average P_F	0.0283
Analytical P_F	0.0288
Upper bound of P_F	0.05

Table 4

Example of a Simulation Problem of 10 Events

Next we describe a larger example of a 10-event case with the following constraints:

- Events 1 and 3 may not overlap
- Event 1 must finish before Event 4 begins,
- There is only one type of resource consumed, and all events consume that resource at a rate of one resource unit per time unit, with the maximum allowable consumption at any time to be 3 units at any time.
- The planning horizon ends at time unit equal to 26

Table 5 summarizes the probability distribution and its parameters of each of the 10 event durations.

Event ID	Type of Dist.	Parameters	Min. Value	Max Value
1	Uni.	NA	5	7
2	Beta	α=4, β=4	1	3
3	Norm	μ=10, σ=.5	NA	NA
4	Tri.	Peak=4	3	5
5	LogN	μ=2, σ=.5	NA	NA
6	Uni.	NA	2	5
7	Beta	α=5, β=5	3	8
8	Uni.	NA	1	3
9	Tri.	Peak=3	2	5
10	Tri.	Peak=4	2	6

Table 5

As in the 5-event case, events in this example are scheduled optimally, with duration of each event fixed at the value where each event has a 99% chance of taking that long or less to complete. Again we apply FMINCON to optimize the above plan of 10 events and Figure 4 illustrates the resulting timeline.

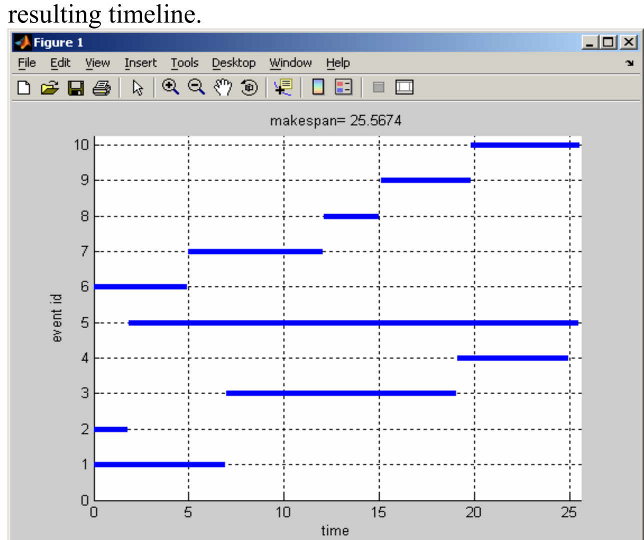

Figure 4

Again we perform 10 Monte Carlo simulations of 5000 runs each to estimate P_F. The results are tabulated in Table 5 below.

Simulation ID	Probability of Schedule (10 Events) Failing
1	0.0424
2	0.0430
3	0.0458
4	0.0448
5	0.0382
6	0.0372
7	0.0358
8	0.0434
9	0.0400
10	0.0430
Average P_F	0.0414
Upper Bound of P_F	0.10

Table 6

6. A Simple Upper Bound of P_F

As shown in Section V, the simulation of P_F can be quite tedious, and it is a function of the optimized conflict-free plan. In this section, we derive an upper bound of P_F that is expressed as a function of $P_{F,i,}\ 1 \le i \le n,$ where $P_{F,i}$=1-$P_{S,i}$ denote the probability that event i would end with a duration of d_i that exceeds the predetermined duration Δ_i as described in Section III. Let us also denote P_S be the probability that the schedule succeeds, meaning it does not violate constraints nor exceeds the planning horizon. It is obvious that

$$P_S \ge P_{S,1}xP_{S,2}...xP_{S,n}, \tag{28}$$

because the term $P_{S,1}xP_{S,2}...xP_{S,n}$, does not take into account all the possible ways in which events may exceed the designated durations determined by $P_{S,i}$, and still have a successful schedule. Thus we have

$$P_F = 1 - P_S \le 1 - P_{S,1}x...xP_{S,n} \le 1 - (1 - P_{F,1}) \cdots (1 - P_{F,n}),$$

which results in an upper bound of P_F given by

$$P_F \le P_{F,1} + P_{F,2} + \cdots P_{F,n} \tag{29}$$

Thus this simple upper bound of P_F does not require tedious simulations as shown in (29), and is just the sum of all $P_{S,i,}\ 1 \le i \le n$. Also this upper bound is independent of the optimization algorithm used to generate the plan.

The upper bounds of P_F for the 5-event case and 10-event case in Section V are 0.05 and 0.1 respectively.

7. Concluding Remarks and Future Work

In this paper, we discuss the problem of planning and scheduling of events with non-deterministic durations. We introduce the probabilistic description of event durations, and derive an analysis approach that quantifies the tradeoff between planning risk and planning efficiency. Also we demonstrate that the general criteria of optimality and rules and constraints for event planning can be described mathematically in terms of linear and non-linear functions and inequalities. This allows the use of customized and COTS constraint optimization algorithms to generate conflict-free plans.

We would like to point out that the aforementioned risk analysis approach in event planning is similar in some respects to telecommunication link analysis. The standard link analysis is a proven statistical estimation technique for evaluating communication system performance and trade-off. Link analysis uses the Design Control Table (DCT) that consists of the calculation and tabulation of the useful signal power and the interfering noise power available at the receiver. Many of the gain and loss parameters that comprise the link are statistical. Each of the statistical link parameter x can be described in terms of a design value x_d, a minimum value x_{min}, a maximum value x_{max}, and a probability distribution function *f(x)* such that that $f(x) \neq 0$ for $x_{min} \le x \le x_{max}$, and $f(x) = 0$ for $x < x_{min}$ and $x > x_{max}$. Some common form of *f(x)* are the uniform, triangular, and truncated Gaussian distributions. From this setup, one can deduce the mean of x (denoted by x_m) and the variance of x (denoted by x_{var}). In link analysis when a large number of independent link parameters are added together (in decibels), one can use a "hand-waving" argument of the Law of Large number to deduce that the net contribution, which is usually expressed in terms of signal-to-noise ratio, has a Gaussian distribution $N(m,s^2)$ where m is the mean and s is the standard deviation (a.k.a. sigma). From this one can design a link and establish link margin policy based on statistical confidence level measured in terms of sigma (e.g. 2-sigma event, 3-sigma event etc.). This analysis is not mathematically rigorous, but decades of experience show that this approach works well to characterize the communication system performance, to support trade-off, and to manage the operation risks associate with the link usage.

In closing, we outline a number of interesting problems that follow up on the results presented in this paper:

- The planning and scheduling approach portrayed in this paper has the inherit efficiency that once a plan is generated, the start times t_o^i's are fixed regardless of how the events are executed. It is possible that one or more events might finish early, and leave a gap before the subsequent events start, thus causing idle time. So are there inline algorithms or heuristics that will check the dependencies among events and kick off the subsequent constraint-free events at earlier times to improve efficiency?

- Given P_F as the requirement, how can one establish robust margin strategies for resource allocation and planning horizon to ensure that the plan would execute successfully?
- How can one perform sensitivity analysis for non-deterministic event planning?
- Can we apply these techniques to perform onboard risk assessment to guide decision making in spacecraft autonomy?

REFERENCES

[1] Optimization Toolbox, Mathworks, http://www.mathworks.com/products/new_products/

[2] ILOG Scheduler, ILOG Inc., http://www.ilog.com/products/scheduler/

[3] William B. Gearhart, "Non-Deterministic Sequence Validation and Verification," JPL Internal Report, June 10, 2004.

[4] Han, S.P., A Globally Convergent Method for Nonlinear Programming, Journal of Optimization Theory and Applications, Vol. 22, No. 3, pp. 297-309, 1977.

[5] Dennis Page, SEQGEN Adaptation Guide, JPL D-28688, May 26, 2004.

ACKNOWLEDGEMENT

This research was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

BIOGRAPHY

Dr. Kar-Ming Cheung *is a Principal Engineer and a Technical Group Supervisor in the Communications Systems Research Section (331) at JPL. His group provides telecom analysis support for JPL missions, and develops the operational telecom analysis and predict generation tools for current and future JPL missions and the DSN. He received NASA's Exceptional Service Medal for his work on Galileo's onboard image compression scheme. He was the Manager of the Network Signal Processing Work Area of NASA's Deep Space Network (DSN) Technology Program. He has authored or co-authored 6 journal papers and over 20 conference papers in the areas of error-correction coding, data compression, image processing, and telecom system operations. Since 1987 he has been with JPL where he is involved in research, development, production, operation, and management of advanced channel coding, source coding, synchronization, image restoration, and link analysis schemes. He got his B.S.E.E. degree from the University of Michigan, Ann Arbor in 1984, his M.S. degree and Ph.D. degree from California Institute of Technology in 1985 and 1987 respectively.*

Mr. Adans Y. Ko *is a Software Systems Design Engineer and Technical Group Supervisor in Section (314) at JPL. He is responsible for the Multi-mission Ground Data System tools, which includes mission planning, sequence generation, and S/C telecom analysis for current and future JPL missions and the DSN. He received NASA's Exceptional Service Medal for his work on Voyager's onboard Computer Command Subsystem for mission to Uranus and Neptune. He was the Development Manager of the Mission Planning and Sequence Subsystem. He got his B.S.C.S. degree from the Utah State University, Logan, Utah. in 1982, his M.B.A. degree from University of California, Los Angeles in 1993.*

Van Dang *is an Associate Engineer in the Avionics Research Technology and Development Group at JPL. Her main area of study are in Numerical Analysis, Optimization, and Statistics. She received her B.A. in Mathematics from Scripps College and her M.S. in Applied Mathematics from Claremont Graduate University.*

David Heckman *holds both a B.S. and an M.A. in mathematics, and he is currently enrolled in the mathematics Ph.D. program at Claremont Graduate University. His mathematical interests include optimization, numerical analysis, stochastic processes, and classical mechanics problems. He finds delight in solving problems involving objects in motion, optimal trajectories, optimal shapes, and computer simulations of dynamic processes.*