Risk Analysis for Resource Planning Optimization

Kar-Ming Chueng
Jet Propulsion Laboratory, California Institute of Technology

9th International Probabilistic Safety Assessment and Management Conference

May 18-23, 2008

What this paper is NOT about

- NOT about anatomy of planning and optimization algorithms
  - But to formulate a risk analysis and planning framework that plugs in different planning and optimization schemes like FMINCON, ILOG, and GA
- NOT about generation of an “optimal” plan
  - But to provide a “near-optimal plan” of non-deterministic events whose probability of failure P_F can be quantified analytically and by simulation
- NOT about tedious mathematical derivations
  - But to demonstrate that non-deterministic events and their relationships (constraints) can be mathematically modeled, and lend itself to mathematical optimization and empirical simulation

Main goals

- The main purpose of this paper is to introduce a risk management approach that allows planners to quantify the risk and efficiency tradeoff in the presence of uncertainties, and to make forward-looking choices in the development and execution of the plan
- Demonstrate a planning and risk analysis framework that tightly integrates mathematical optimization, empirical simulation, and theoretical analysis techniques to solve complex problems

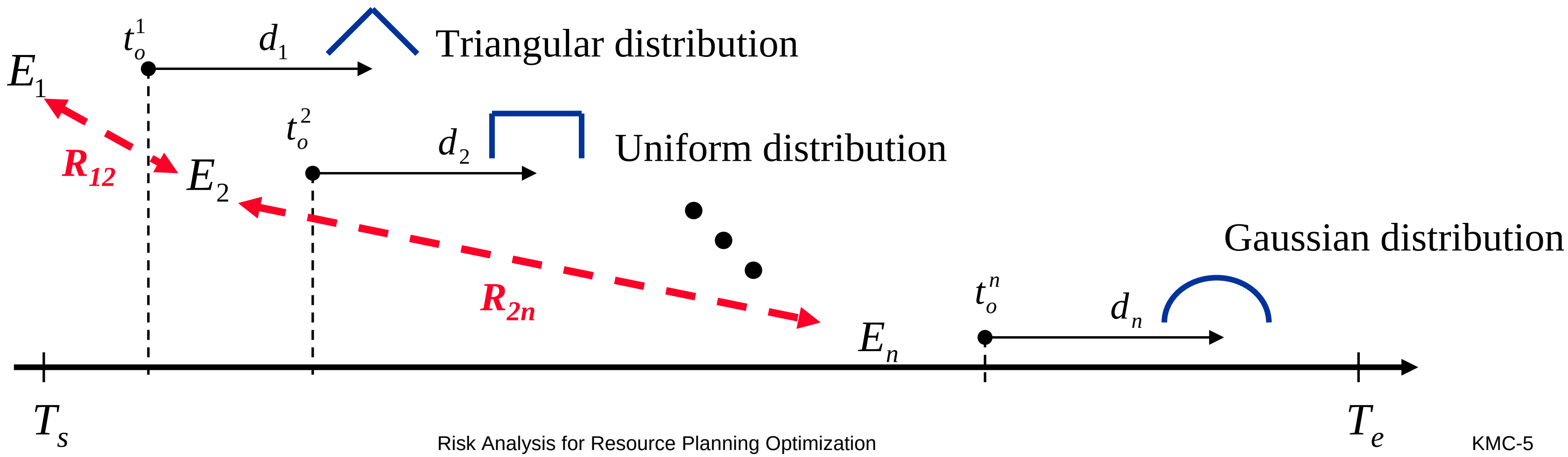
Problem statement (1)

- Extending link analysis techniques to resource planning optimization in the presence of uncertainties
  - Standard link analysis is a proven statistical risk analysis technique for evaluating communication system performance and trade-off
  - Many of the gain/loss parameters (in dB's) of the link are statistical
    - Parameter x with designed value x_d, minimum value, x_{min}, maximum value x_{max}, and a probability function $f(x)$, result in x_{mean} and x_{var}
  - With the 'hand-waving' assumption that the sum of all gain/loss link parameters has a Gaussian distribution with distribution $N(m,\sigma^2)$, one can design a link and establish link margin policy based on statistical confidence level measured in terms of σ (i.e. n-sigma event)

  - Non-deterministic events has variable time durations
  - Extend the link performance analysis (in dB's) to non-deterministic event planning (in time)

Problem statement (2)

- Some notations
  - Planning horizon $[T_s, T_e]$: given start time T_s, given end time T_e, all events must fit within $[T_s, T_e]$
  - Event E_i: start time t_o^i, duration d_i, where t_o^i is the state variables to optimize, and d_i is a random variable that has a unimodal probability distribution function $p_i(d_i)$ with mean m_i and standard deviation σ_i
  - A plan consists of a number of events within the planning horizon, and events E_i and E_j might bear certain pair-wise relationship R_{ij}
  - There are one or more resource limits that cannot be exceeded



Problem statement (3)

- Some definitions of terms
  - Planning is the process of a priori scheduling the events within the planning horizon
  - There are one or more objective functions that the plan is trying to optimize subjected to the given rules and constraints
  - A plan is said to be successfully executed if
    - All events in the plan can be accommodate within the planning horizon
    - There is no resource usage that exceeds the maximum allowable limit
    - There is no violation to the set of pre-defined rules and constraints

Applications (1)

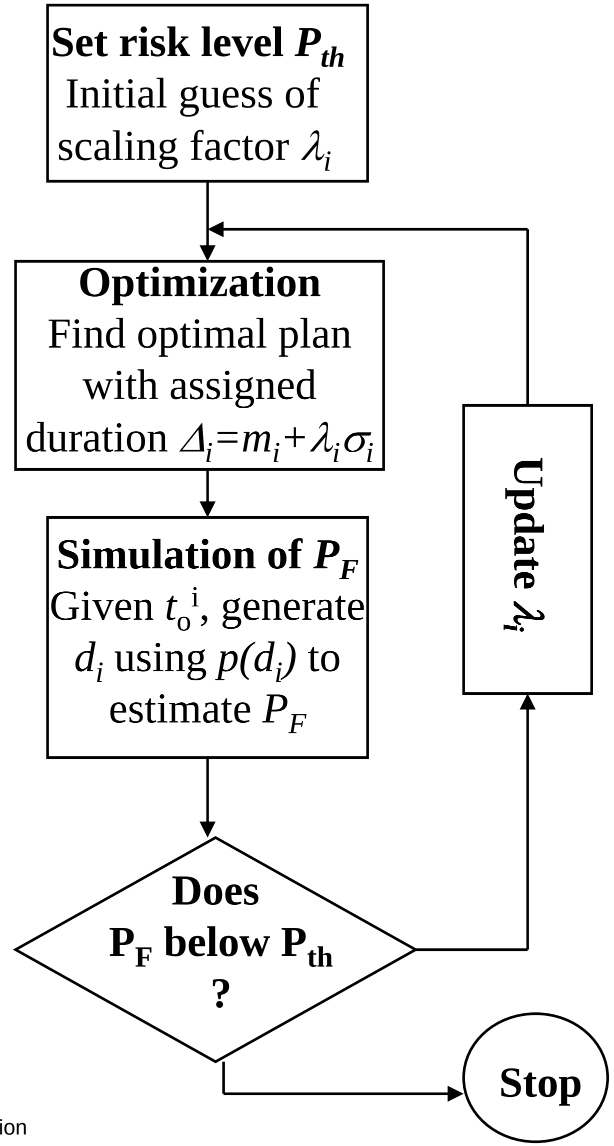
- Space mission planning and sequencing
  - Mission planning/sequencing translates science intents and spacecraft health and safety requests from the users into activities in the mission plan
  - Non-deterministic spacecraft events: star-tracker to acquire a star, data volume per pass, slew, … etc.
  - Spacecraft resources: power/energy, data rate/data volume, thermal limits, onboard storage, CPU etc.
  - Event-driven spacecraft activities: an activity could be contingent upon the complete of other activities, upon the state of the spacecraft and/or estimated resources, or triggered by real-time events such as observation of a supernova explosion

Applications (2)

- Risk analysis for cost and schedule planning
  - Model budget (resource) and schedule (duration) and their uncertainties
  - Model tasks dependencies
- Risk analysis for communication network planning
  - Model link durations and their uncertainties
    - Time uncertainty to transmit a certain fix data volume in the presence of retransmission (e.g. Prox-1)
  - Model link availabilities as resources
    - Number of users in a multiple access scheme
    - Data rates
  - Model link dependencies
    - Store-and-forward relay link: forbidden synchronic
    - Bent-pipe relay link: inclusion

Risk analysis approach by iterative simulation and optimization

- Given an acceptable risk level P_{th}, find a plan with $P_F \leqslant P_{th}$ by iterative optimization and simulation
- Plan is intentionally sub-optimal to ensure a stable solution
  - Start time $t_o^{\,i}$ is not dependent upon the completion time of any prior events
  - Ensure successful execution of plan as long as $d_i \leqslant \Delta_i$
- Simulation always converge
- P_F is always "well-behaved", i.e. increasing the task duration Δ_i will always yield lesser events to be accommodated but higher probability or completion or vice versa



Mathematical representation of non-deterministic events and constraints (1)

- Examples of objective Functions
  - Given start times $t_o^1, t_o^2, \ldots t_o^n$ (state variables to optimize)

$$f_1(t_0^1,\ldots,t_0^n) = \max_i\{t_o^i + d_i\}$$

$$f_2(t_0^1,\ldots,t_0^n) = \sum_{i=1}^{n} t_o^i$$

$$f_3(t_0^1,\ldots,t_0^n) = \sum_{i=1}^{n} t_o^i + d_i$$

f_1 : Minimizing maximum end time

f_2 : Minimizing initial time occurrence of all events

f_3 : Minimizing end time of all events

f_n : Priority weighted versions of the above

Mathematical representation of non-deterministic events and constraints (2)

- Example of linear constraints
  - Ranges of start time t_o^i

$$\overline{x} = \begin{bmatrix} t_o^1 \\ \mathrm{M} \\ t_0^n \end{bmatrix}_{nx\ 1} \qquad \overline{lb} = \begin{bmatrix} T_{\min}^1 \\ \mathrm{M} \\ T_{\min}^n \end{bmatrix}_{nx\ 1} \qquad \overline{ub} = \begin{bmatrix} T_{\max}^1 \\ \mathrm{M} \\ T_{\max}^n \end{bmatrix}_{nx\ 1}$$

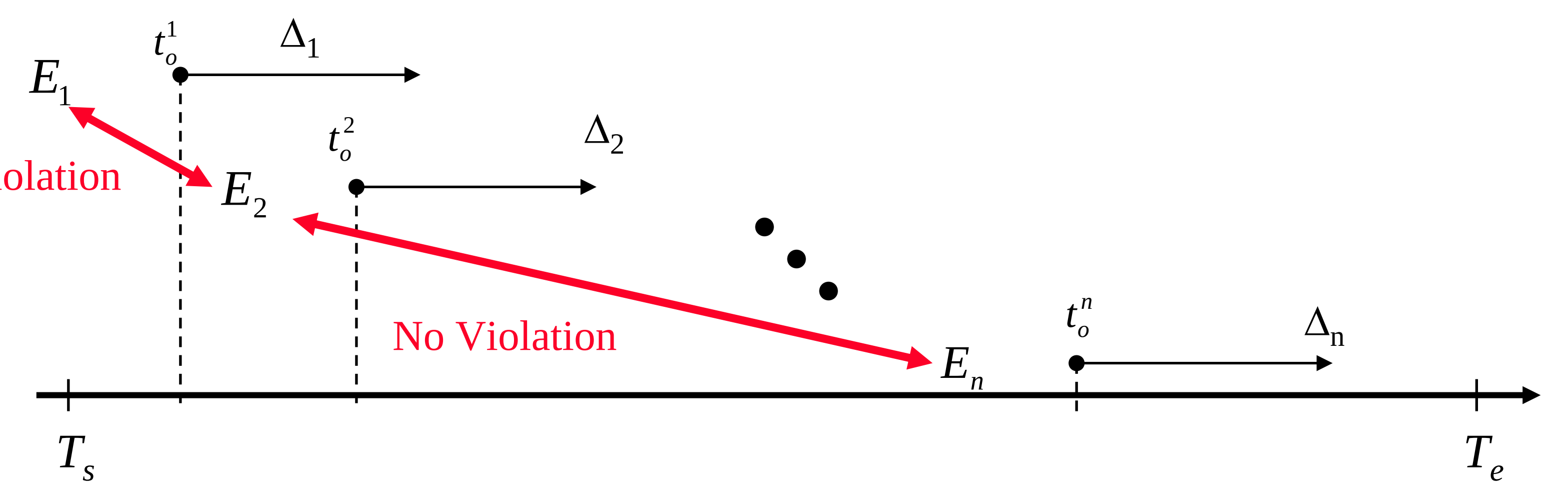
$$\overline{lb} \le \overline{x} \le \overline{ub} \rightarrow \begin{bmatrix} T_{\min}^1 \\ \mathrm{M} \\ T_{\min}^n \end{bmatrix} \le \begin{bmatrix} t_o^1 \\ \mathrm{M} \\ t_0^n \end{bmatrix} \le \begin{bmatrix} T_{\max}^1 \\ \mathrm{M} \\ T_{\max}^n \end{bmatrix}$$

Mathematical representation of non-deterministic events and constraints (3)

- An example of non-linear constraints (with explanation)
  - Forbidden synchronic: when two given events are both scheduled, they must not occur simultaneously at any point in time

Direct form: $\max(t_o^i + \Delta_i, t_0^j + \Delta_j) - \min(t_o^i, t_0^j) \geq \Delta_i + \Delta_j$

Alternate form: $[\Delta_i + \Delta_j - \left|2(t_o^i - t_0^j) + \Delta_i - \Delta_j\right|] \leq 0$



Mathematical representation of non-deterministic events and constraints (4)

- Other examples of non-linear constraints (with no explanation)
  - Inclusion: if event *i* is scheduled, then event *j* must be initiated in some chosen time interval $[w_o^j, w_f^j]$

$$\left(\left|2t_o^j - w_o^j - w_f^j\right| + w_o^j - w_f^j\right) \leq 0$$

  - Exclusion: if event *i* is scheduled, then event *j* must not be initiated in some chosen time interval $[w_o^j, w_f^j]$

$$\left(w_f^j - w_o^j - \left|2t_o^j - w_o^j - w_f^j\right|\right) \leq 0$$

  - Others: precedence relationships, resource constraints, etc.

Empirical results and theoretical results (1)

- Theoretical result: a simple upper bound of P_F
  - Denote $P_{F,i}$ the probability that event i would end with a duration d_i that exceeds the predetermined duration Δ_i, and $P_{S,i} = 1 - P_{F,i}$
  - Denote P_S the probability that the schedule succeeds, meaning it does not violate constraints nor exceeds the planning horizon; it is obvious that $P_S \geqslant P_{S,1} P_{S,2} \ldots P_{S,n}$, because $P_{S,1} P_{S,2} \ldots P_{S,n}$ does not take into account all the possible ways in which event may exceed the designated durations determined by $P_{S,I}$, and still have a successful schedule
  - Therefore

$$P_F = 1 - P_S \leq 1 - P_{S,1} x \ldots x P_{S,n} \leq 1 - (1 - P_{F,1}) \cdots (1 - P_{F,n})$$

  - Which results in an upper bound of PF given by

$$P_F \leq P_{F,1} + P_{F,2} + \cdots P_{F,n}$$

- The upper bound of P_F can be used to guide the adjustment of λ_i in the iterative optimization/simulation process

Empirical results and theoretical results (2)

- Theoretical result: Saddle-Point approximation of P'_F of an Ensemble of Tasks in Tandem
  - In task planning, a common situation is that there are a number of tasks that are required to execute in tandem, sometime with a constraint on overall duration
  - If no dependencies between these tasks with other tasks, one can treat them as a single task to simplify downstream analysis and optimization
  - The probability that the total duration of tasks exceed α, $P'_F(z > \alpha)$, can be approximated by

$$\boxed{q_+(\alpha) \approx \frac{e^{\psi(s_0)}}{\sqrt{2\pi\psi''(s_0)}}}$$

  - See next chart for outline of derivation

Empirical results and theoretical results (3)

- Some notations
  - $x_1, x_2, \ldots x_n$ are n independent random variable with pdf $f_{xi}(x_i)$
  - z is the sum of $x_1, x_2, \ldots x_n$
  - $\Psi_{xi}(s)$ is the characteristic function of x_i, and $\Psi_z(s)$ is the characteristic function of z
  - $q_+(\alpha)$ is the tail probability of z

$$\psi(s) = -s\alpha + Log\ \Psi_z(s) - Log\ s$$

$$z = \sum_{i=1}^{n} x_i \qquad f_z(z) = f_{x_1}(x_1) * f_{x_2}(x_2) * \ldots * f_{x_n}(x_n)$$

$$\Psi_z(s) = \int_{-\infty}^{\infty} e^{sz} f_z(z)dz \qquad \Psi_z(s) = \Psi_{x_1}(s)\Psi_{x_2}(s) \ldots \Psi_{x_n}(s)$$

$$q_+(\alpha) = \int_{\alpha}^{\infty} f_z(z)dz \qquad \boxed{q_+(\alpha) \approx \frac{e^{\psi(s_0)}}{\sqrt{2\pi\psi''(s_0)}}}$$

- Analysis challenges
  - Evaluation of pdf of sum of n variables requires $n-1$ nested integration
  - Inverse of $\Psi_z(s)$ is usually extremely difficult, if not impossible

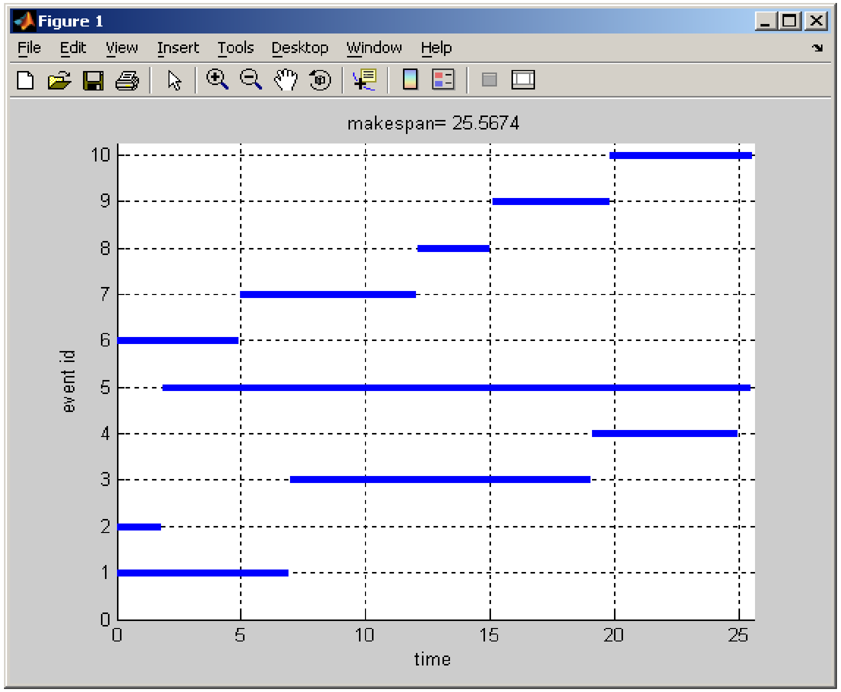
Empirical results and theoretical results (4)

- 10-event case toy problem
  - Events 1 and 3 may not overlap
  - Event 1 must finish before event 4 begins
  - Each event consumes 1 unit of resource, limit 3 at any time
  - PDF and its parameters of each of the ten 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

Empirical results and theoretical results (5)

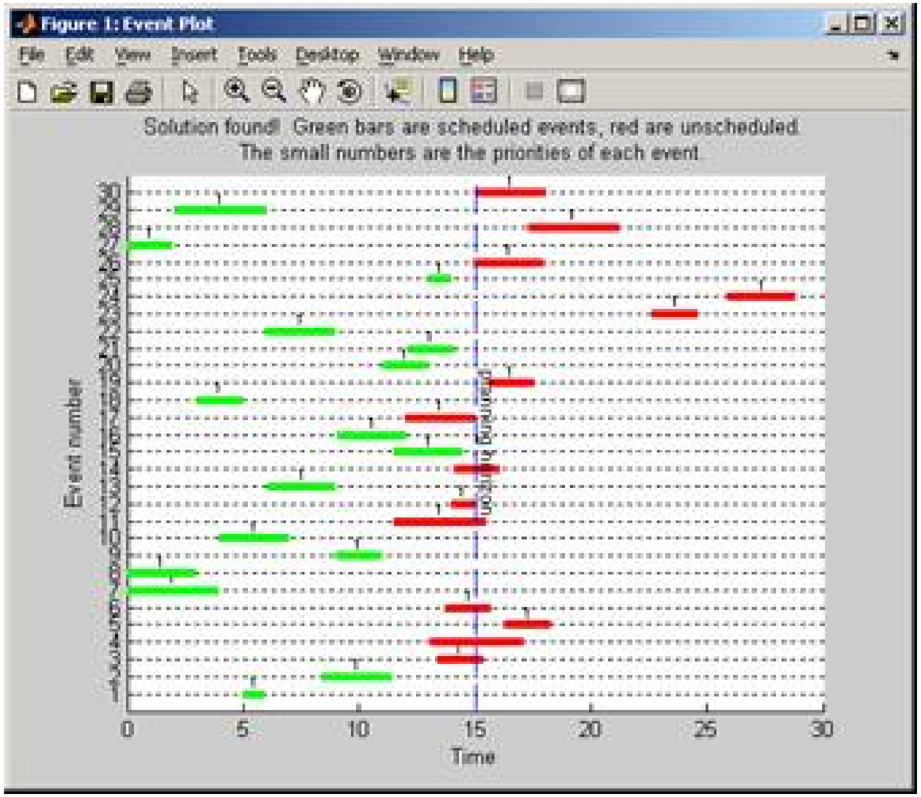
- 10-event case optimization and simulation results
  - Set durations Δ_i such that each event has a 99% confidence of successful completion

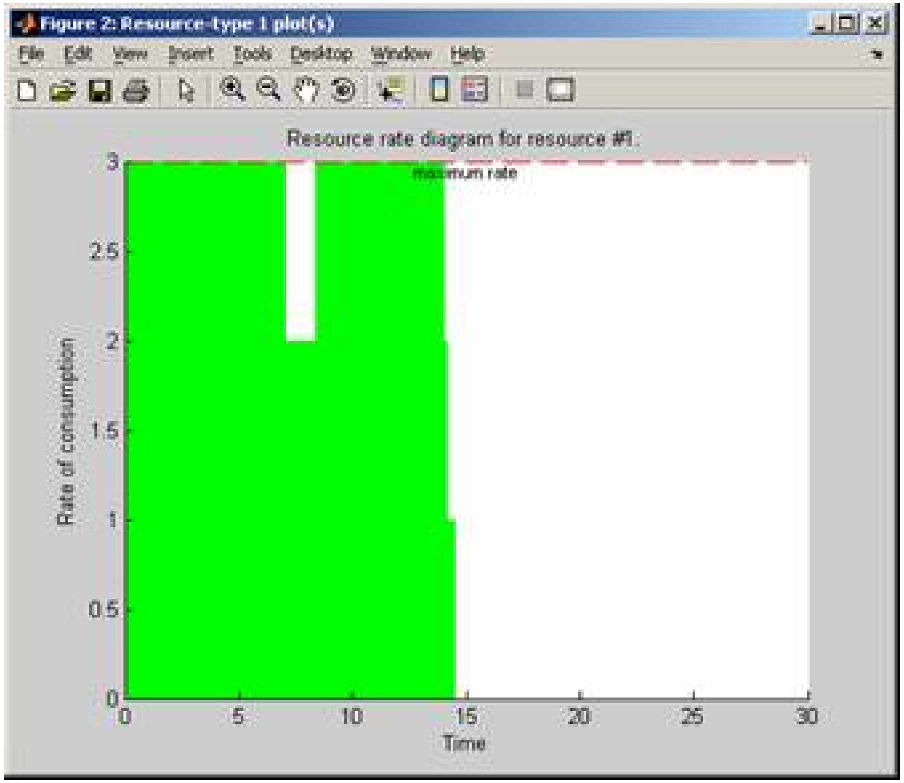


Simulation ID	Probability of Schedule (10 Events) Failing (5000 runs)
1	0.0424
2	0.0430
3	0.0458
4	0.0448
5	0.0382
6	0.0372
7	0.0358
8	0.0434
9	0.0400
10	0.0430
Ave. P_F	0.0414
Upper Bound of P_F	0.10

Empirical results and theoretical results (6)

- 30-event case
  - 2 precedence relations, 1 exclusion relation, 1 resource limit of 3

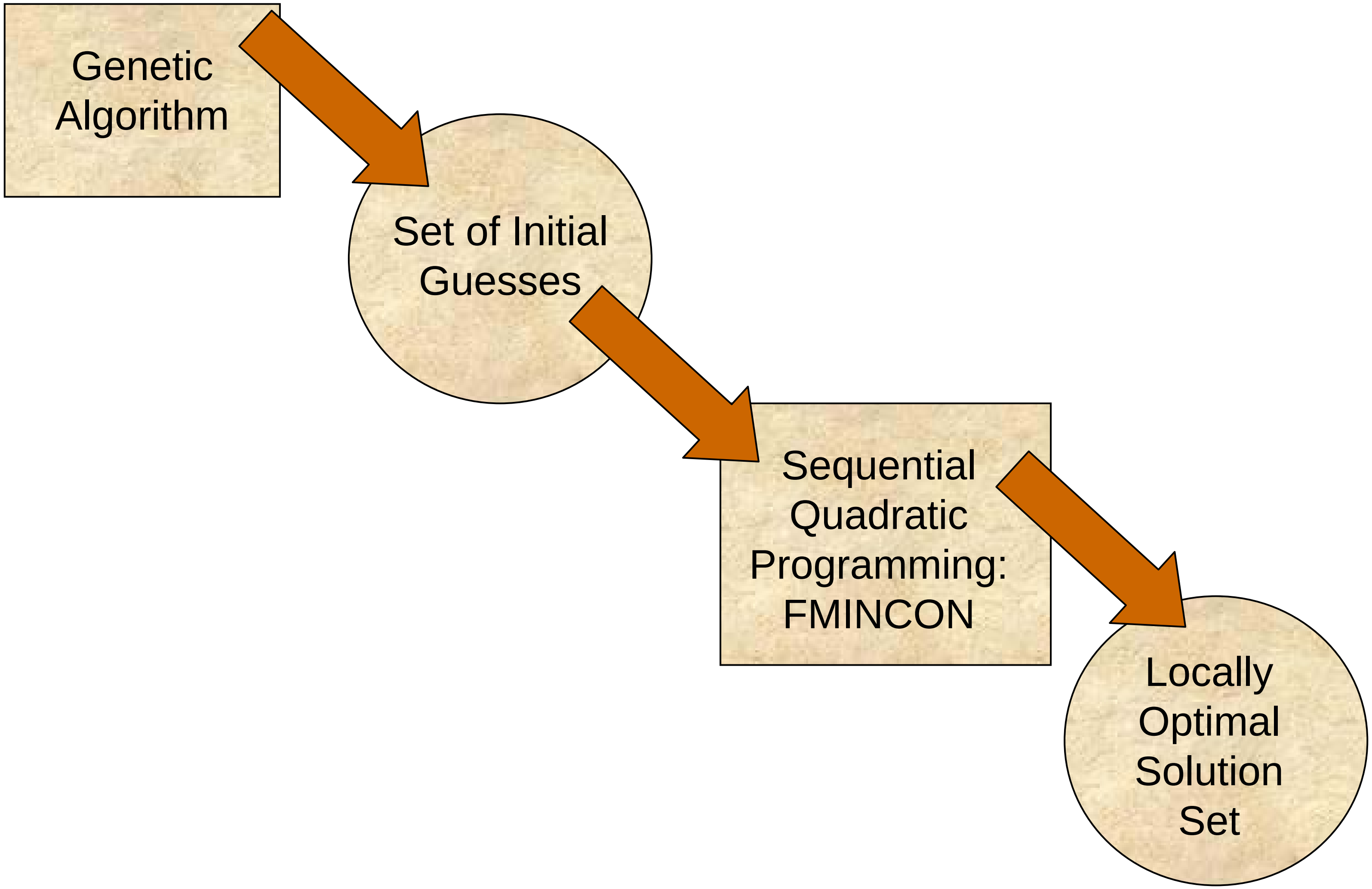




Using Stochastic Optimization to Find a Good Initial Point (1)

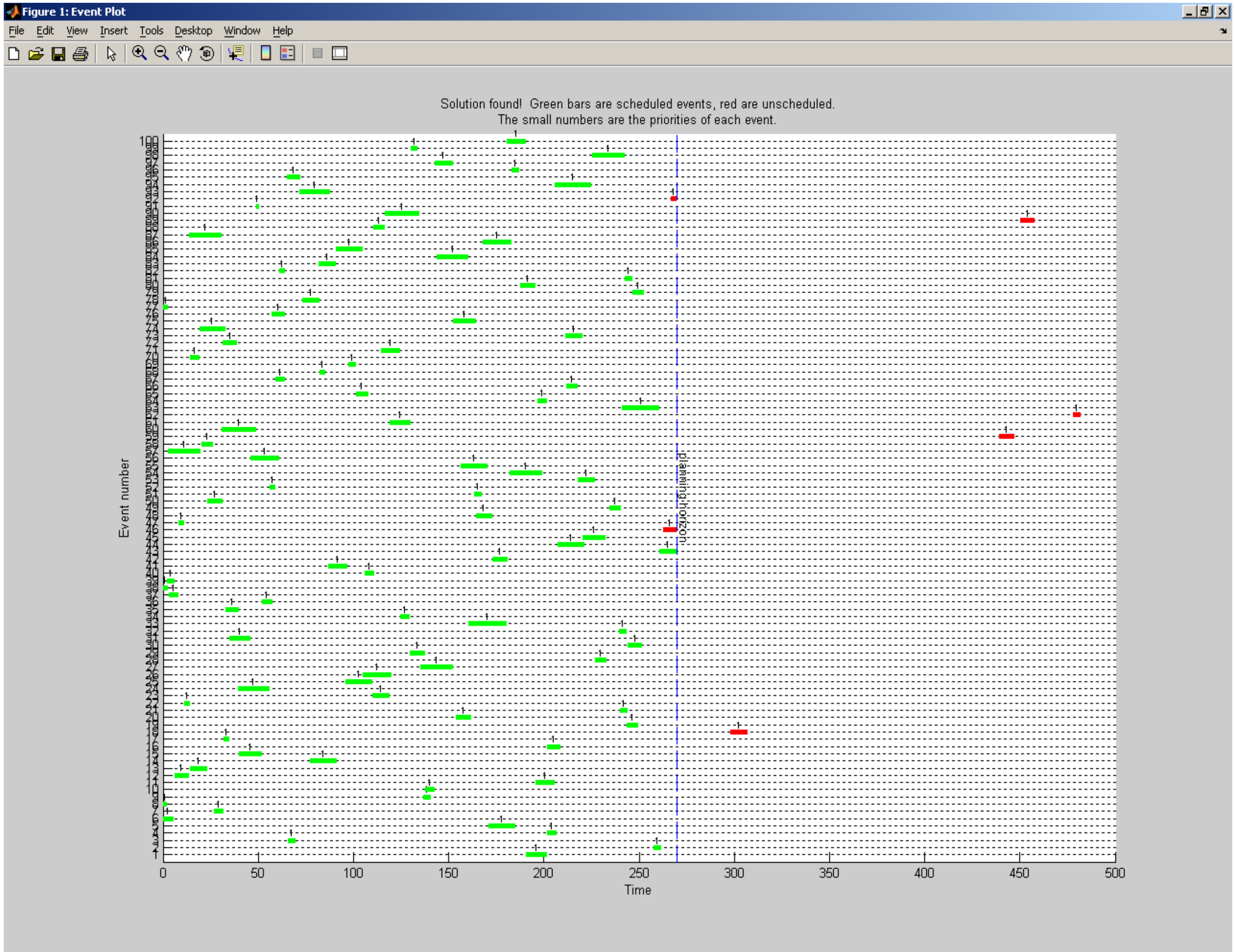
- Challenges of optimization
  - Speed and optimization performance depends strongly on the initial guess of the state vector $[t_o^1, t_o^2, \ldots t_o^n]^T$
  - A bad guess results in slow convergence and/or poor locally-optimal solution
- Improved optimization using stochastic optimization algorithm
  - Use stochastic optimization algorithm (e.g. genetic algorithm) to find a set of viable and promising state vectors to serve as initial guesses
  - Use the initial guesses as input to more sophisticated optimization schemes (e.g. Sequential Quadratic Programming in Matlab's FMINCON) to generate a set of locally-optimal solutions
  - Obtain an "overall" optimal solution out of all the local optima by subjecting them to a probabilistic simulation to determine likelihood of failure and to compare objective values

Using Stochastic Optimization to Find a Good Initial Point (1)



Using Stochastic Optimization to Find a Good Initial Point (2)

- 100 event case, 40 constraints, 2 resources limit of 4



Using Stochastic Optimization to Find a Good Initial Point (3)

- Resource#1 usage profile of 100 event case