CALCULUS OF VARIATIONS IN DISCRETE SPACE FOR CONSTRAINED NONLINEAR DYNAMIC OPTIMIZATION

Yixin Chen and Benjamin W. Wah

Department of Electrical and Computer Engineering
and the Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA

ICTAI 2002, USA

November 4, 2002

Outline

- Introduction
  - Discrete-space dynamic optimization problems
  - Existing approaches
- Calculus of variations in discrete time and discrete state space
  - Necessary and sufficient conditions for constrained local minima
  - Variational search algorithm with node dominance
- Implementations in ASPEN
  - Some sample results
- Conclusions

INTRODUCTION

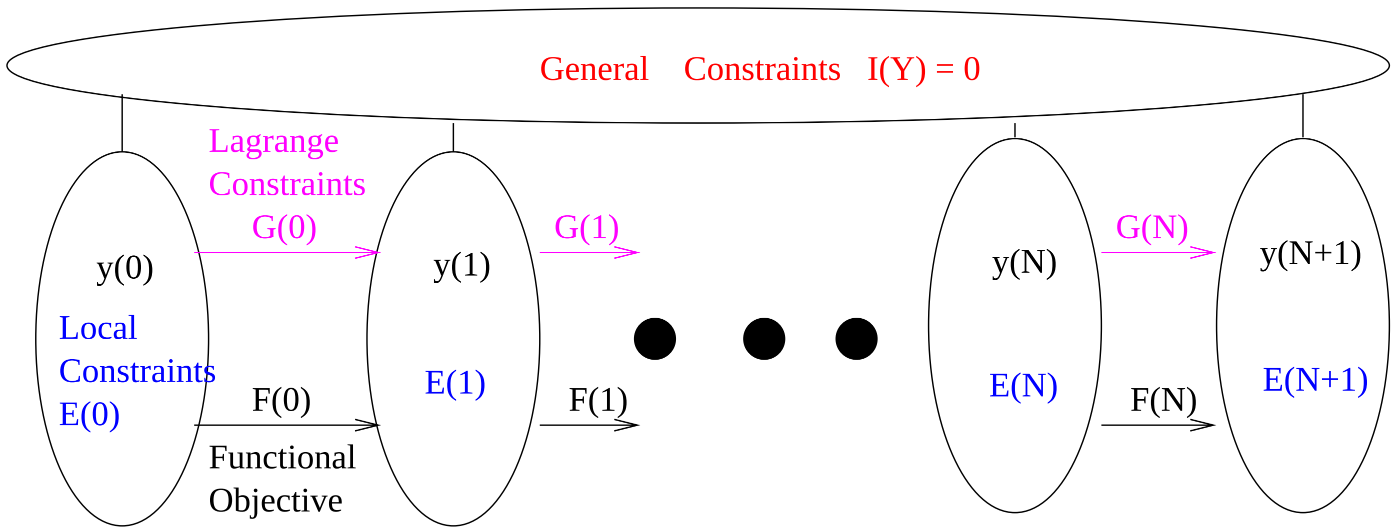
Dynamic Optimization Problems

Optimization problems with time varying dynamic variables

State \ Time	Continuous	Discrete
Continuous	Continuous−time Continuous−state	Discrete−time Continuous−state
Discrete	Continuous−time Discrete−state	Discrete−time Discrete−state

Control theory and classical theory of calculus of variations

Discrete-Time Discrete-State Constrained Dynamic Problems



$$\text{minimize} \quad J[\{y(j)\}] = \sum_{j=0}^{N} F(j, y(j+1), y(j)) \tag{1}$$

$$\text{s.t.} \quad G(j, y(j+1), y(j)) = 0, \quad j = 0, 1, \cdots, N \tag{2}$$

$$E(j, y(y)) = 0, \tag{3}$$

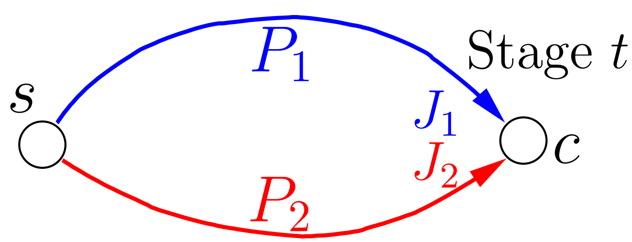
$$I(Y) = 0 \tag{4}$$

where $y(j)$ is defined in *discrete* space $\mathcal{Y}$,

F, G and I are *not* necessarily continuous or differentiable

Unconstrained Problems or Problems with Lagrange Constraints

- Path Dominance: Principle of Optimality
  - Principle of Optimality in dynamic programming applied on feasible state c



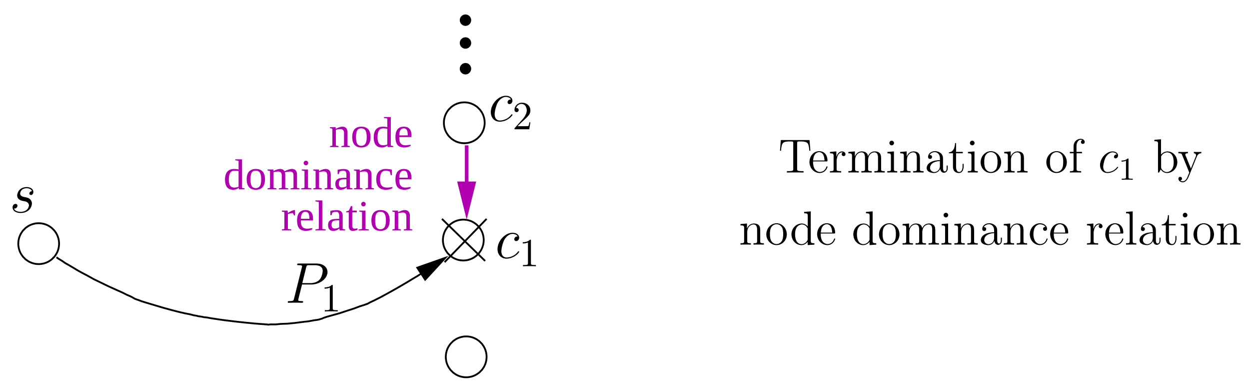
If c lies on the optimal path between s and d and

$$J_2 \leq J_1 \quad \Longrightarrow \quad P_2 \rightarrow P_1$$

- Polynomial worst-case complexity: $O\left(N|\mathcal{Y}|^2\right)$

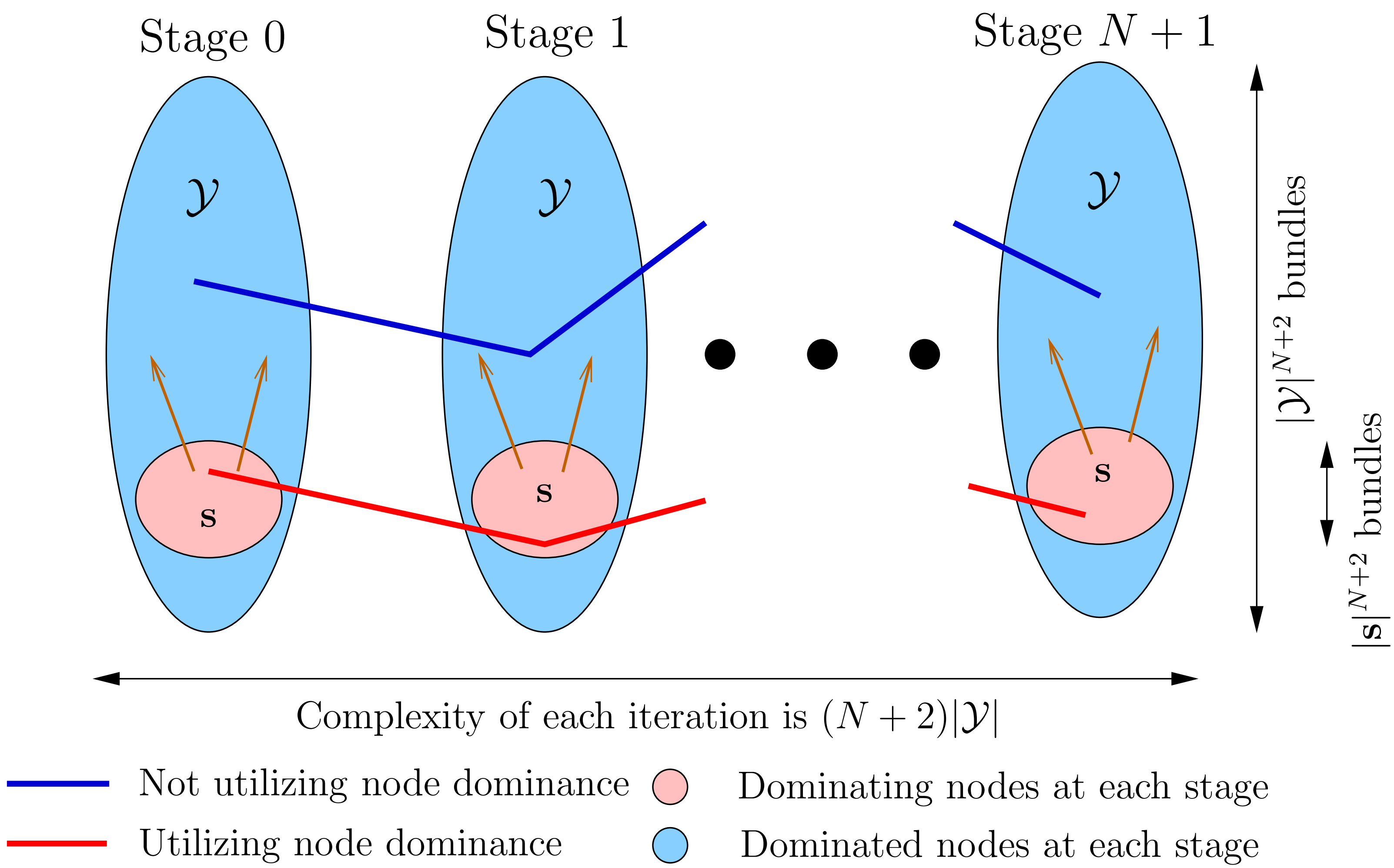
Problems with General Constraints

- Path dominance not applicable
  - A dominating path may become infeasible due to general constraints
  - Exponential worst-case complexity: $O\left(|\mathcal{Y}|^N\right)$, assuming NP hard
- Node dominance applicable



  - Worst-case complexity without path dominance: $O\left(N|\mathcal{Y}||\mathbf{s}|^N\right)$
    - Assuming $|\mathbf{s}|$ nodes are not pruned in each stage
    - Substantially less than $O(|\mathcal{Y}|^N)$ when $|\mathbf{s}| \ll |\mathcal{Y}|$

Benefits of Using Node Dominance



Worst-Case Complexities in Path and Node Dominance

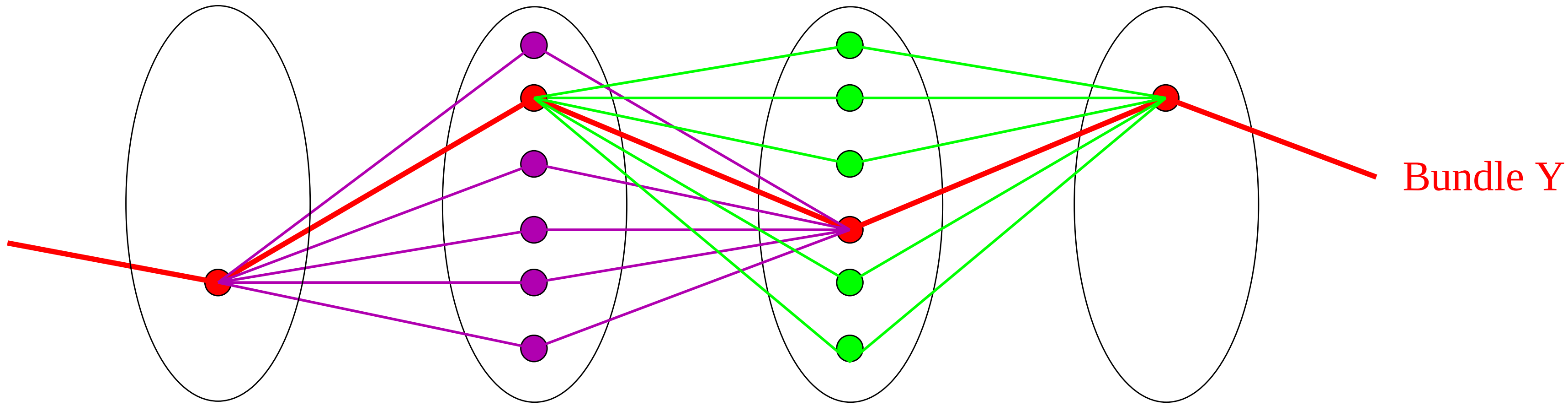
Constraint Type	Without General Constraints (Conventional DP)		With General Constraints (Without Path Dominance)	
Dominance Used	With Path Dominance	With Path and Node Dominance	Without Node Dominance	With Node Dominance
Complexity	$O\left(N\|\mathcal{Y}\|^2\right)$	$O\left(N\|\mathcal{Y}\| + N\|\mathbf{s}\|^2\right)$	$O\left(\|\mathcal{Y}\|^N\right)$	$O\left(N\|\mathcal{Y}\|\|\mathbf{s}\|^N\right)$

Node dominance works well when $|\mathbf{s}| \ll |\mathcal{Y}|$

CALCULUS OF VARIATIONS IN DISCRETE SPACE

Constrained Local-Minimum Bundle in Discrete Space

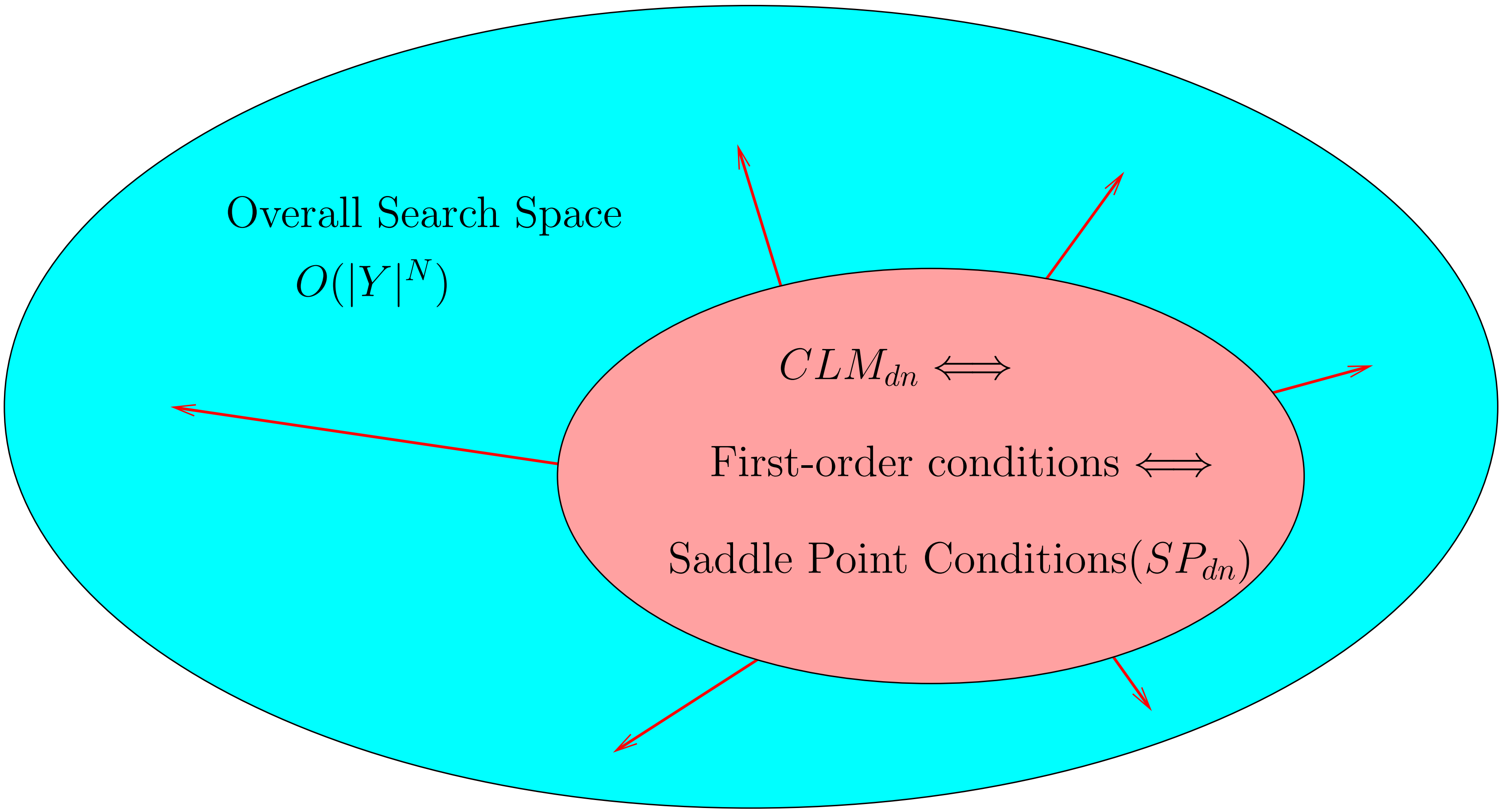
- State vector s in discrete state space $\mathcal{Y}$ has user-defined discrete neighborhood
- Discrete neighborhood of bundle $Y = \{y(j)\}$ is the union of discrete neighborhoods of all stages, each defined on neighborhoods of states in each stage



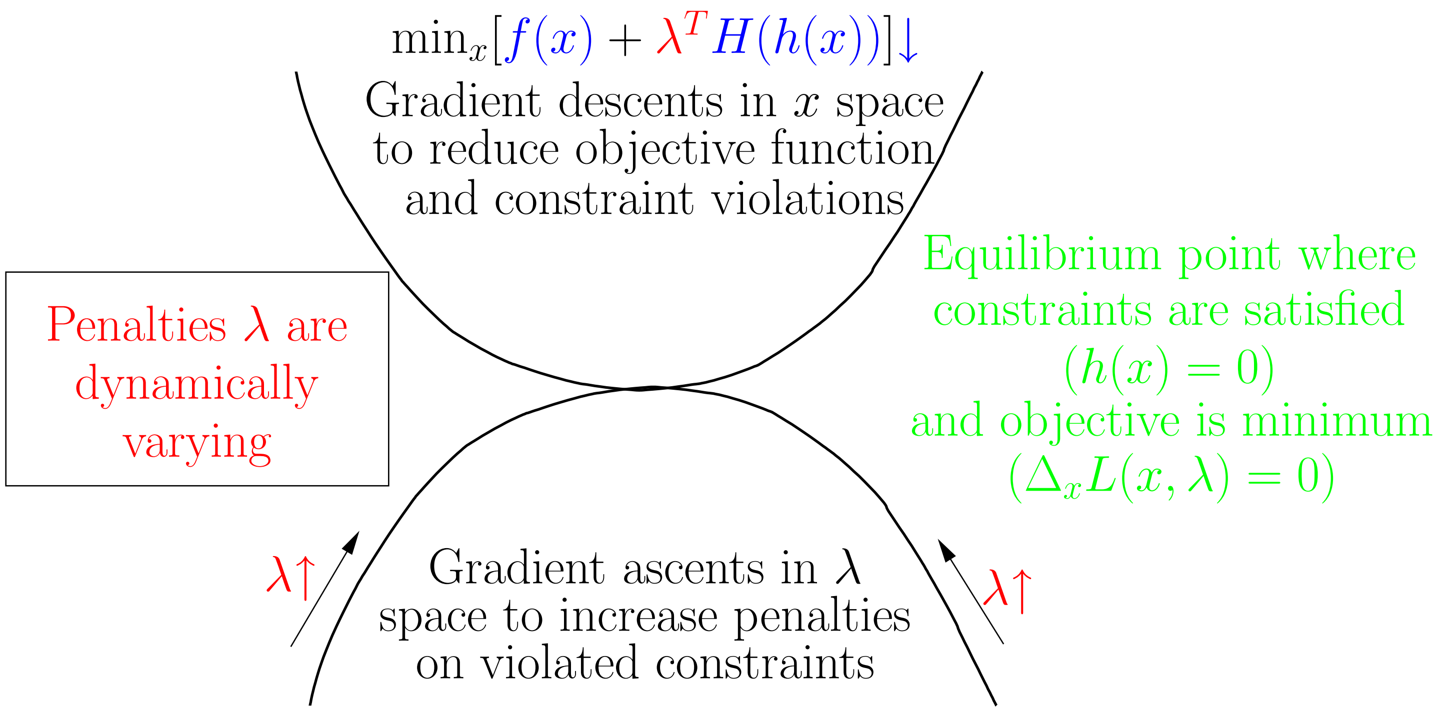
- Bundle Y is a constrained local minimum in discrete space (CLM_{dn}) if
  - Y is feasible
  - No feasible bundle in $\mathcal{N}_b(Y)$ has better functional value than $J[Y]$

Solving Overall Problem using Discrete-Space Lagrangian Theory

- First-order necessary and sufficient conditions based on the theory of Lagrange multipliers in discrete space



Intuitive Meaning Behind Saddle Points



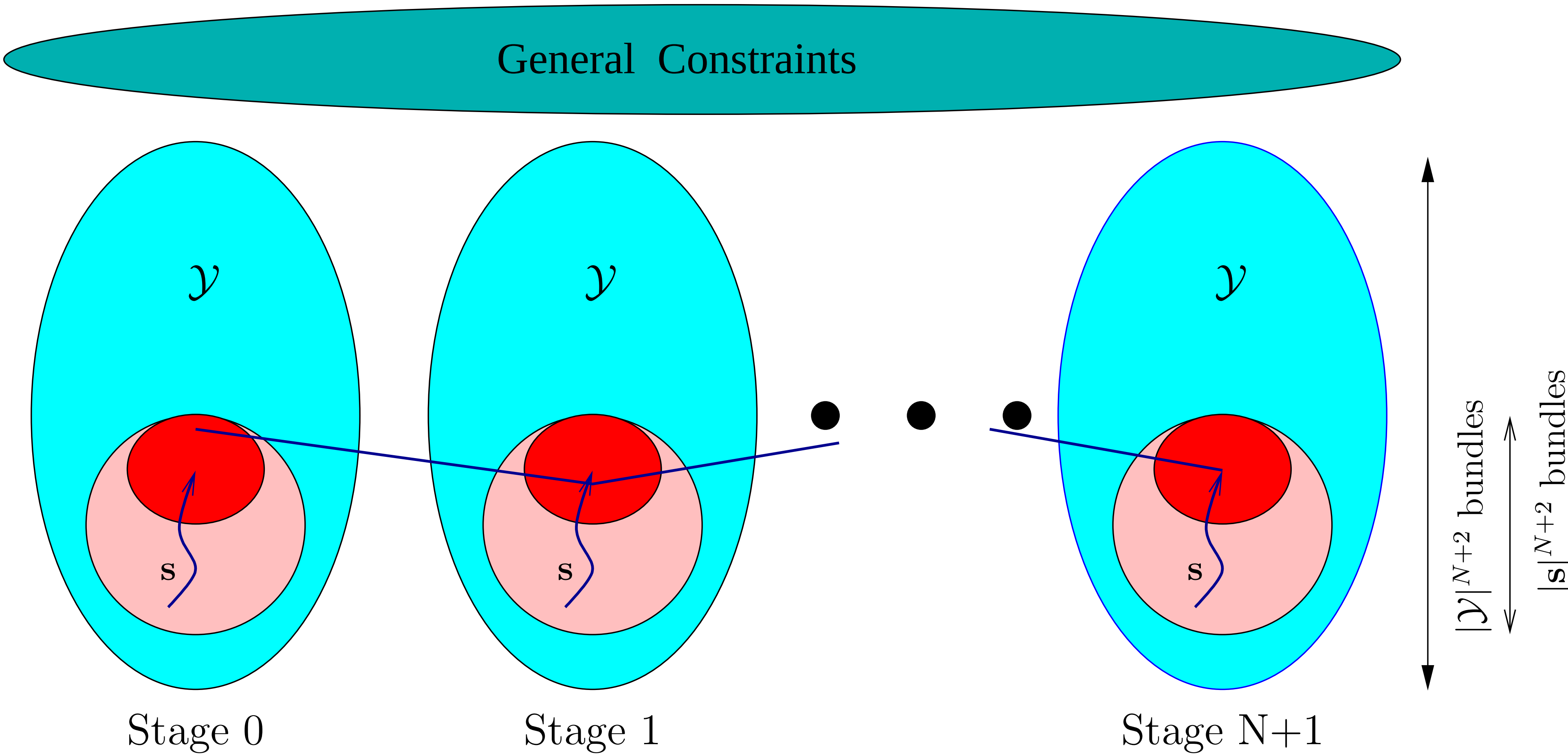
Discrete-Time Discrete-State Euler-Lagrange Equation

- Decompose first-order condition into Discrete-Space Euler-Lagrange Equations (ELE_{dn}) for each stage
- Decompose SP_{dn} in into $N + 2$ distributed saddle-point conditions DSP_{dn}
- Distributed necessary and sufficient conditions for CLM_{dn}

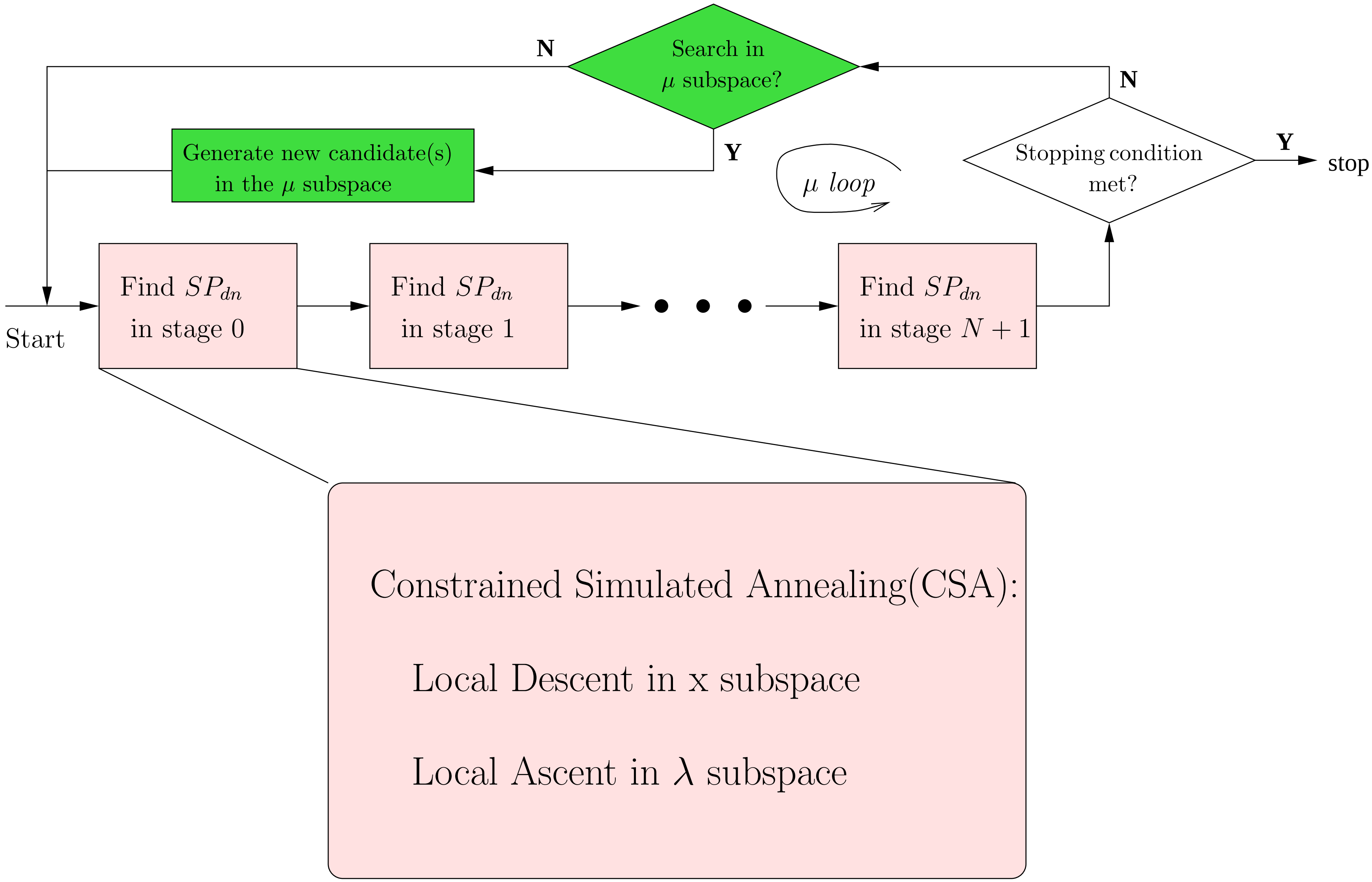
$$CLM_{dn} \equiv ELE_{dn} \equiv DSP_{dn}$$

  - Only necessary when there are general constraints
- Differences from continuous variational calculus theory
  - Do not require differentiability or continuity of functions
  - Continuous ELE conditions are only necessary but not sufficient, even in the absence of general constraints

Solving Distributed Subproblems Using Node Dominance



Heuristic Search Procedure For Finding DSP_{dn}



DEMONSTRATIONS ON ASPEN

ASPEN Planner

- Automated Scheduling and Planning Environment at Jet Propulsion Laboratory
- ASPEN models have discrete time horizons
  - Each time point is a stage
  - Adjacent time points can be collapsed into a single stage
  - Current implementation: maximum 100 stages
- ASPEN repair/optimization actions provide promising descent directions in state-variable subspace
  - Repair actions: resolve conflicts
  - Optimization actions: optimize preferences
- ASPEN does not have the UNDO mechanism

Distributed Lagrangian Formulation

- Assign each conflict c_i a unique Lagrange Multiplier λ_i
- Augmented distributed Lagrangian function of stage t:

$$\Gamma_{dn}(t) = -w_s \cdot \text{Score} + \sum_{c_i \in C(t)} \lambda_i * H(c_i) + \sum_{c_i \in C(t)} \frac{H(c_i)^2}{2}$$

  - w_s: weight of score
  - Score: preference score of schedule
  - $C(t)$: set of conflicts whose time duration intersects with stage t
  - $H(c_i)$: non-negative value assigned to a conflict reflecting its degree of violation
    - $H(c_i) = 1$ in current implementation

Distributed Heuristic Search for Finding SP_{dn} in Stage t

- Descent of $\Gamma_{dn}(t)$ in state subspace
  - Choose probabilistically from repair actions and optimization actions
  - Select random feasible action at each choice point
  - Apply selected action to current schedule in a child process
  - Evaluate $\Gamma_{dn}(t)$ of the new schedule
  - Accept new schedule according to Metropolis probability controlled by a geometrically decreasing temperature
  - Repeat action in parent process if accepted; otherwise, discard result of child process (overcome lack of UNDO but limited to 3685 forks)
- Ascent of $\Gamma_{dn}(t)$ in Lagrange-multiplier subspace

$$\lambda_i \longleftarrow \lambda_i + \alpha_i H(c_i)$$

SOME SAMPLE RESULTS

Benchmark CX1-PREF

- Citizen Explorer-I satellite design and operation planning benchmark
  - Multiple competing preferences to be optimized
  - Problem generator to generate different problem instances

    $perl \quad probgen.pl \quad < random\ seed > \quad < number\ of\ orbits >$

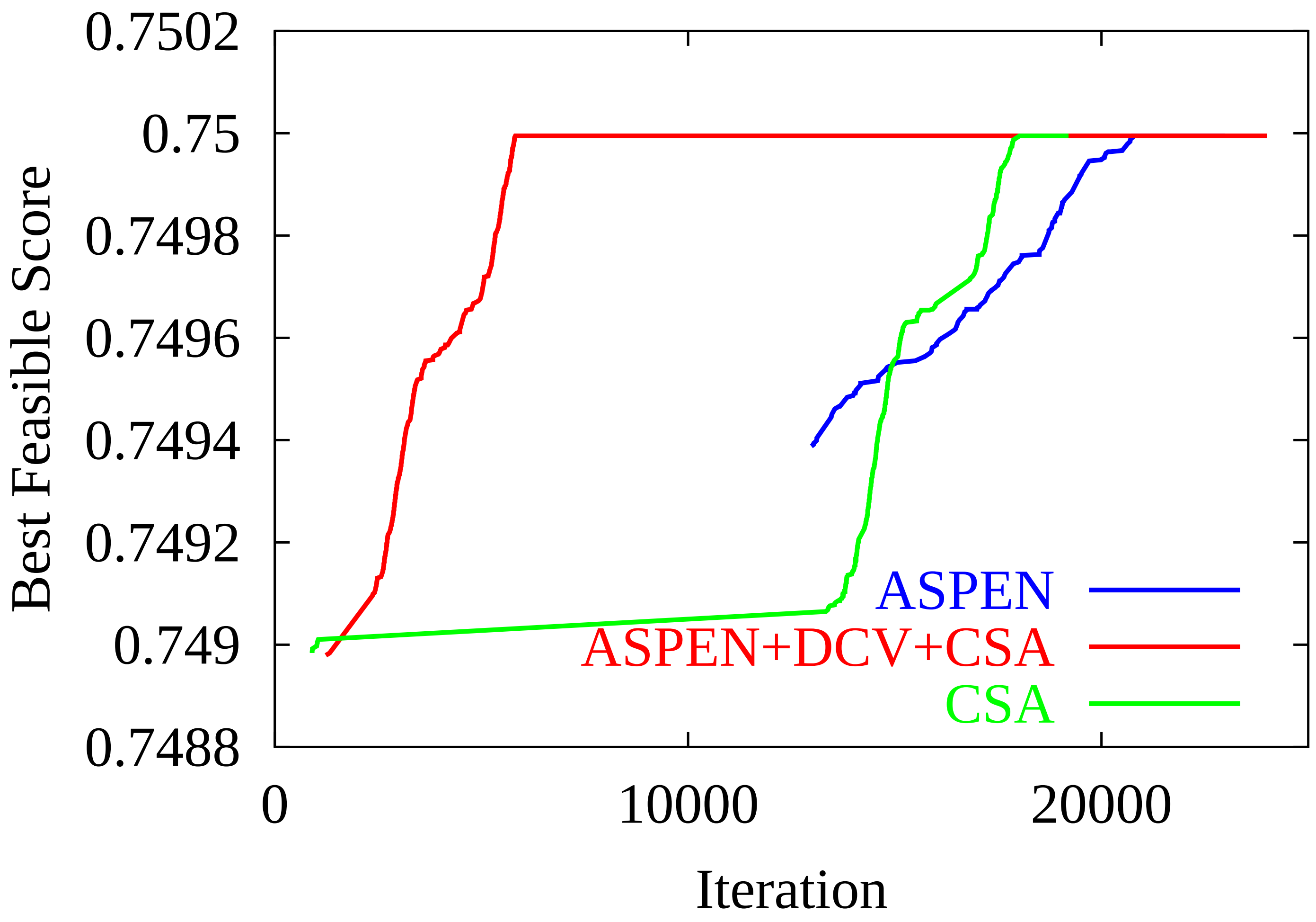
- ASPEN search setting:

  a) Find feasible schedule using *repair*

  b) Optimize score using *optimize* (default 200 iterations)

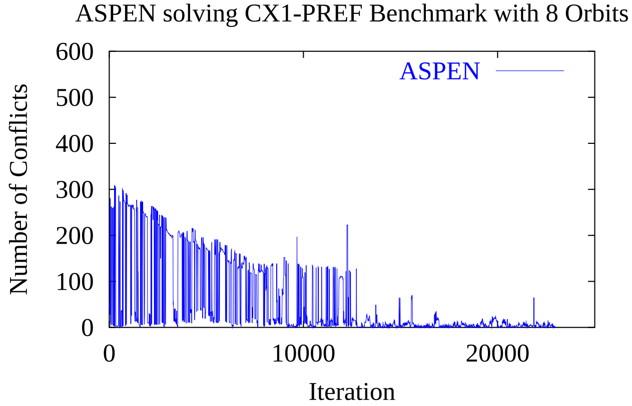
  c) repeat (a) and (b)

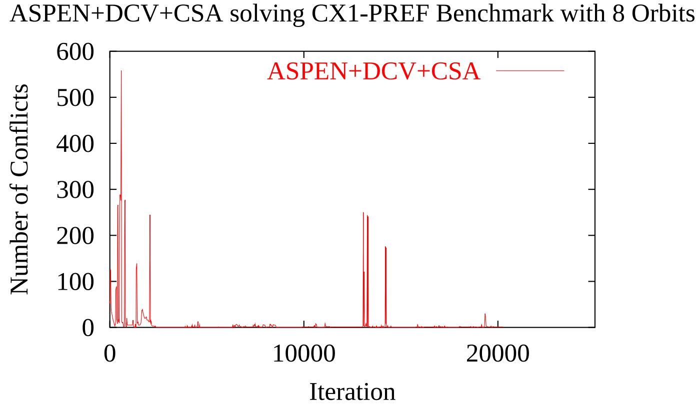
Best Feasible Solution on an 8-Orbit Problem



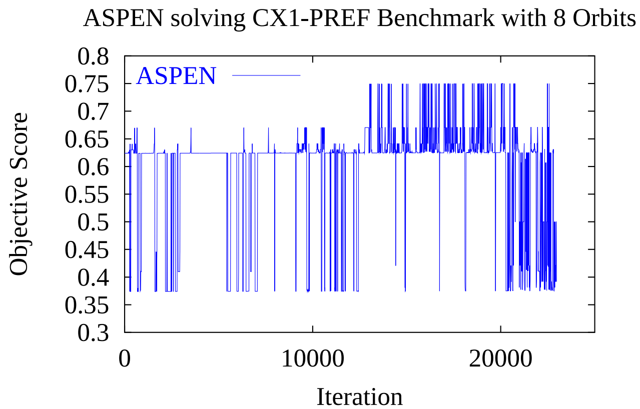
Search Progress on an 8-Orbit Problem

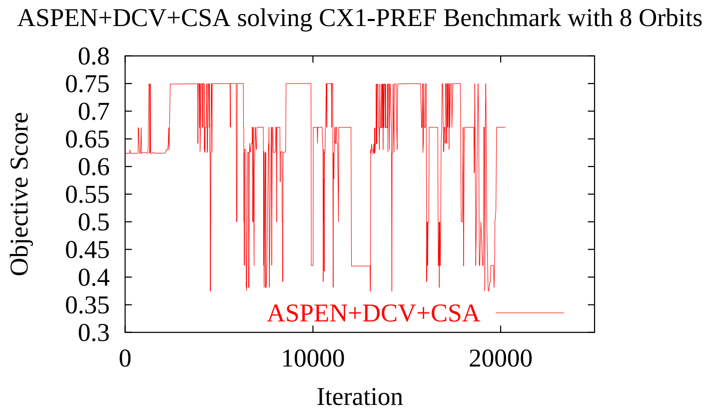
- Conflicts vs. Iteration:



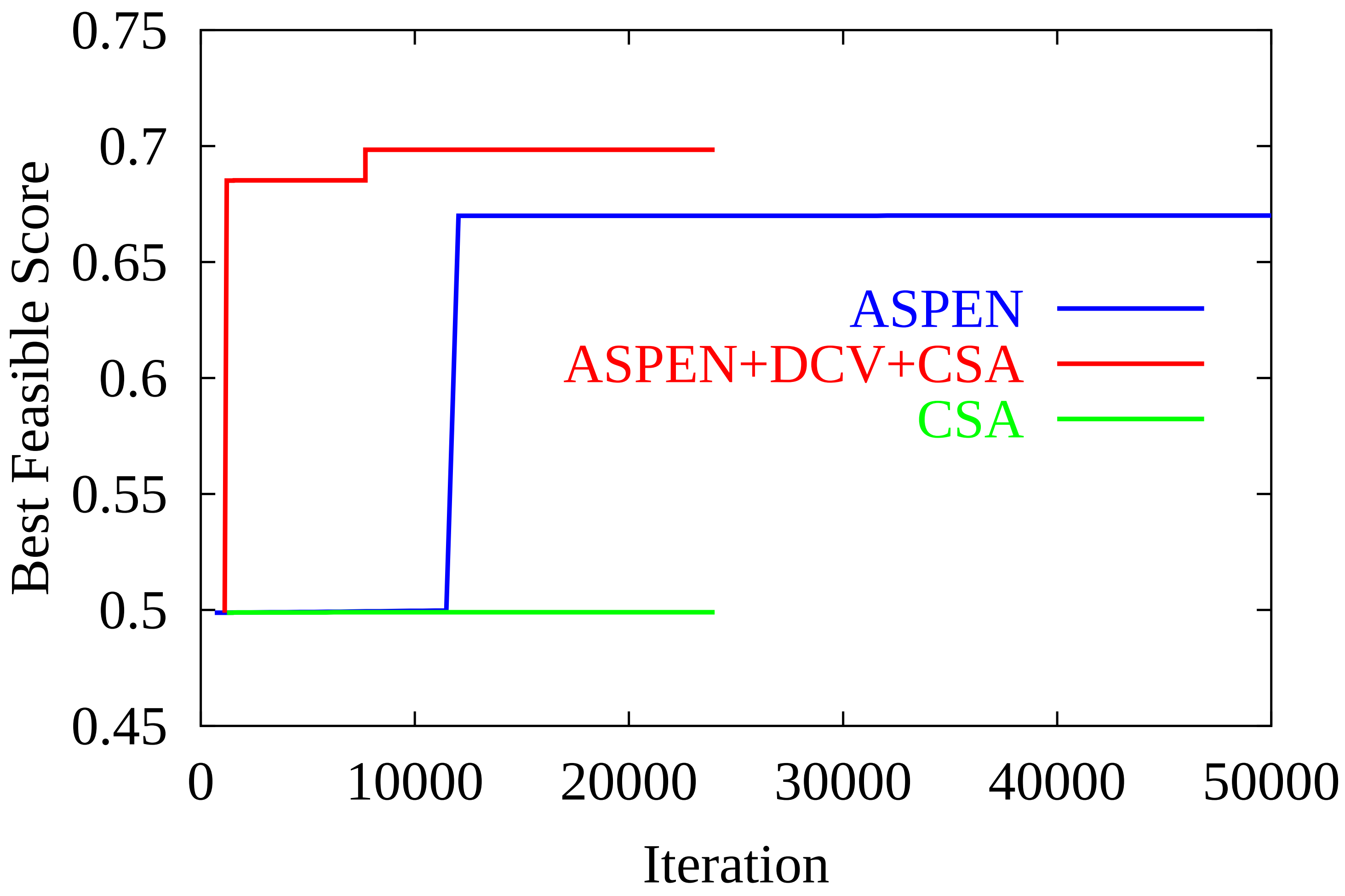


- Score vs. Iteration:



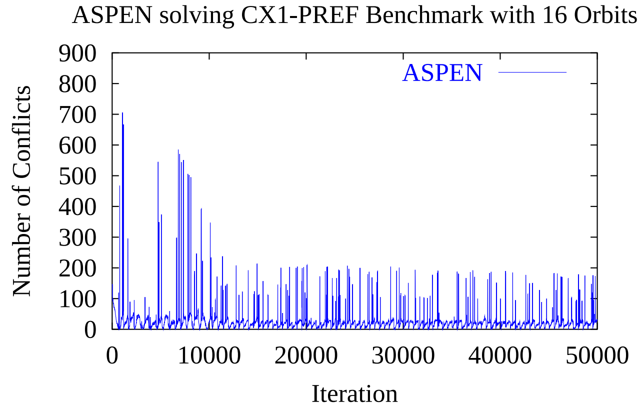


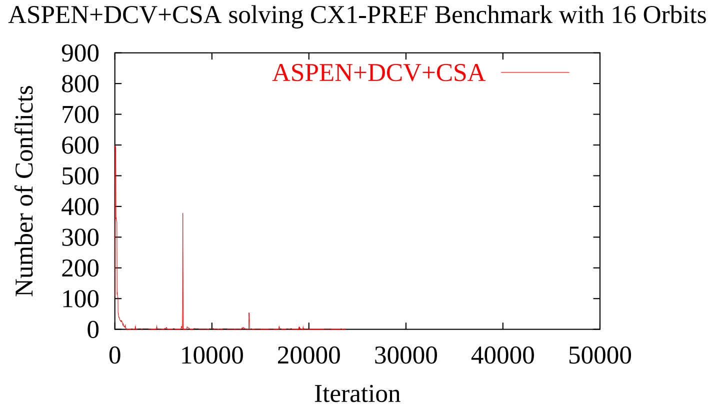
Best Feasible Solution on a 16-Orbit Problem



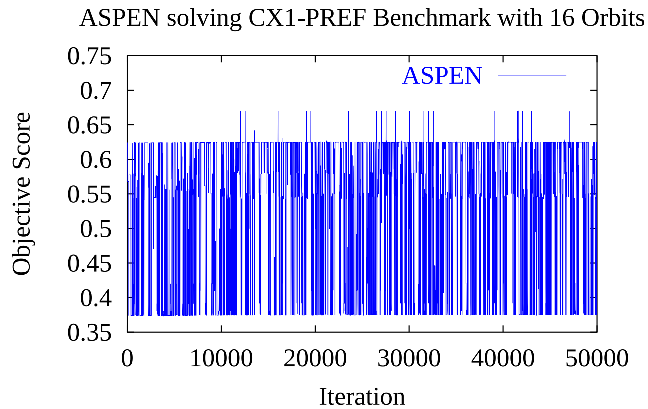
Search Progress on a 16-Orbit Problem

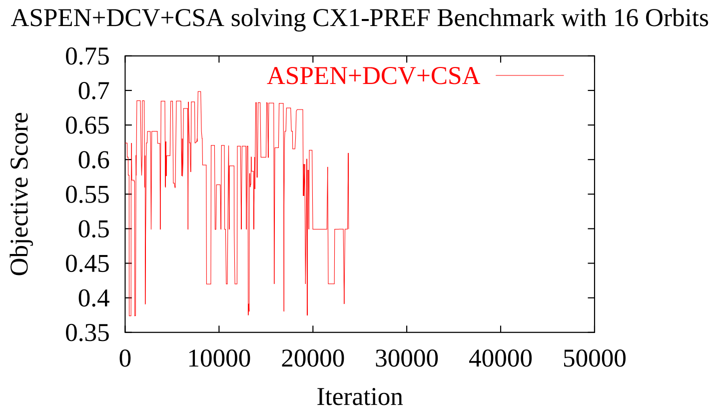
- Conflicts vs. Iteration:



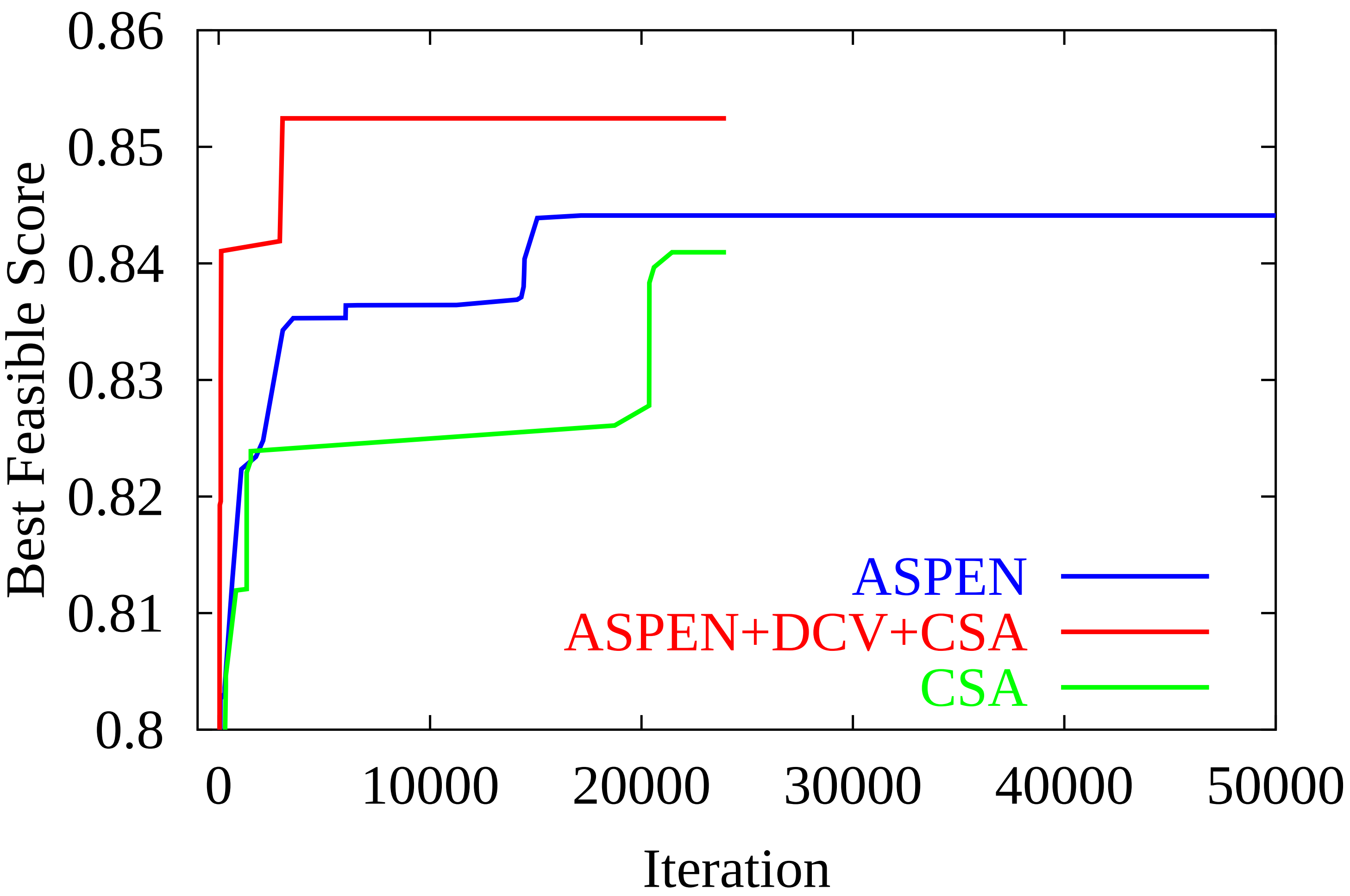


- Score vs. Iteration:



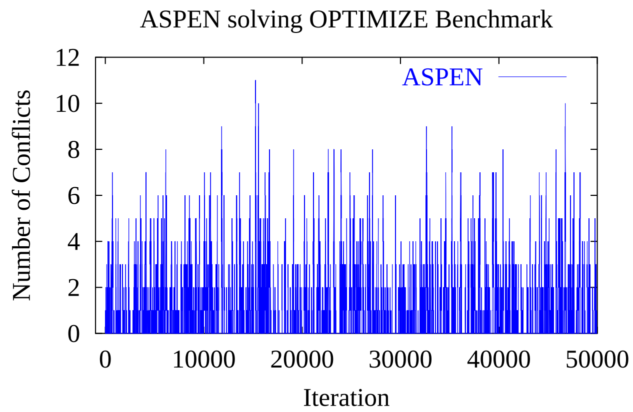


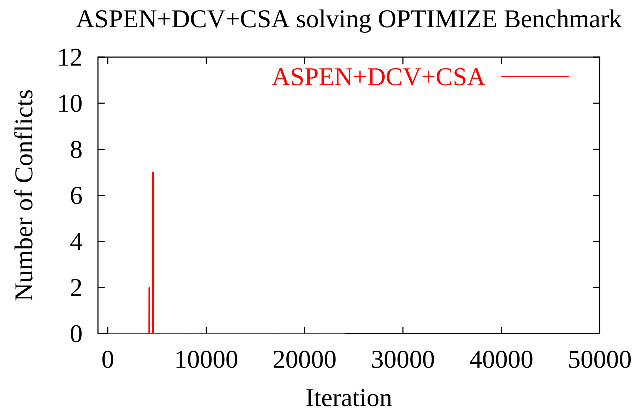
Best Feasible Solution on OPTIMIZE Benchmark



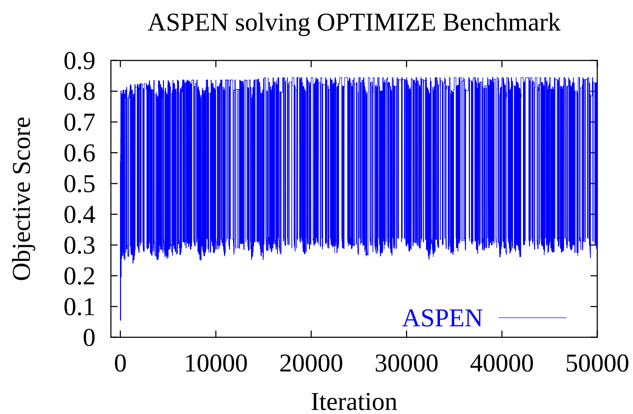
Search Progress on OPTIMIZE Benchmark

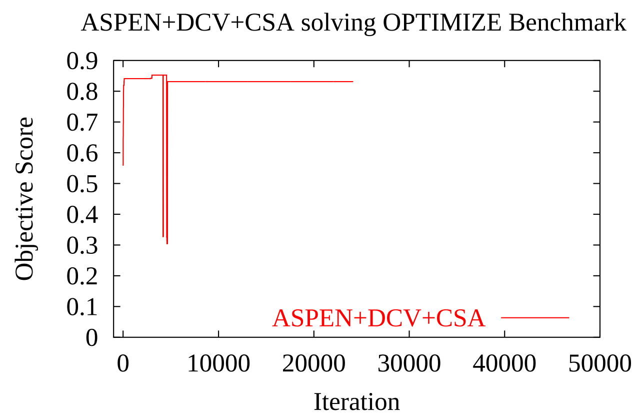
- Conflicts vs. Iteration:



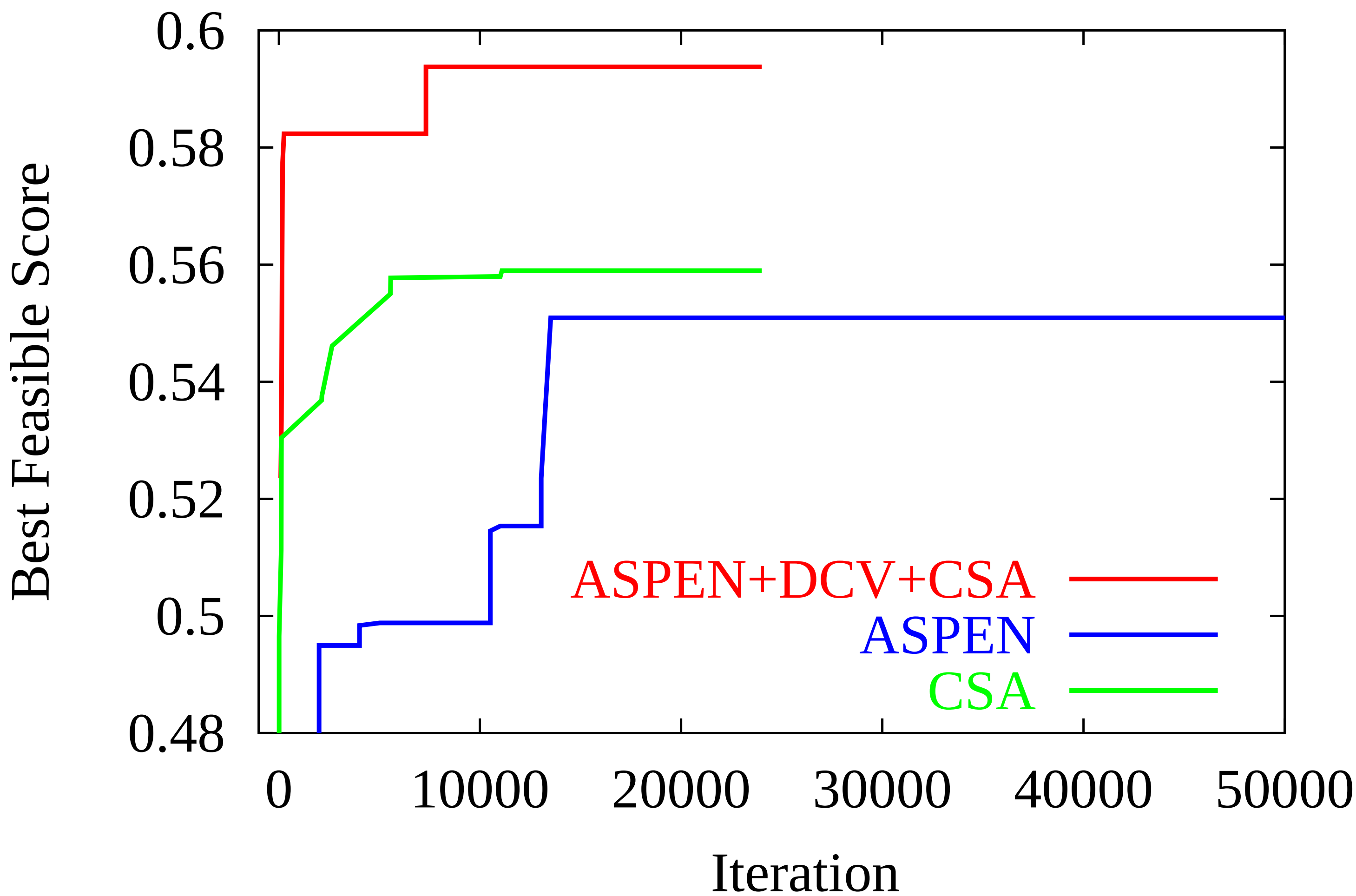


- Score vs. Iteration:



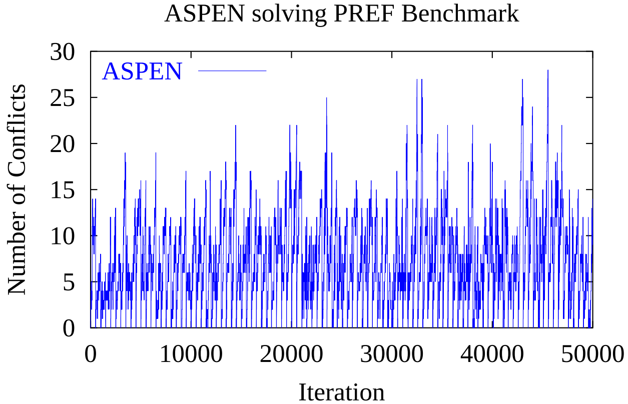


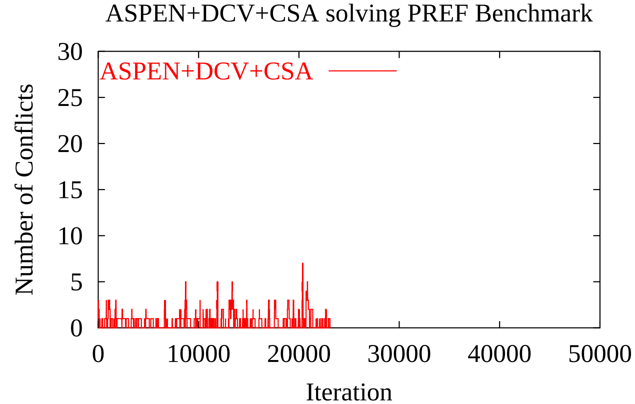
Best Feasible Solution on PREF Benchmark



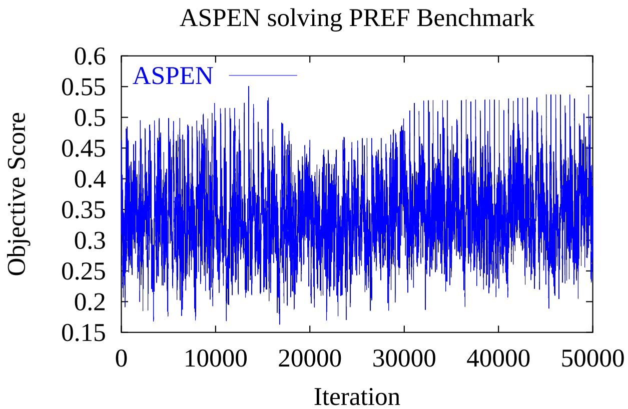
Search Progress on PREF Benchmark

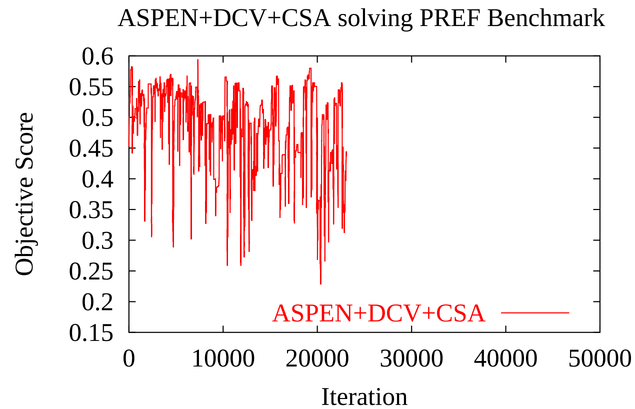
- Conflicts vs. Iteration:





- Score vs. Iteration:





Conclusions

- Extension of calculus of variations in continuous space to discrete space
- Partitioning general necessary conditions into distributed necessary conditions
- Relying on theory of Lagrange multipliers for discrete constrained optimization
- Significant reduction in the base of the exponential complexity
- Significant improvement in search times with equal or better quality in planning and scheduling problems as compared to those of ASPEN