PARTITIONING OF TEMPORAL PLANNING PROBLEMS IN MIXED SPACE USING THE THEORY OF EXTENDED SADDLE POINTS

Benjamin W. Wah and Yixin Chen
Department of Electrical and Computer Engineering
and the Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
http://manip.crhc.uiuc.edu

Presented at the 2003 IEEE Int'l Conf. on Tools for AI

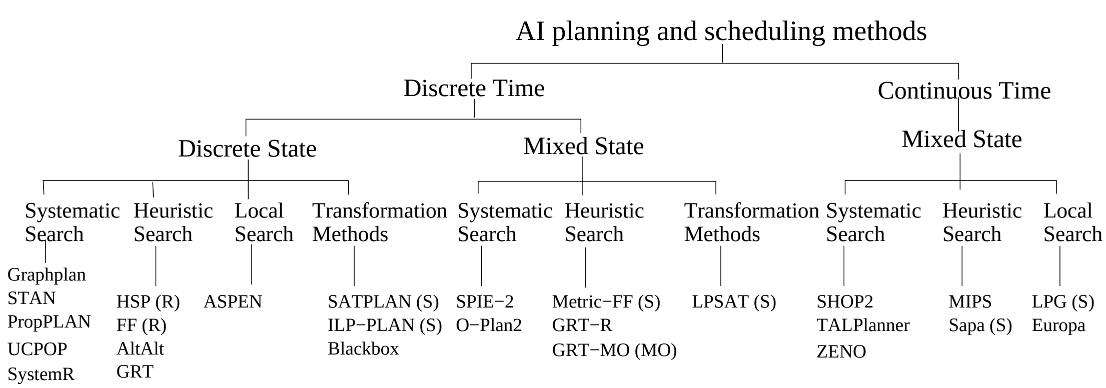
Research supported by NASA and NSF

Outline

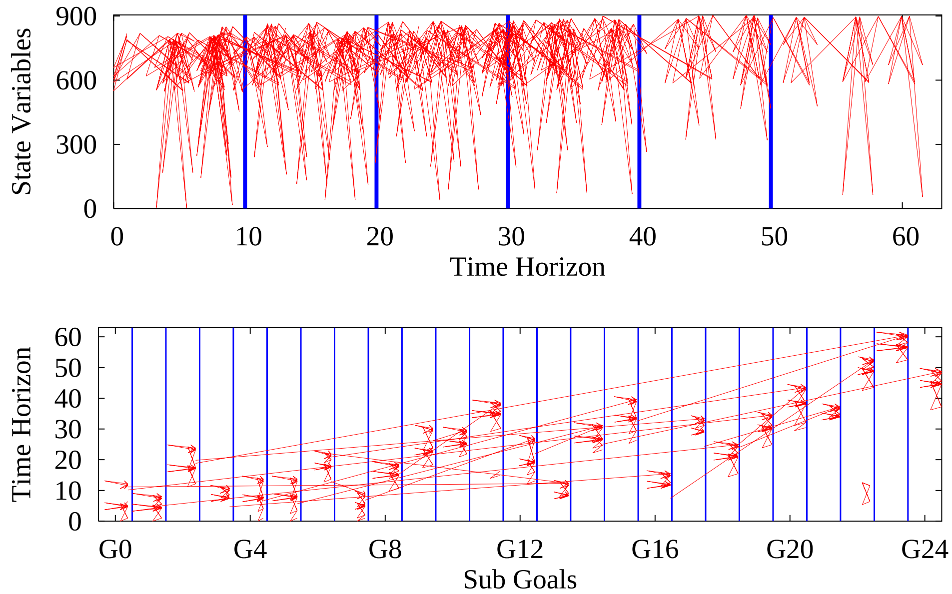
- Research problem addressed
  - Limitation of existing planning methods
- Theory of Lagrange multipliers for mixed constrained optimization
  - Necessary and sufficient extended saddle-point condition
  - Iterative implementations
- Partitioning of variable space
  - Distributed extended saddle-point condition
  - Distributed Iterative implementations
- Experimental results on MIPS and PDDL2.1
- Conclusions

INTRODUCTIONS

A Classification of Existing Approaches in Planning



Partitioning by Time versus Partitioning by SubGoals

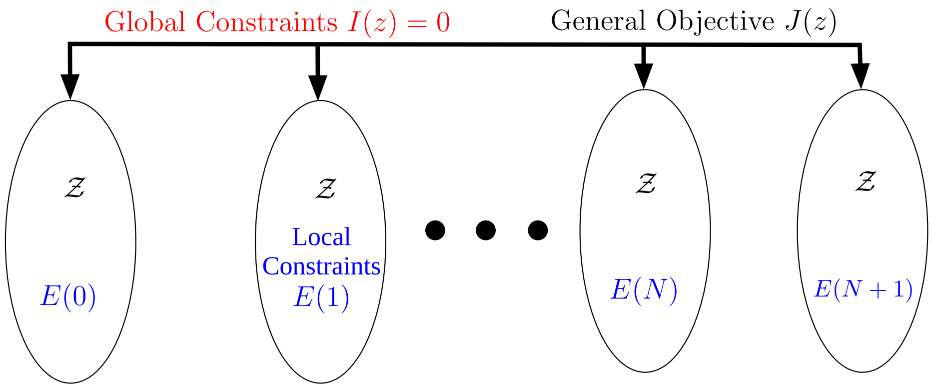


Research Problem Addressed

Partition (mixed-space and continuous-time) temporal planning problems and develop methods for resolving global constraints across partitions

- Partitioned problems have lower time and space complexity
- Overall problem can be solved better and more efficiently

Mathematical Formulation

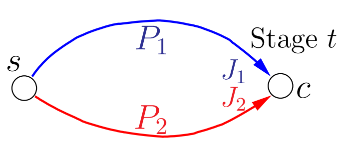


$$
\begin{aligned}
\min_{z} \quad & J(z) \\
\text{subject to} \quad & E(z(j)) = 0, \quad j = 0, 1, \cdots, N+1 \\
& I(z) = 0
\end{aligned}
$$

where $z(j)$ is defined in *mixed* space $\mathcal{Z}$ of stage j,
E, I and J_i are *not* necessarily continuous and differentiable

Dynamic Programming Cannot Be Applied

- Path dominance on multi-stage search with local constraints
  - Principle of Optimality applied on feasible state c



If c lies on the optimal path between s and d and
$J_2 \leq J_1 \quad \Longrightarrow \quad P_2 \to P_1$

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

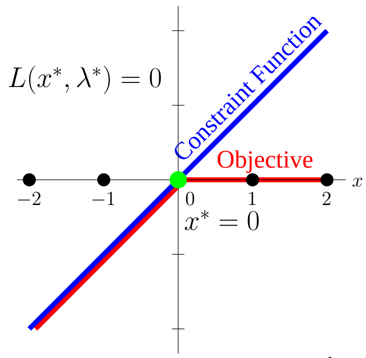
- Path dominance is not applicable when there are global constraints
  - A dominating path early on may become infeasible due to global constraints that got violated later
  - Exponential search space: $O\left(|\mathcal{Z}|^{N+2}\right)$

Penalty-Based Methods Do Not Always Work

Penalty-based methods

- By choosing suitable penalties in a penalty function, a local minimum of the penalty function corresponds to a feasible local minimum of the objective

Counter-example



$$\min_{x \in \{-2,-1,0,1,2\}} f(x) = \begin{cases} 0 & x \geq 0 \\ x & x < 0 \end{cases}$$

subject to $x = 0$

Penalty formulation

- $L(x, \lambda) = f(x) + \lambda\, x$
- Hypothesize $L(x, \lambda^*) \geq L(x^*, \lambda^*) = 0$

No λ^* exists when solving

$$\begin{cases} 0 = L(0, \lambda^*) \leq L(-1, \lambda^*) \\ 0 = L(0, \lambda^*) \leq L(1, \lambda^*) \end{cases}$$

$$\Longrightarrow \begin{cases} \lambda^* \leq -1 \\ \lambda^* \geq 0 \end{cases}$$

Mathematical Programming Methods

Continuous Methods: unique λ cannot be found in distributed methods

- Necessary KKT condition
- Sufficient saddle point condition

MINLP methods: require the function of subproblems to be convex or factorable

- Generalized Benders Decomposition
- Outer Approximation
- Branch-and-Reduce Methods

THEORY OF EXTENDED SADDLE POINTS FOR MIXED CONSTRAINED OPTIMIZATION

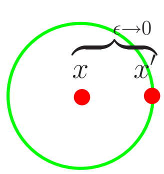
Mixed Neighborhood $\mathcal{N}_m(z)$ of Point z

- $\mathcal{N}_m(z) = \mathcal{N}_m(x, y) = \Big\{(x', y) \mid x' \in \mathcal{N}_c(x)\Big\} \bigcup \Big\{(x, y') \mid y' \in \mathcal{N}_d(y)\Big\}$

Continuous Subspace: $\mathcal{N}_c(x)$

x is a vector of continuous variables

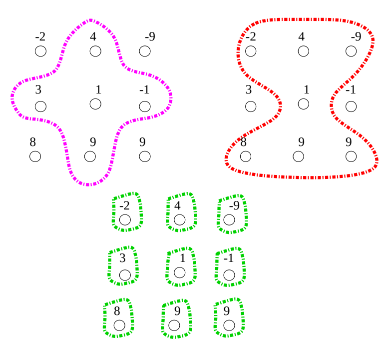
Neighborhood defined by open sphere



Discrete Subspace: $\mathcal{N}_d(y)$

y is a vector of discrete variables

User defined neighborhood

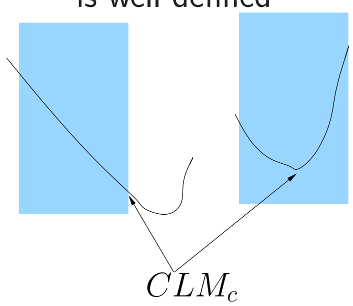


Constrained Local Minimum (CLM)

- Feasible z is CLM_m in mixed space if $J(z) \le J(z') \ \forall$ feasible $z' \in \mathcal{N}_m(z)$

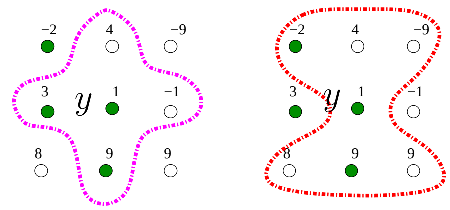
Continuous Subspace: CLM_c

- Feasible local minimum when compared to feasible points inside an open sphere
- Whether point x is a CLM_c is well defined



Discrete Subspace: CLM_d

- Feasible local minimum with respect to neighboring feasible points
- Whether point y is a CLM_d depends on $\mathcal{N}_d(y)$



CLM_d at y Not CLM_d at y

Lagrangian Formulation of Mixed Optimization Problem

- Transformed Lagrangian function with extended Lagrange multipliers γ and μ

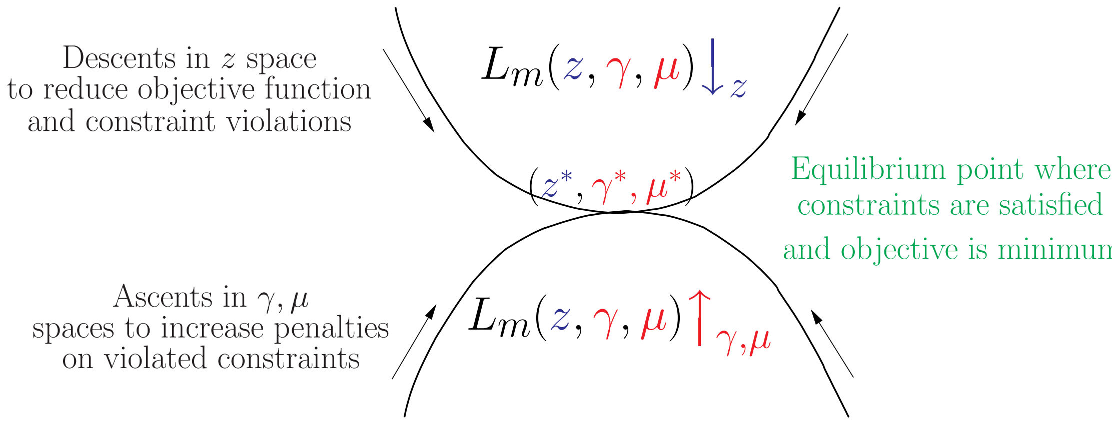
$$L_m(z, \gamma, \mu) = J(z) + \sum_{t=0}^{N+1} \gamma^T(t) \cdot \left| E(z(t)) \right| + \mu^T \cdot \left| I(z) \right|$$

- Necessary and sufficient Extended Saddle-Point Condition (ESPC)
  - z^* is a CLM_m iff (z^*, γ^*, μ^*) is a mixed-neighborhood saddle point (SP_m)

$$L_m(z^*, \gamma, \mu) \le L_m(z^*, \gamma^*, \mu^*) \le L_m(z, \gamma^*, \mu^*)$$

- (z^*, γ^*, μ^*) is at
  - Local minimum of L_m with respect to z
  - Local maximum of L_m with respect to γ and μ
- Condition is true for $\gamma^{**} > \gamma^*$ and $\mu^{**} > \mu^*$

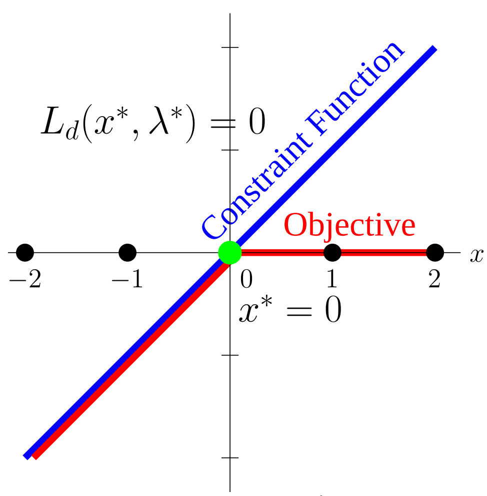
Intuitive Meaning Behind Saddle Points



Although γ^* and μ^* always exists,

- Their search in mixed space may be very time consuming
- The search of $\gamma^{**} > \gamma^*$ and $\mu^{**} > \mu^*$ is much easier

Continuing from the Previous Example



$$\min_{x \in \{-2,-1,0,1,2\}} f(x) = \begin{cases} 0 & x \geq 0 \\ x & x < 0 \end{cases}$$

subject to $x = 0$

Transformed Lagrangian function

- $L_d(x, \lambda) = f(x) + \lambda\,|x|$
- Find λ^* such that $L_d(x, \lambda^*) \geq L_d(x^*, \lambda^*)$

Solving

$$\begin{cases} 0 = L_d(0, \lambda^*) \leq L_d(-1, \lambda^*) \\ 0 = L_d(0, \lambda^*) \leq L_d(1, \lambda^*) \end{cases}$$

leads to $\lambda^* \geq 1$

Pick $\lambda^* = 1$
Saddle-point condition applies for $\lambda^{**} > \lambda^*$

Iterative Implementation

Algorithm needs to look for $\gamma^{**} > \gamma^*$ and $\mu^{**} > \mu^*$

$L_m(z,\gamma,\mu)\uparrow_{\gamma,\mu}$ to find γ^{**}, μ^{**}

Outer Loop

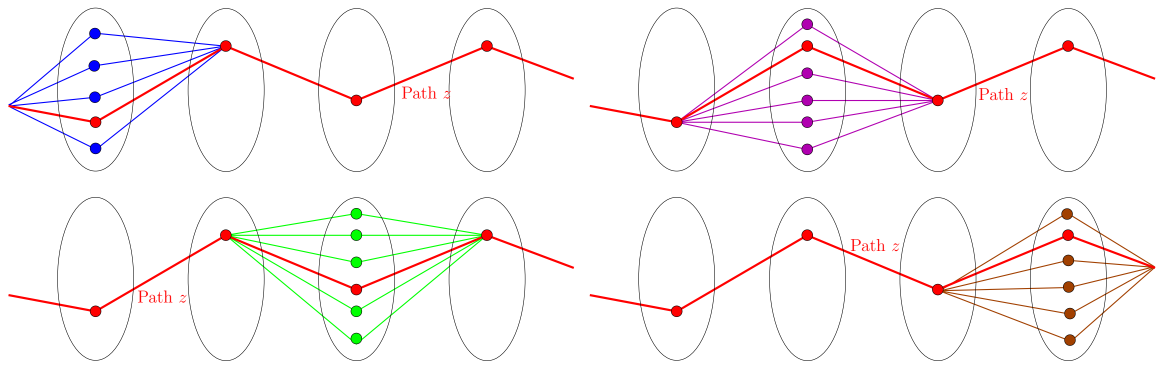
$L_m(z,\gamma,\mu)\downarrow_{z}$ to find z^*

Inner Loop

PARTITIONING OF ESPC FOR SEPARABLE NEIGHBORHOODS

Separable Neighborhoods

$\mathcal{N}_p(z)$ (mixed neighborhood of path $z = (z(0), \ldots, z(N+1))^T$) is the union of mixed neighborhoods in each stage, while keeping the path fixed in other stages



Path z is a constrained local minimum in mixed space (CLM_m) iff

- z is feasible
- No feasible path in $\mathcal{N}_p(z)$ has better objective value than $J(z)$

Decomposition of Lagrangian Function into Stages

Decompose Lagrangian function

$$L_m(z, \gamma, \mu) = J(z) + \sum_{t=0}^{N+1} \gamma^T(t) \cdot \left|E(z(t))\right| + \mu^T \cdot \left|I(z)\right|$$

into distributed Lagrangian function for stage t, $t = 0, \cdots, N+1$,

$$\Gamma_m(z, \gamma(t), \mu) \;=\; J(z) + \gamma(t) \cdot \left|E(z(t))\right| + \mu \cdot \left|I(z)\right|$$

Distributed Necessary & Sufficient ESPC for CLM_m

- Path z is a CLM_m if and only if it satisfies
  - Distributed Necessary & Sufficient ESPC for all $t = 0, 1, \cdots, N+1$

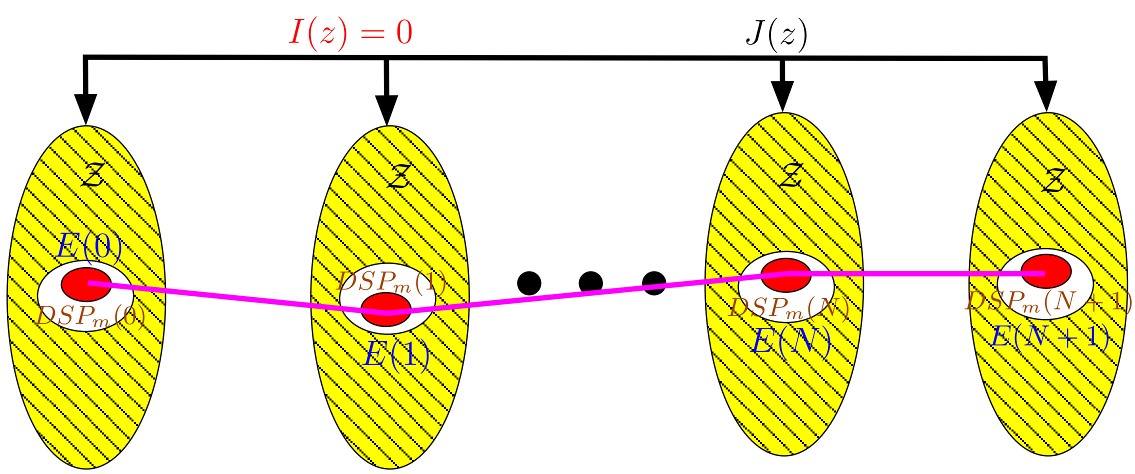
$$\Gamma_m(z^*, \gamma(t)', \mu^*) \le \Gamma_m(z^*, \gamma(t)^*, \mu^*) \le \Gamma_m(z', \gamma(t)^*, \mu^*)$$

$$L_m(z^*, \gamma^*, \mu) \le L_m(z^*, \gamma^*, \mu^*)$$

$$\text{for all } z' = (z(0), \ldots, z(t-1), z(t)', z(t+1), \ldots, z(N+1)) \in \mathcal{N}_p^{(t)}(z^*)$$

  - Condition is also true for all $\gamma(t)^{**} > \gamma(t)^*$ and $\mu^{**} > \mu^*$

Reduced Search Space for Finding Feasible/Optimal Paths



Significant reduction in complexity

Iterative Implementation

$L_m(z,\gamma,\mu)\uparrow_\mu$ to find μ^{**}

$\Gamma_m(z,\gamma,\mu)\uparrow_{\gamma(0)}$ to find $\gamma^{**}(0)$

$\Gamma_m(z,\gamma,\mu)\downarrow_{z(0)}$

• • •

$\Gamma_m(z,\gamma,\mu)\uparrow_{\gamma(N+1)}$ to find $\gamma^{**}(N+1)$

$\Gamma_m(z,\gamma,\mu)\downarrow_{z(N+1)}$

Observation:

- Based on a separable neighborhood, the combined local minimum of Γ_m in all subspaces is the local minimum of L_m

Equivalent Search:

$L_m(z,\gamma,\mu)\uparrow_\mu$ to find μ^{**}

$L_m(z,\gamma,\mu)\uparrow_\gamma$ to find γ^{**}

$L_m(z,\gamma,\mu)\downarrow_z$ to find z^*

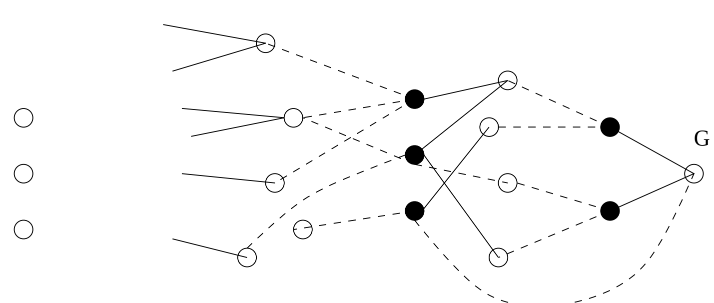
EXPERIMENTAL RESULTS ON MIPS

MIPS+DIS Algorithm

procedure MIPS+DIS
    compute the relevant actions for each goal fact;
    compute the partial orders among goal facts;
    generate an initial ordered goal list of goal facts;
    set $iter \longleftarrow 0$;
    repeat
        for each goal fact in the goal list
            call modified MIPS to solve the subproblem;
        end_for
        if (feasible plan found)
            call PERT to generate & evaluate a parallel plan;
            decrease some Lagrange multipliers;
        else increase Lagrange multipliers γ on unsatisfied global constraints;
        $iter \leftarrow iter + 1$;
        if ($iter \% \tau == 0$) dynamically re-order the goals;
    until no change on z and γ in an iteration;
end_procedure

Search-space reduction for a subproblem

- For each goal fact
  - Backward relevance analysis to get a relevance list of relevant actions



- Possible improvement: tighter reduction

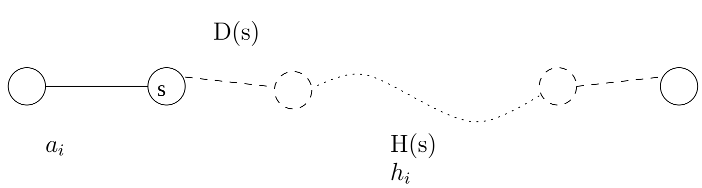
Ordering of goals

- Difficult goals be resolved before easier ones
- Two levels of partial ordering:
  1. Ascending number of irrelevant actions
  2. Descending minimum number of preconditions
- Example: (at person3 city1) $\rightarrow$ (at truck1 city2) $\rightarrow$ (at driver2 city1)
- Possible improvement
  - Heuristic distance from current state to each goal
  - Dynamic ordering during search

Modified MIPS

- Modify heuristic function for A^* search

$$H'(s) = H(s) + D(s) + \sum_{i=1}^{N_G} (\gamma_i a_i + \zeta_i h_i), \tag{1}$$



- Prune nodes not in relevance list at node expansion

Heuristic objective

- A heuristic objective $D(s)$ to measure solution quality:

$$D(s) = \alpha_D * n_d, \tag{2}$$

- α_D is a weighting factor (0.01 in our experiments),
- n_d is the number of action dependencies in the relaxed plan from s to goal
- Possible improvement:
  - Apply PERT and compute the objective function at each s

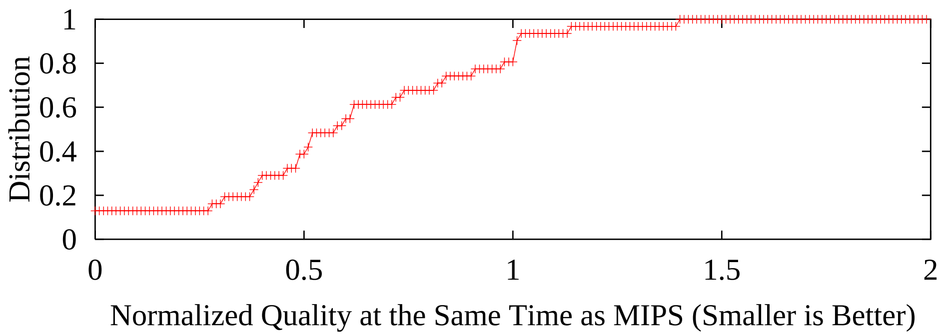
Lagrange Multipliers

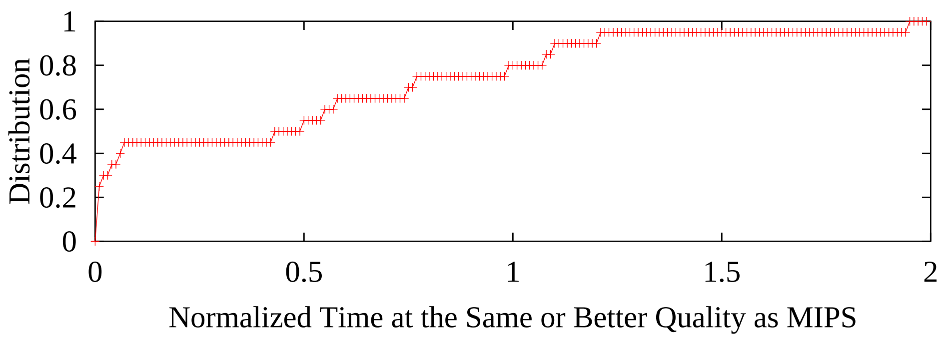
- If a feasible plan is not found:
  - Increase Lagrange multipliers of those unsatisfied goal facts
- If a feasible plan is found:
  - Call PERT to generate parallel plan and evaluate quality
  - Decrease some Lagrange multipliers
- Possible improvement:
  - May periodically decrease Lagrange multipliers even when a feasible plan is not found

Experimental Results

- 140 problems in 7 domains
- 75 problems solvable by MIPS in 1 second:
  - MIPS+DIS improves 67 in solution quality
- 43 problems solvable by MIPS in 10^3 seconds:
  - 80% better trade-off in time and quality by MIPS+DIS
- 22 problems unsolvable by MIPS in 10^3 seconds:
  - 15 solvable by MIPS+DIS in 10^3 seconds:

Distribution of Quality and Time of MIPS+DIS





Conclusions

- Partitioning of mixed constrained optimization in temporal planning
  - Distributed method to resolve global constraints across partitions
  - Significant reduction in search space by reducing the base of the exponential complexity
- Few parameters to tune in algorithm
- Significant improvement on PDDL2.1 planning problems