Value Driven Landmarks for Oversubscription Planning

Daniel Muller & Erez Karpas

Technion - Israel Institute of Technology

June 24, 2018

mullerdm@gmail.com

Value Driven Landmarks for Oversubscription Planning

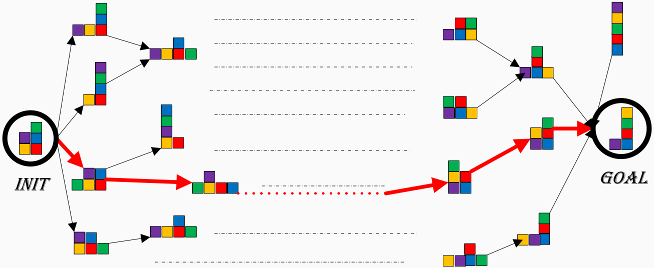
Daniel Muller & Erez Karpas

Technion - Israel Institute of Technology
June 24, 2018

mullerdm@gmail.com

Background

Classical Planning Problem



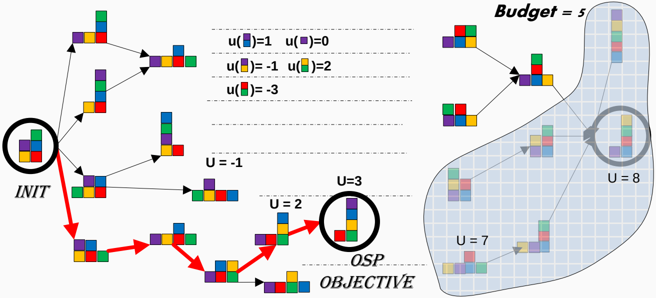
Objective

Find a sequence of actions from a given initial state to achieve a ***goal state***

Optimal Classical Planning

constrained to the ***cheapest*** cost sequence of actions which achieves the goal

Oversubscription (OSP) Planning Problem



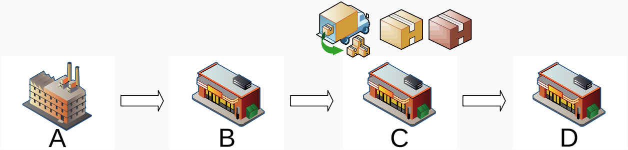
Objective

Find a sequence of actions from a given initial state to *a* ***valuable state***, within a ***limited budget*** on action cost

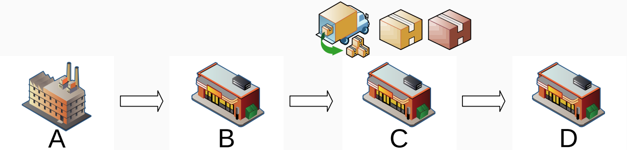
Optimal OSP

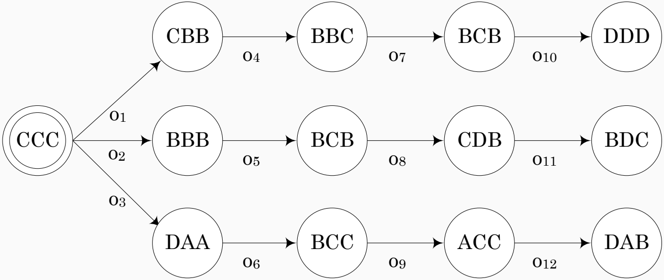
Constrained to finding a sequence of actions which achieves a state with ***maximal utility***

Logistics Task - Search Space Example



Logistics Task - Search Space Example





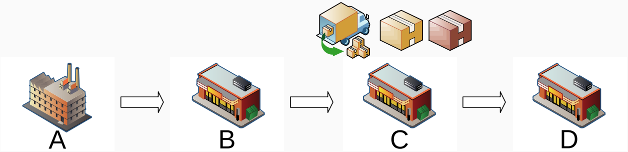
Formalism

Classical Planning

$$\Pi = \langle \mathbf{V}, \mathbf{s_0}, G; \mathbf{A}, \mathbf{c} \rangle$$

Oversubscription Planning (OSP)

$$\Pi = \langle \mathbf{V}, \mathbf{s_0}, u; \mathbf{A}, \mathbf{c}, b \rangle$$



State Variables

- $V = \{v_1, \ldots, v_n\}$.
- $S = dom(v_1) \times \cdots \times dom(v_n)$.

Initial State

- $s_0 \in S$.

Actions

- $A = \{a_1, \ldots, a_n\}$.
- Each $a \in A$ represented by a pair;
  $\langle \text{pre}(a), \text{eff}(a), c \rangle$.
- $\text{pre}(a), \text{eff}(a)$ are partial assignments to V.

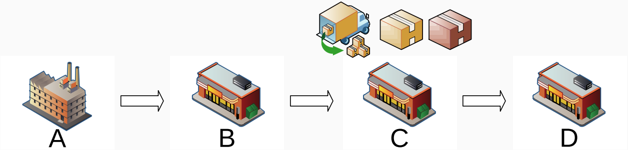
Formalism

Classical Planning

$$\Pi = \langle \mathbf{V}, \mathbf{s_0}, G; \mathbf{A}, \mathbf{c} \rangle$$

Oversubscription Planning (OSP)

$$\Pi = \langle \mathbf{V}, \mathbf{s_0}, u; \mathbf{A}, \mathbf{c}, b \rangle$$



Initial State Representation

- v_{truck1} = at(truck1, loc-A)
- v_{pkg1} = at(pkg1, loc-A)
- v_{pkg2} = at(pkg2, loc-A)

Action Representation

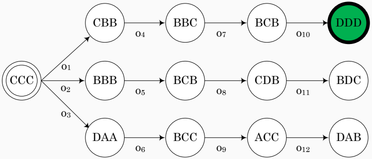
- a = *load-truck pkg1 truck1 loc-A*
- pre(a) = {at(pkg1, loc-A)}
- eff(a) = {in(pkg1, truck1)}
- $c(a) = 1$

Hard Goals vs. Utility&Budget

Classical Planning

$$\Pi = \langle V, s_0, \mathbf{G}; A, c \rangle$$

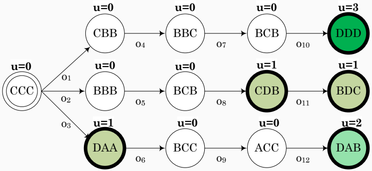
- G is a partial assignment on state variables
- $G \subseteq V$.



Oversubscription Planning (OSP)

$$\Pi = \langle V, s_0, \mathbf{u}; A, c, \mathbf{b} \rangle$$

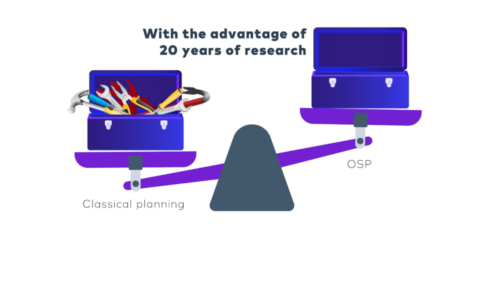
- u is a **state utility function** $u : S \to \mathbb{R}$.
- $b \in \mathbb{R}^{0+}$ is an **cost budget** allowed for the task over **actions**.



Recent Advances in Heuristic Search OSP

Smith, 2004

- presented the OSP problem

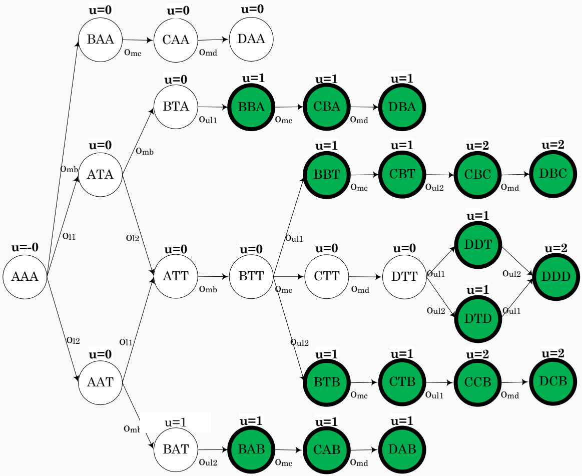


Mirkis and Domshlak, 2014

- Logical landmarks for goal reachability ("Bring me Something")
- Budget Reduction

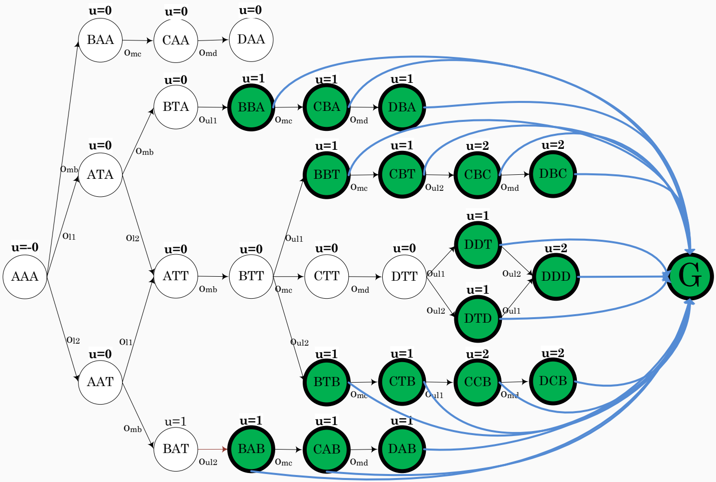
"Bring me Something" Landmarks in OSP - Logistics Task

Budget = 4



"Bring me Something" Landmarks - Compilation

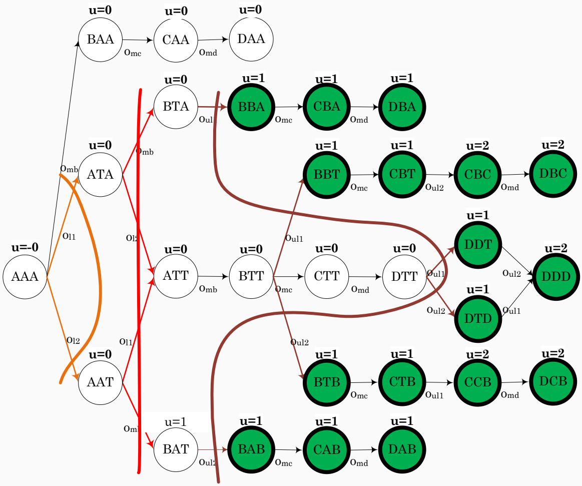
Budget = 4



"Bring me Something" Landmarks - Compilation

Budget = 4

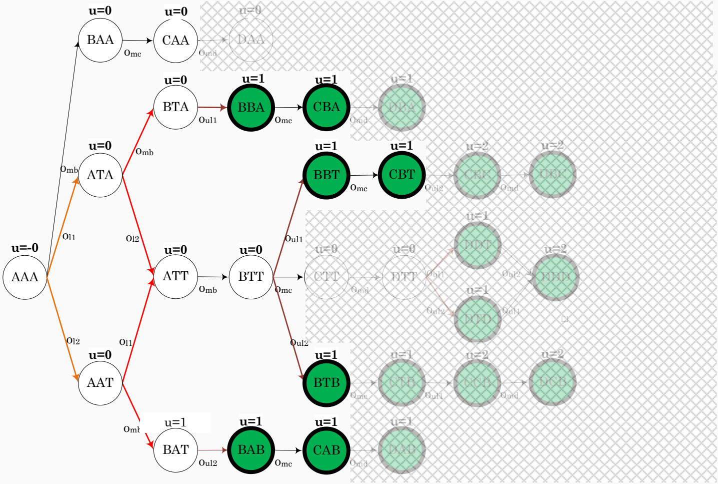
$$\boxed{(a_1 \vee a_2) \wedge (a_3 \vee a_4 \vee a_5 \vee a_6) \wedge (a_8 \vee a_9 \vee a_{10} \vee a_{11} \vee a_{12} \vee a_{13})}$$



"Bring me Something" Landmarks - Budget Reduction

Discounted Actions & Reduced Budget = ~~4~~ 1

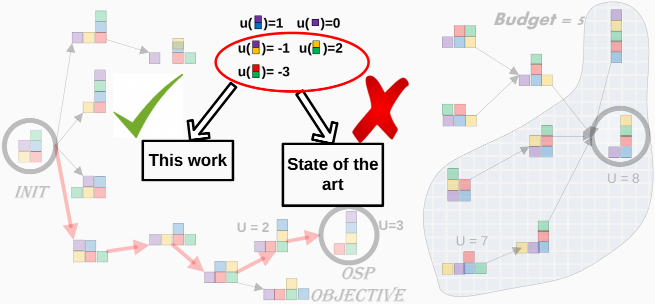
$$\boxed{(a_1 \vee a_2) \wedge (a_3 \vee a_4 \vee a_5 \vee a_6) \wedge (a_8 \vee a_9 \vee a_{10} \vee a_{11} \vee a_{12} \vee a_{13})}$$



The Gap in Heuristic OSP Search

Limited Expressivity

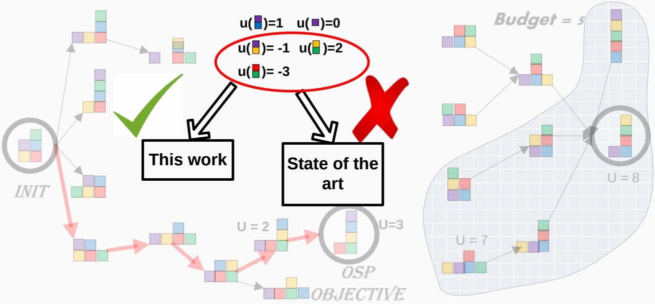
- 0-binary utility functions
- non-negative utility



Bridge The Gap in Heuristic OSP Search

Arbitrary Additive Utility Functions

- non 0-binary utility functions
- negative utility functions



Motivation & Challenges

Why Negative Values Effects Need a Special Treatment?

Classical Planing & Non-Negative OSP

Strive to collect the goals (or valuable facts) which are a partial assignment to state variables

Why Negative Values Effects Need a Special Treatment?

General OSP

Consider an entire state

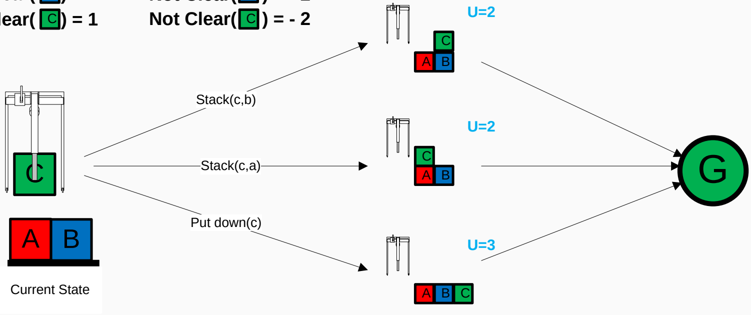
Avoiding collection of facts carrying negative values

In State Negative Interactions Between Facts

When the focus is on goal or valuable facts, we have 3 goal states

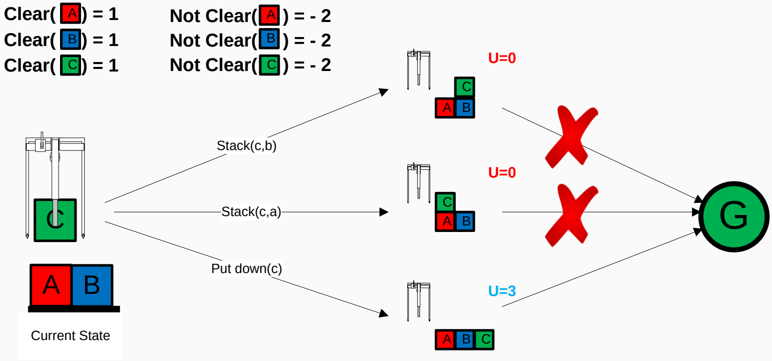
Clear(A) = 1
Clear(B) = 1
Clear(C) = 1

Not Clear(A) = - 2
Not Clear(B) = - 2
Not Clear(C) = - 2



In State Negative Interactions Between Facts

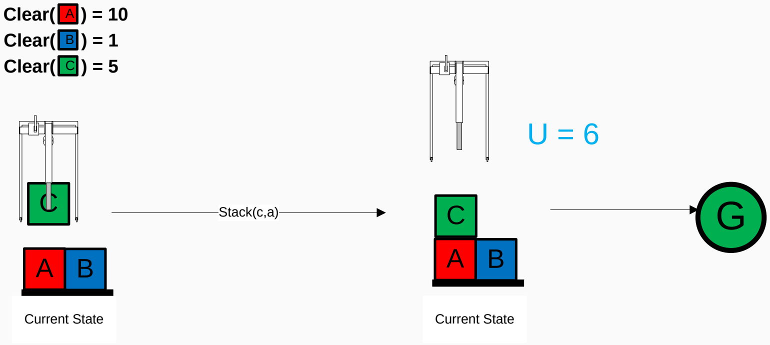
When considering a wider set of facts, we stay with one goal state



Good is not Enough

Negative Interactions Between States

When the focus is just on target state, we have a goal

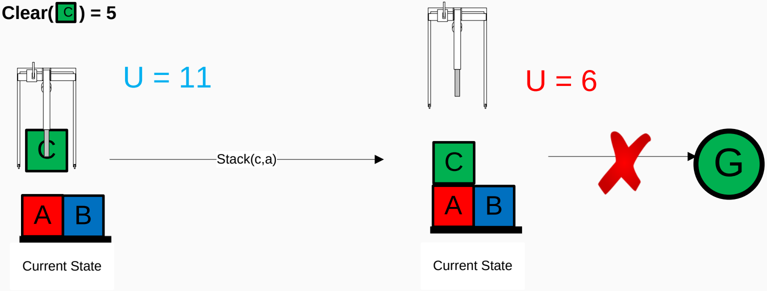


In OSP we want Better

Negative Interactions Between States

When we consider also action origin state, it seems to be less attractive

Clear(A) = 10
Clear(B) = 1
Clear(C) = 5



Relative Evaluation of Achievements

Focus on **improving** state-variables rather than on **collecting** valuable facts:

- **characterization** of the **process** of searching for an improvement
  - we define the **utility of actions**
  - exploiting descriptiveness of actions
  - discover **properties** of optimal plans on the level of **actions** and **sequences of actions**
  - synergistic combination of these properties allows handling **arbitrary additive utility** functions over multi-valued variables
  - define an effective search space

Relative Evaluation of Achievements

Focus on **improving** state-variables rather than on **collecting** valuable facts:

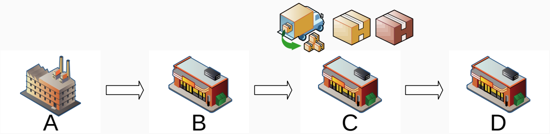
- **characterization** of the **process** of searching for an improvement
- Relative evaluation of plans (eliminate dependence on initial state)

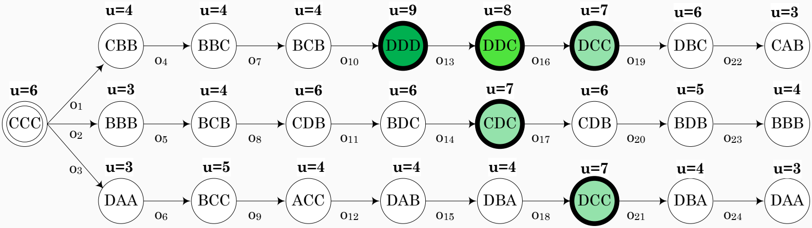
  dynamic reference point for improvement

  optimal and **satisficing** planning
- treat real-world **investment scenarios**
- domain and utility setting independent

Optimal Plan Properties

High Level Overview





Utility Function Definition for Actions

Definition

For an OSP action a, the ***net utility*** of a is

$$u_a^{net}(a) = \sum_{v \in \mathcal{V}(\mathsf{eff}(a))} [u_v(\mathsf{eff}(a)[v]) - u_v(\mathsf{pre}(a)[v])]$$

Utility Function Definition for Actions

Definition

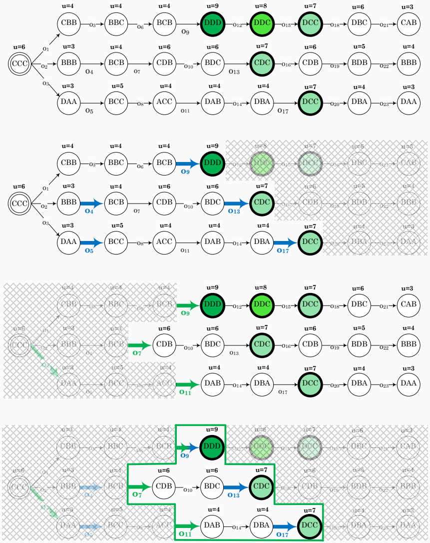
For an OSP action a, the ***net utility*** of a is

$$u_a^{net}(a) = \sum_{v \in \mathcal{V}(\mathsf{eff}(a))} [u_v(\mathsf{eff}(a)[v]) - u_v(\mathsf{pre}(a)[v])]$$

Captures State Dependencies

- Traces total benefit of an action with respect to its origin state
- In OSP, goal could be achieved just through an action with total positive net value

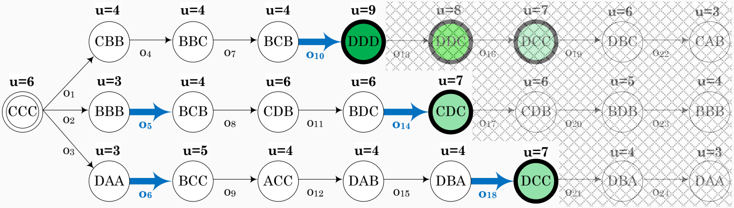
High Level Overview - Improving Approach



Net Positive Suffix Termination

For each plan π, there is plan π' that:

- ends with a net positive utility value action a_{last},
- is at most as costly as π, and
- is at least as valuable as π



Improving Approach

Dynamic Reference Point to Improve

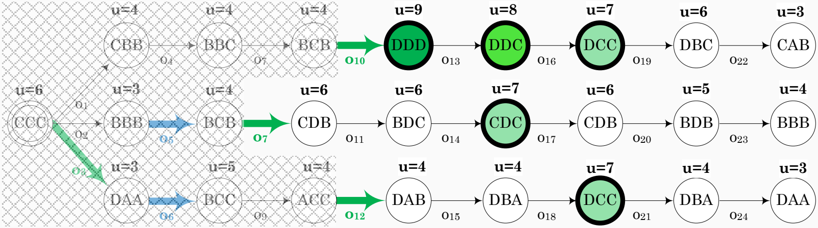
Given an OSP task $\Pi = \langle V, s_0, u; A, c, b\rangle$ and a state s, let **improved(s)** be the set of all propositions $\langle v/d\rangle \in \mathcal{D}$ such that $u_v(d) > u_v(s[v])$.

Gross Positive Actions

Given an OSP task $\Pi = \langle V, s_0, u; A, c, b\rangle$ and a state s, the **gross positive actions** relative to s $A_{gPos}(s) \subseteq A$ of Π are $A_{gPos}(s) = \{a \in A, \text{eff}(a) \cap \text{improved}(s) \neq \emptyset\}$.

Gross Positive Prefix Lemma

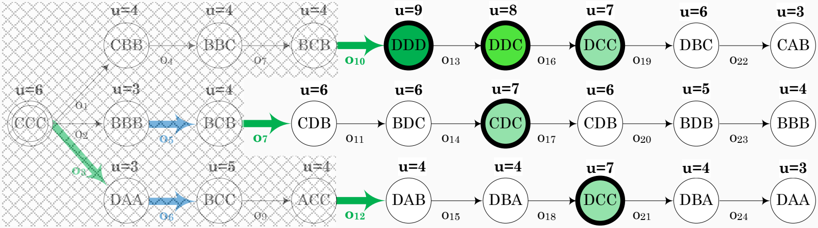
Gross Positive Prefix



Maintained Achievements Lemma

Each improving plan over a given state

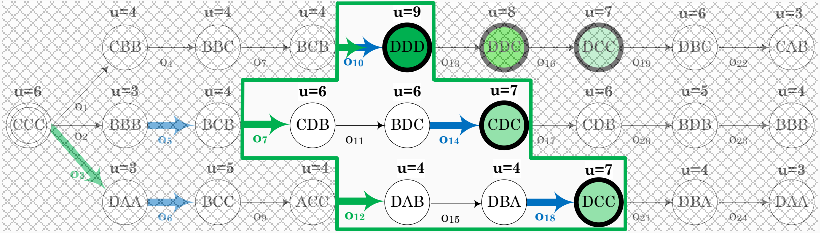
- starts with a gross positive action a_{gPos}, and
- a valuable fact achieved by a_{gPos} is maintained along the suffix.



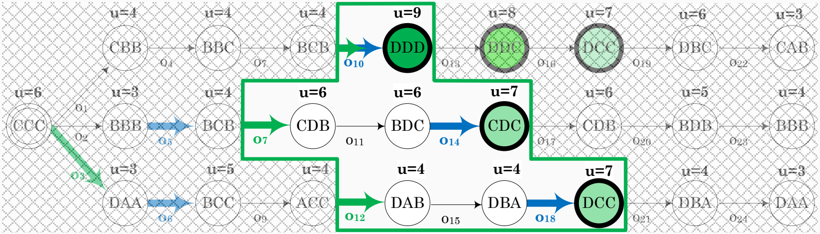
The Window of Opportunity Theorem

synergistic criteria of optimal plan

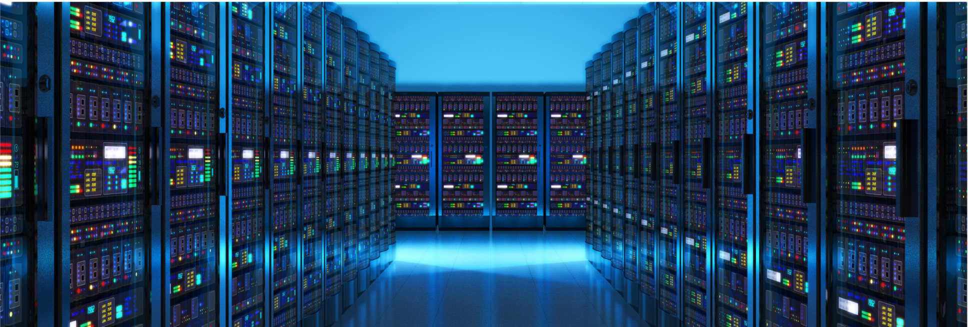
- starts with a gross positive actions
- temporal maintained achievement
- terminates with a net positive action



A Novel Landmarks Generation Approach

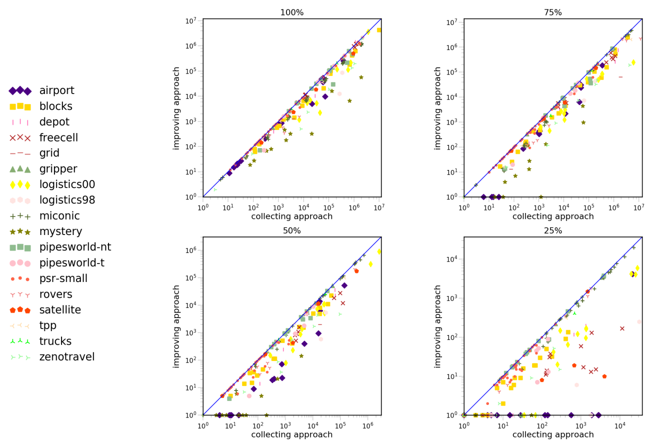


Empirical Evaluation



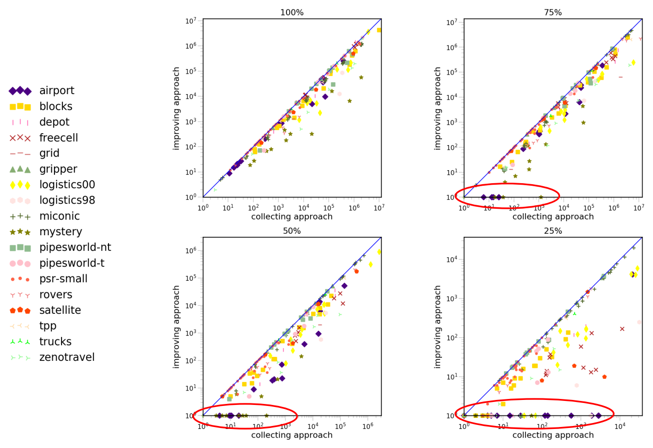
Improving vs. Collecting Approach - Expanded Nodes

Utility setting $u(dom(v)) \in \{0, 1, 2\}$



Improving vs. Collecting Approach - Solved Without a Search

Utility setting $u(dom(v)) \in \{0, 1, 2\}$



Improving vs. Collecting Approach - Reduced Budget

Utility setting $u(dom(v)) \in \{\{0, 1, 2\}\}$

	Reduced Budget		Discounted Actions		Effectiveness Score	
	Improving	*Collecting*	*Improving*	*Collecting*	*Improving*	*Collecting*
airport (17)	**484**	21	76318	**20375**	**157.68**	970.24
blocks (16)	**154**	18	215	**142**	**1.40**	7.89
depot (3)	**32**	3	569	**21**	17.78	**7.00**
freecell (11)	**54**	11	**5851**	15331	**108.35**	1393.73
grid (2)	**26**	15	2132	**1059**	82.00	**70.60**
gripper (4)	**60**	4	**172**	224	**2.87**	56.00
logistics00 (11)	**189**	11	212	**118**	**1.12**	10.73
logistics98 (3)	**40**	3	114	**40**	**2.85**	13.33
miconic (40)	**586**	121	2892	**1691**	**4.94**	13.98
mystery (11)	**49**	15	5142	**2122**	**104.94**	141.47
pipesw-nt (12)	**39**	28	**14751**	18500	**378.23**	660.71
pipesw-t (6)	**12**	6	9016	**5664**	**751.33**	944.00
psr-small (47)	**167**	55	2330	**1534**	**13.95**	27.89
rovers (6)	**44**	14	**380**	456	**8.64**	32.57
satellite (4)	**46**	10	264	**120**	**5.74**	12.00
tpp (6)	**62**	25	**155**	158	**2.50**	6.32
trucks (2)	**26**	7	**602**	931	**23.15**	133.00
zenotravel (8)	**57**	8	1213	**407**	**21.28**	50.88
total (209)	**2127**	375	122328	**68893**	**57.51**	183.71

Negative Utility Setting

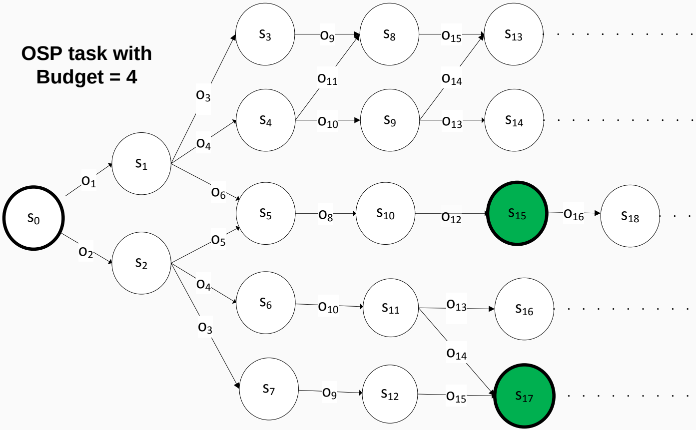
Utility setting $u(dom(v)) \in \{\{-1, 0, 1\}\}$

	25%			50%			75%			100%		
	Expanded	*Solved*	*Time*	*Expanded*	*Solved*	*Time*	*Expanded*	*Solved*	*Time*	*Expanded*	*Solved*	*Time*
airport (15)	0	15(15)	47.76	44599	15(6)	74.02	151860	15(6)	149.11	203372	15(0)	69.38
blocks (17)	405	17(4)	21.17	24379	17(1)	32.09	1945724	17(1)	144.81	14218611	17(1)	706.26
depot (3)	25	3(2)	11.42	16928	3(1)	17.38	382154	3(0)	43.46	13900	2(0)	8.88
freecell (5)	48	5(1)	77.51	2463	5(0)	122.81	14644	5(0)	132.95	24733	5(0)	141.51
grid (2)	0	2(2)	35.49	2021	2(1)	36.76	1142555	2(0)	111.88	1409996	2(0)	399.31
gripper (3)	821	3(1)	3.27	10258	3(0)	6.84	34214	3(0)	10.25	60207	3(0)	14.25
logistics00 10)	1043	10(2)	14.41	54470	10(0)	22.84	479866	10(0)	40.51	1854165	10(0)	108.12
logistics98 (2)	9	2(1)	1.97	10957	2(0)	3.98	94961	2(0)	8.23	327929	2(0)	21.95
miconic (40)	12746	40(15)	44.23	639797	40(6)	134.43	4748596	40(1)	425.22	11337706	40(0)	933.18
mystery (11)	0	11(11)	49.48	160	11(10)	51.71	12516	11(5)	67.39	395511	11(0)	121.13
pipesw-nt 11)	250	11(2)	48.19	21490	11(0)	99.09	290079	11(0)	178.4	2248028	11(0)	677.66
pipesw-t (5)	71	5(1)	45.68	6386	5(0)	71.05	95533	5(0)	88.71	356736	5(0)	115.91
rovers (4)	15	4(1)	2.6	7563	4(0)	5.58	2637	4(0)	5.18	10181	4(0)	7.21
satellite (4)	30	4(2)	4.53	13949	4(0)	9.97	187176	4(0)	24.52	1013108	4(0)	104.54
tpp (5)	0	5(5)	2.61	1932	5(1)	3.54	14455	5(1)	3.35	28278	5(0)	4.49
trucks (2)	730	2(1)	2.92	21946	2(0)	6	77382	2(0)	9.62	82969	2(0)	10.88
zenotravel (7)	34	7(4)	11.94	10594	7(1)	19.94	229144	7(1)	30.14	1477266	7(1)	85.49
total (146)	16227	146(70)	425.18	882892	146(27)	718.03	9903496	146(15)	1473.72	35062696	145(2)	3630.15


THANK YOU

Examples

Implications on Search Space

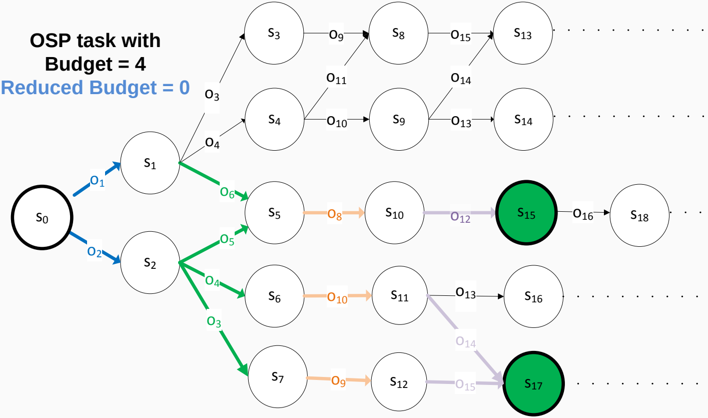


Implications on Search Space

$$(a_1 \vee a_2) \wedge (a_3 \vee a_4 \vee a_5 \vee a_6) \wedge (a_8 \vee a_9 \vee a_{10}) \wedge (a_{12} \vee a_{14} \vee a_{15})$$

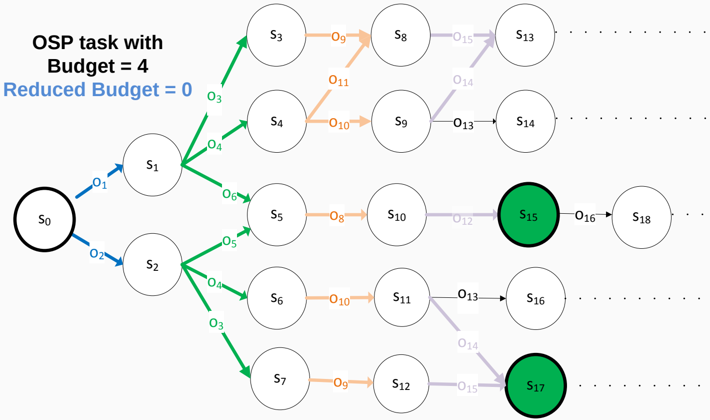
OSP task with Budget = 4

Reduced Budget = 0



Implications on Search Space

$$(a_1 \lor a_2) \land (a_3 \lor a_4 \lor a_5 \lor a_6) \land (a_8 \lor a_9 \lor a_{10}) \land (a_{12} \lor a_{14} \lor a_{15})$$

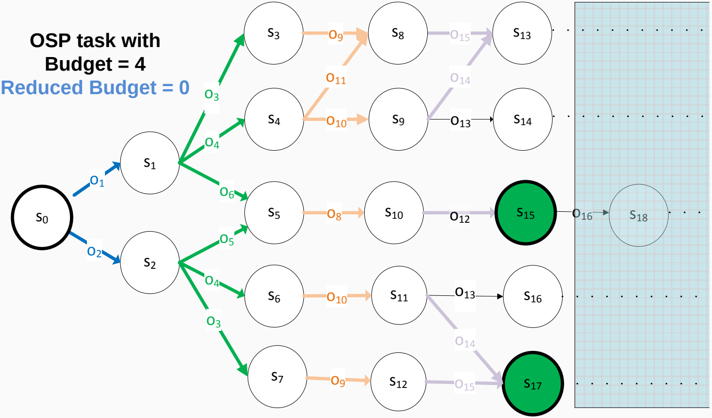


Implications on Search Space

$$(a_1 \vee a_2) \wedge (a_3 \vee a_4 \vee a_5 \vee a_6) \wedge (a_8 \vee a_9 \vee a_{10}) \wedge (a_{12} \vee a_{14} \vee a_{15})$$

OSP task with Budget = 4

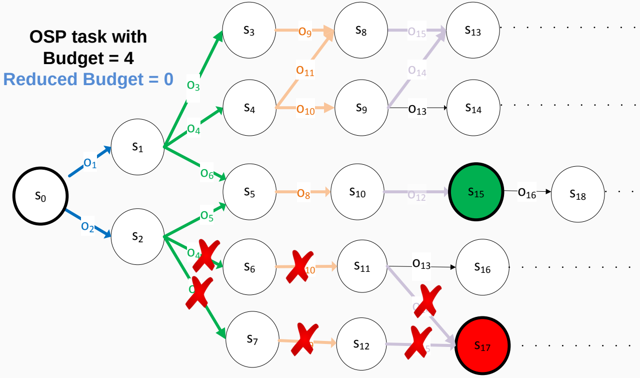
Reduced Budget = 0



Implications on Search Space

$$\boxed{\cancel{(a_1 \vee a_2) \wedge (a_3 \vee a_4 \vee a_5 \vee a_6) \wedge (a_8 \vee a_9 \vee a_{10}) \wedge (a_{12} \vee a_{14} \vee a_{15})}}$$

$$\boxed{(a_1 \vee a_2) \wedge (a_5 \vee a_6) \wedge a_8 \wedge a_{13}}$$

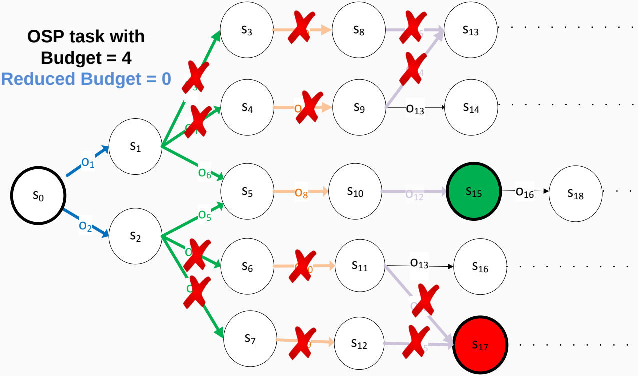


Implications on Search Space

$$\cancel{(a_1 \vee a_2) \wedge (a_3 \vee a_4 \vee a_5 \vee a_6) \wedge (a_8 \vee a_9 \vee a_{10} \vee a_{11}) \wedge (a_{14} \vee a_{15})}$$

$$\boxed{(a_1 \vee a_2) \wedge (a_5 \vee a_6) \wedge a_8 \wedge a_{13}}$$

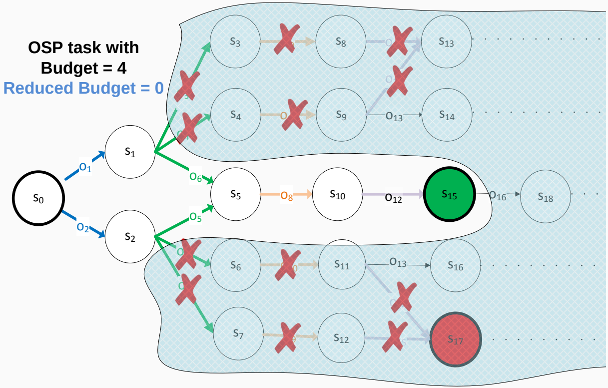
OSP task with
Budget = 4
Reduced Budget = 0



Implications on Search Space

$$\cancel{(a_1 \lor a_2) \land (a_3 \lor a_4 \lor a_5 \lor a_6) \land (a_8 \lor a_9 \lor a_{10} \lor a_{11}) \land (a_{14} \lor a_{15})}$$

$$\boxed{(a_1 \lor a_2) \land (a_5 \lor a_6) \land a_8 \land a_{13}}$$



Example - Dynamic Relative Improvement Point

$$dom(v_1) = \{at(pkg1, loc-A), at(pkg1, loc-B) \ldots at(pkg1, loc-H)\}$$
$$dom(v_2) = \{at(pkg2, loc-A), at(pkg2, loc-B) \ldots at(pkg2, loc-H)\}$$
.
.
.
$$dom(v_7) = \{at(pkg7, loc-A), at(pkg7, loc-B) \ldots at(pkg7, loc-H)\}$$

Dynamic Relative Improvement Point

A	170	12	8.5	0.6	-1.5	-10	-190
B	130	17	7.5	0.5	-5	-7	-140
C	140	11	8	0.2	-3.5	-5	-150
D	100	14	6.5	0.4	-3	-6	-200
E	150	10	7	0	-2.5	-4	-170
F	110	15	6	0.1	-4	-8	-180
G	120	16	5	0.7	-4.5	-9	-130
H	160	13	5.5	0.3	-2	-3	-160
	v_1	v_2	v_3	v_4	v_5	v_6	v_7

Dynamic Relative Improvement Point

170	12	8.5	0.6	-1.5	-10	-190
130	17	7.5	0.5	-5	-7	-140
140	11	8	0.2	-3.5	-5	-150
100	14	6.5	0.4	-3	-6	-200
150	10	7	0	-2.5	-4	-170
110	15	6	0.1	-4	-8	-180
120	16	5	0.7	-4.5	-9	-130
160	13	5.5	0.3	-2	-3	-160
v_1	v_2	v_3	v_4	v_5	v_6	v_7

Dynamic Relative Improvement Point

170	17	8.5	0.7	-1.5	-3	-130
160	16	8	0.6	-2	-4	-140
150	15	7.5	0.5	-2.5	-5	-150
140	14	7	0.4	-3	-6	-160
130	13	6.5	0.3	-3.5	-7	-170
120	12	6	0.2	-4	-8	-180
110	11	5.5	0.1	-4.5	-9	-190
100	10	5	0	-5	-10	-200
v_1	v_2	v_3	v_4	v_5	v_6	v_7

u(s*)	-39.8

Dynamic Relative Improvement Point

170	17	8.5	0.7	-1.5	-3	-130
160	16	8	0.6	-2	-4	-140
150	15	7.5	0.5	-2.5	-5	-150
140	14	7	0.4	-3	-6	-160
130	13	6.5	0.3	-3.5	-7	-170
120	12	6	0.2	-4	-8	-180
110	11	5.5	0.1	-4.5	-9	-190
100	10	5	0	-5	-10	-200
v_1	v_2	v_3	v_4	v_5	v_6	v_7

u(s*)	-35.3

Dynamic Relative Improvement Point

170	17	8.5	0.7	-1.5	-3	-130
160	16	8	0.6	-2	-4	-140
150	15	7.5	0.5	-2.5	-5	-150
140	14	7	0.4	-3	-6	-160
130	13	6.5	0.3	-3.5	-7	-170
120	12	6	0.2	-4	-8	-180
110	11	5.5	0.1	-4.5	-9	-190
100	10	5	0	-5	-10	-200
v_1	v_2	v_3	v_4	v_5	v_6	v_7

u(s*)	-20.3

Dynamic Relative Improvement Point

170	17	8.5	0.7	-1.5	-3	-130
160	16	8	0.6	-2	-4	-140
150	15	7.5	0.5	-2.5	-5	-150
140	14	7	0.4	-3	-6	-160
130	13	6.5	0.3	-3.5	-7	-170
120	12	6	0.2	-4	-8	-180
110	11	5.5	0.1	-4.5	-9	-190
100	10	5	0	-5	-10	-200
v_1	v_2	v_3	v_4	v_5	v_6	v_7

u(s*)	12

Dynamic Relative Improvement Point

170	17	8.5	0.7	-1.5	-3	-130
160	16	8	0.6	-2	-4	-140
150	15	7.5	0.5	-2.5	-5	-150
140	14	7	0.4	-3	-6	-160
130	13	6.5	0.3	-3.5	-7	-170
120	12	6	0.2	-4	-8	-180
110	11	5.5	0.1	-4.5	-9	-190
100	10	5	0	-5	-10	-200
v_1	v_2	v_3	v_4	v_5	v_6	v_7

u(s*)	18.5

Dynamic Relative Improvement Point

170	17	8.5	0.7	-1.5	-3	-130
160	16	8	0.6	-2	-4	-140
150	15	7.5	0.5	-2.5	-5	-150
140	14	7	0.4	-3	-6	-160
130	13	6.5	0.3	-3.5	-7	-170
120	12	6	0.2	-4	-8	-180
110	11	5.5	0.1	-4.5	-9	-190
100	10	5	0	-5	-10	-200
v_1	v_2	v_3	v_4	v_5	v_6	v_7

u(s*)	51.7


THANK YOU