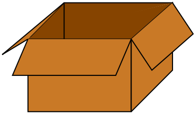
Pareto Optimal Solutions for Smoothed Analysts

Ankur Moitra, IAS

joint work with Ryan O'Donnell, CMU

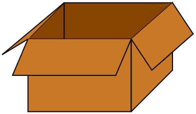
September 26, 2011

values	v_1	$\cdots$	v_i	v_{i+1}	$\cdots$	v_n
weights	w_1	$\cdots$	w_i	w_{i+1}	$\cdots$	w_n



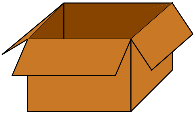
W

values	v_1	...	v_i	v_{i+1}	...	v_n
weights	w_1	...	w_i	w_{i+1}	...	w_n

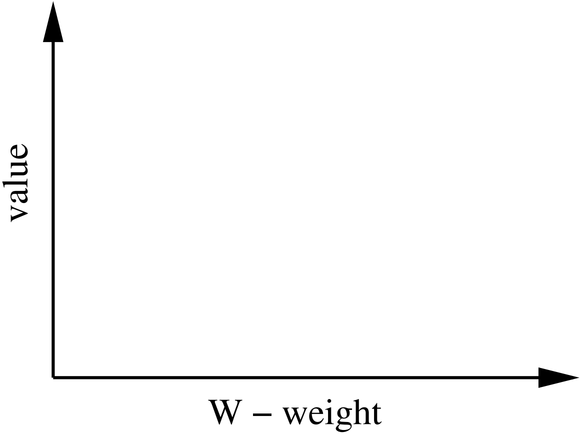


W

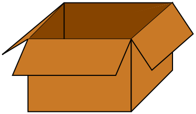
values	v_1	…	v_i	v_{i+1}	…	v_n
weights	w_1	…	w_i	w_{i+1}	…	w_n



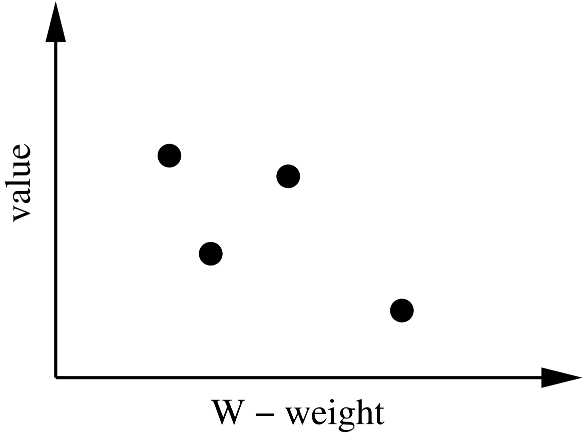
W



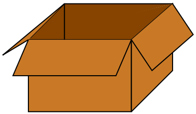
values	v_1	...	v_i	v_{i+1}	...	v_n
weights	w_1	...	w_i	w_{i+1}	...	w_n



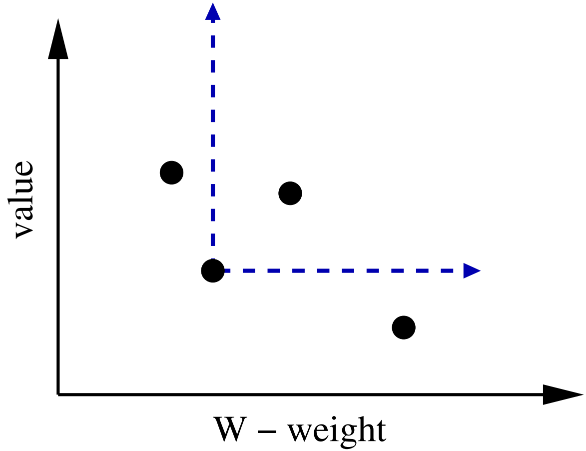
W

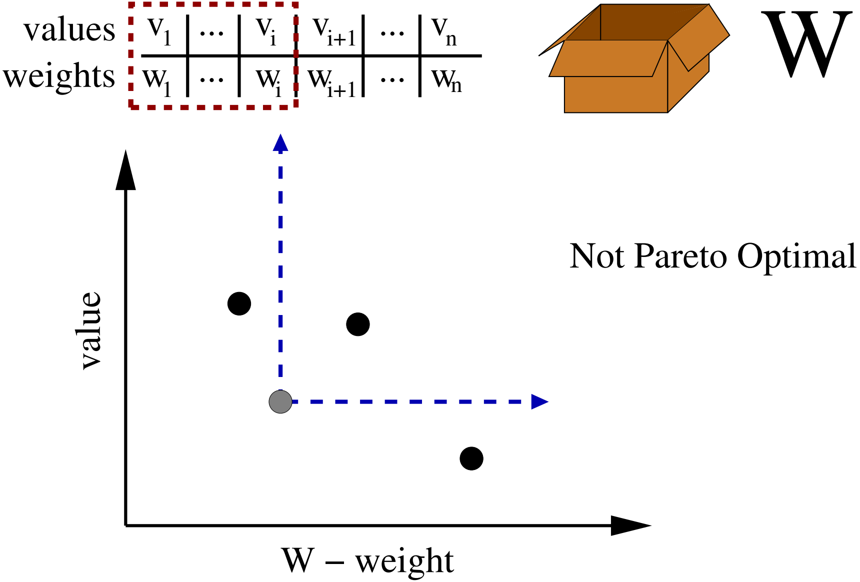


values	v_1	$\cdots$	v_i	v_{i+1}	$\cdots$	v_n
weights	w_1	$\cdots$	w_i	w_{i+1}	$\cdots$	w_n

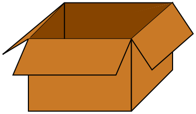


W



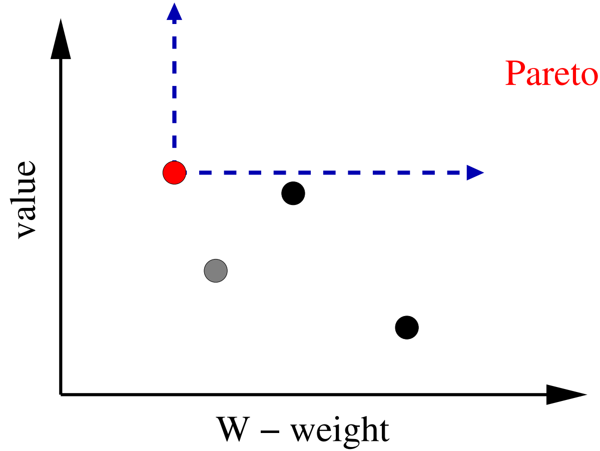

values
weights
v1 … vi vi+1 … vn
w1 … wi wi+1 … wn
W
Not Pareto Optimal
value
W – weight

values	v_1	$\cdots$	v_i	v_{i+1}	$\cdots$	v_n
weights	w_1	$\cdots$	w_i	w_{i+1}	$\cdots$	w_n

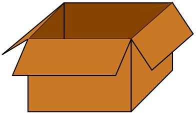


W

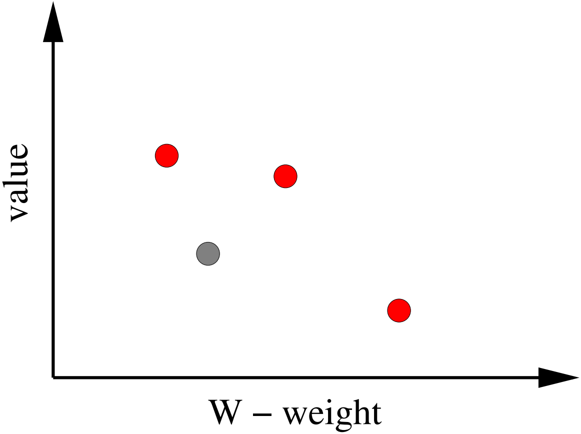
Pareto Optimal!



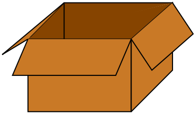
values	v_1	$\cdots$	v_i	v_{i+1}	$\cdots$	v_n
weights	w_1	$\cdots$	w_i	w_{i+1}	$\cdots$	w_n



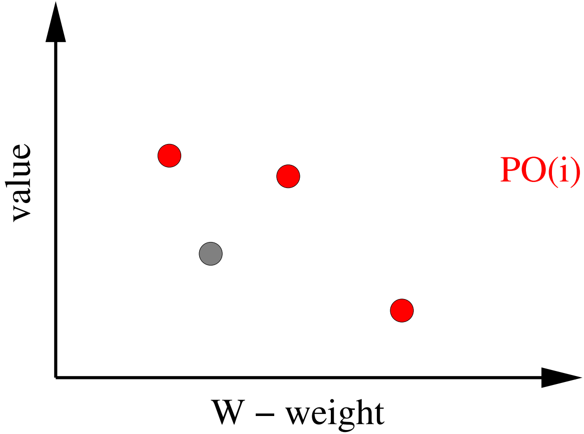
W



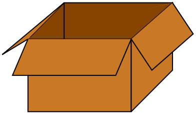
values	v_1	…	v_i	v_{i+1}	…	v_n
weights	w_1	…	w_i	w_{i+1}	…	w_n



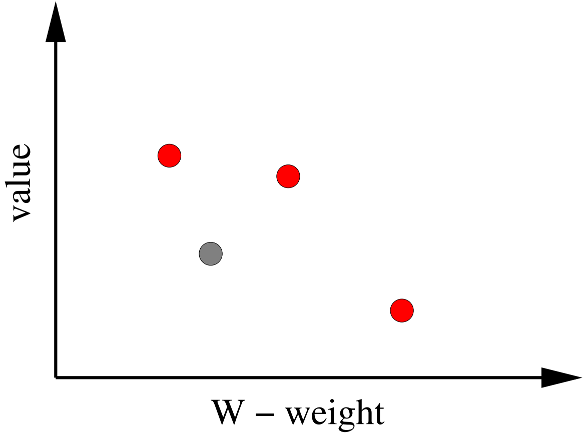
W



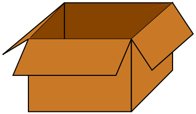
values	v_1	...	v_i	v_{i+1}	...	v_n
weights	w_1	...	w_i	w_{i+1}	...	w_n



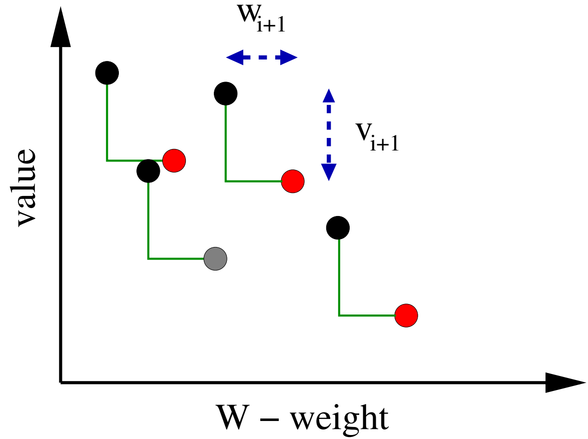
W



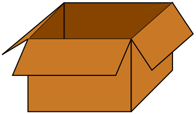
values	v_1	…	v_i	v_{i+1}	…	v_n
weights	w_1	…	w_i	w_{i+1}	…	w_n



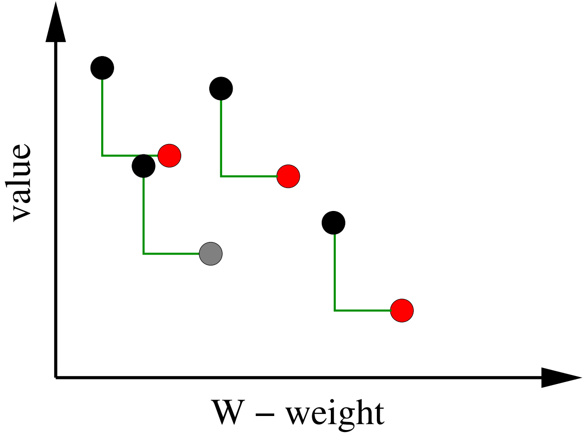
W



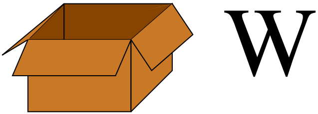
values	v_1	...	v_i	v_{i+1}	...	v_n
weights	w_1	...	w_i	w_{i+1}	...	w_n

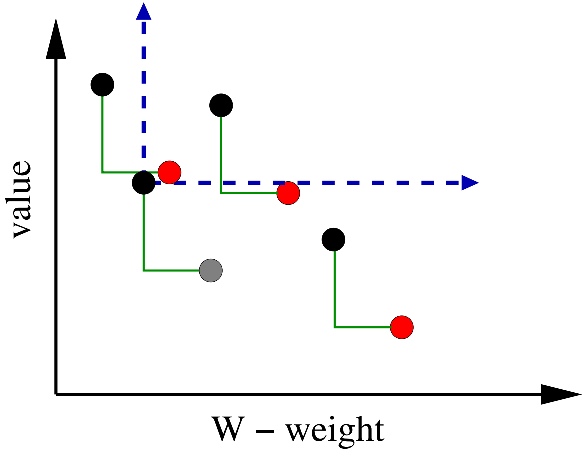


W

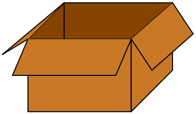


values	v_1	…	v_i	v_{i+1}	…	v_n
weights	w_1	…	w_i	w_{i+1}	…	w_n

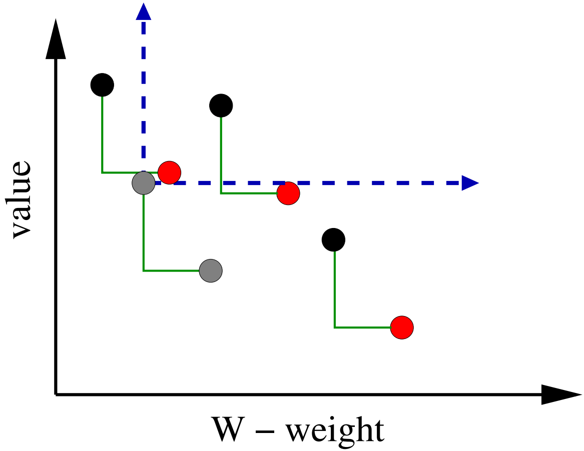




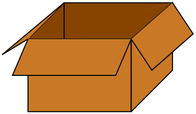
values	v_1	…	v_i	v_{i+1}	…	v_n
weights	w_1	…	w_i	w_{i+1}	…	w_n



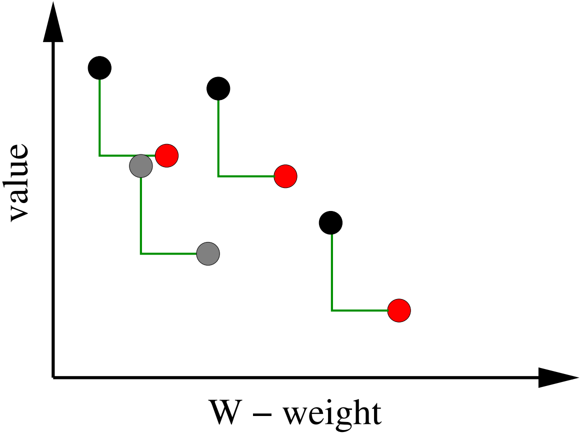
W

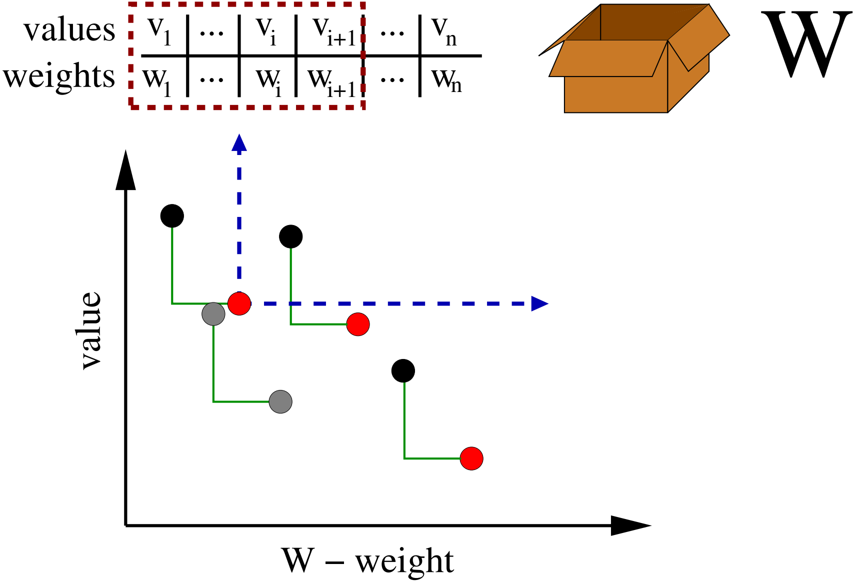


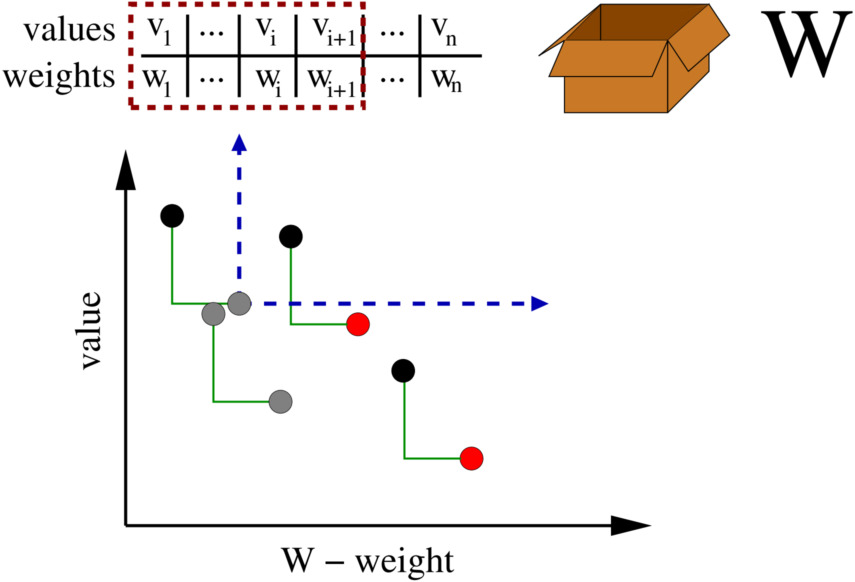
values	v_1	…	v_i	v_{i+1}	…	v_n
weights	w_1	…	w_i	w_{i+1}	…	w_n



W

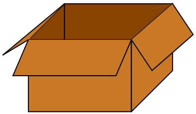



values
weights
v1 … vi vi+1 … vn
w1 … wi wi+1 … wn
W
value
W – weight

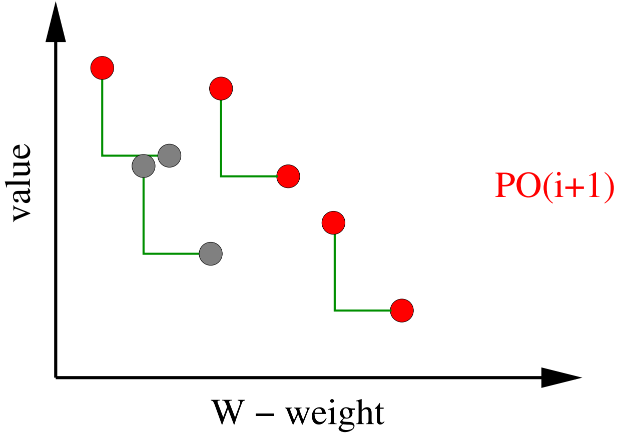


values
weights
v_1 … v_i v_{i+1} … v_n
w_1 … w_i w_{i+1} … w_n
W
value
W – weight

values	v_1	…	v_i	v_{i+1}	…	v_n
weights	w_1	…	w_i	w_{i+1}	…	w_n



W



Dynamic Programming with Lists (Nemhauser, Ullmann):

Dynamic Programming with Lists (Nemhauser, Ullmann):

- $PO(i+1)$ can be computed from $PO(i)$...

Dynamic Programming with Lists (Nemhauser, Ullmann):

- $PO(i+1)$ can be computed from $PO(i)$...

  ... in linear (in $|PO(i)|$) time

Dynamic Programming with Lists (Nemhauser, Ullmann):

- $PO(i+1)$ can be computed from $PO(i)$...

  ... in linear (in $|PO(i)|$) time

- an optimal solution can be computed from $PO(n)$

Dynamic Programming with Lists (Nemhauser, Ullmann):

- $PO(i+1)$ can be computed from $PO(i)$...

  ... in linear (in $|PO(i)|$) time

- an optimal solution can be computed from $PO(n)$

Bottleneck in many algorithms: enumerate the set of Pareto optimal solutions

Results of Beier and Vöcking (STOC 2003, STOC 2004)

Smoothed Analysis of Pareto Curves:

Results of Beier and Vöcking (STOC 2003, STOC 2004)

Smoothed Analysis of Pareto Curves:

- Polynomial bound on $|PO(i)|$s (in two dimensions) in the framework of Smoothed Analysis (Spielman, Teng)

Results of Beier and Vöcking (STOC 2003, STOC 2004)

Smoothed Analysis of Pareto Curves:

- Polynomial bound on $|PO(i)|$s (in two dimensions) in the framework of Smoothed Analysis (Spielman, Teng)
- Knapsack has polynomial smoothed complexity:

Results of Beier and Vöcking (STOC 2003, STOC 2004)

Smoothed Analysis of Pareto Curves:

- Polynomial bound on $|PO(i)|$s (in two dimensions) in the framework of Smoothed Analysis (Spielman, Teng)
- Knapsack has polynomial smoothed complexity:
  - first NP-hard problem that is (smoothed) easy

Results of Beier and Vöcking (STOC 2003, STOC 2004)

Smoothed Analysis of Pareto Curves:

- Polynomial bound on $|PO(i)|$s (in two dimensions) in the framework of Smoothed Analysis (Spielman, Teng)
- Knapsack has polynomial smoothed complexity:
  - first NP-hard problem that is (smoothed) easy
  - generalizes long line of results on random instances

Capturing Tradeoffs

Capturing Tradeoffs

Question

What if the precise objective function (of a decision maker) is unknown?

Capturing Tradeoffs

Question

What if the precise objective function (of a decision maker) is unknown?

e.g. travel planning: prefer lower fare, shorter trip, fewer transfers

Capturing Tradeoffs

Question

What if the precise objective function (of a decision maker) is unknown?

e.g. travel planning: prefer lower fare, shorter trip, fewer transfers

Question

Can we algorithmically help a decision maker?

Capturing Tradeoffs

Question

What if the precise objective function (of a decision maker) is unknown?

e.g. travel planning: prefer lower fare, shorter trip, fewer transfers

Question

Can we algorithmically help a decision maker?

Pareto curves capture tradeoffs among competing objectives

Proposition

Only useful approach if Pareto curves are small

Proposition

Only useful approach if Pareto curves are small

Confirmed empirically e.g. Müller-Hannemann, Weihe: German train system

Proposition

Only useful approach if Pareto curves are small

Confirmed empirically e.g. Müller-Hannemann, Weihe: German train system

Question

Why should we expect Pareto curves to be small?

Proposition

Only useful approach if Pareto curves are small

Confirmed empirically e.g. Müller-Hannemann, Weihe: German train system

Question

Why should we expect Pareto curves to be small?

Caveat: Smoothed Analysis is not a complete explanation

The Model

Adversary chooses:

The Model

Adversary chooses:

- $Z \subset \{0,1\}^n$ (e.g. spanning trees, Hamiltonian cycles)

The Model

Adversary chooses:

- $Z \subset \{0,1\}^n$ (e.g. spanning trees, Hamiltonian cycles)
- an adversarial objective function $OBJ_1 : Z \to \Re^+$

The Model

Adversary chooses:

- $Z \subset \{0, 1\}^n$ (e.g. spanning trees, Hamiltonian cycles)
- an adversarial objective function $OBJ_1 : Z \rightarrow \Re^+$
- $d - 1$ linear objective functions

$$OBJ_i(\vec{x}) = [w_1^i, w_2^i, \ldots w_n^i] \cdot \vec{x}$$

The Model

Adversary chooses:

- $Z \subset \{0,1\}^n$ (e.g. spanning trees, Hamiltonian cycles)
- an adversarial objective function $OBJ_1 : Z \to \Re^+$
- $d-1$ linear objective functions

$$OBJ_i(\vec{x}) = [w_1^i, w_2^i, \ldots w_n^i] \cdot \vec{x}$$

... each w_j^i is a random variable on $[-1,+1]$ – density is bounded by ϕ

History

Let PO be the set of Pareto optimal solutions...

History

Let PO be the set of Pareto optimal solutions...

- [Beier, Vöcking STOC 2003] ($d = 2$), $E[|PO|] = O(n^5\phi)$

History

Let PO be the set of Pareto optimal solutions...

- [Beier, Vöcking STOC 2003] $(d = 2)$, $E[|PO|] = O(n^5\phi)$
- [Beier, Vöcking STOC 2004] $(d = 2)$, $E[|PO|] = O(n^4\phi)$

History

Let PO be the set of Pareto optimal solutions...

- [Beier, Vöcking STOC 2003] ($d = 2$), $E[|PO|] = O(n^5\phi)$
- [Beier, Vöcking STOC 2004] ($d = 2$), $E[|PO|] = O(n^4\phi)$
- [Beier, Röglin, Vöcking IPCO 2007] ($d = 2$), $E[|PO|] = O(n^2\phi)$ (tight)

History

Let PO be the set of Pareto optimal solutions...

- [Beier, Vöcking STOC 2003] ($d = 2$), $E[|PO|] = O(n^5 \phi)$
- [Beier, Vöcking STOC 2004] ($d = 2$), $E[|PO|] = O(n^4 \phi)$
- [Beier, Röglin, Vöcking IPCO 2007] ($d = 2$), $E[|PO|] = O(n^2 \phi)$ (tight)
- [Röglin, Teng, FOCS 2009] $E[|PO|] = O((n\sqrt{\phi})^{f(d)})$
  ... where $f(d) = 2^{d-1}(d+1)!$

History

Let PO be the set of Pareto optimal solutions...

- [Beier, Vöcking STOC 2003] ($d = 2$), $E[|PO|] = O(n^5 \phi)$
- [Beier, Vöcking STOC 2004] ($d = 2$), $E[|PO|] = O(n^4 \phi)$
- [Beier, Röglin, Vöcking IPCO 2007] ($d = 2$), $E[|PO|] = O(n^2 \phi)$ (tight)
- [Röglin, Teng, FOCS 2009] $E[|PO|] = O((n\sqrt{\phi})^{f(d)})$
  ... where $f(d) = 2^{d-1}(d+1)$!
- [Dughmi, Roughgarden, FOCS 2010] any FPTAS can be transformed to a truthful in expectation FPTAS

Our Results

Theorem

$E[|PO|] \leq 2 \cdot (4\phi d)^{d(d-1)/2} n^{2d-2}$

Our Results

Theorem

$E[|PO|] \leq 2 \cdot (4\phi d)^{d(d-1)/2} n^{2d-2}$

...answers a conjecture of Teng

Our Results

Theorem

$E[|PO|] \leq 2 \cdot (4\phi d)^{d(d-1)/2} n^{2d-2}$

...answers a conjecture of Teng

[Bently et al, JACM 1978]: 2^n points sampled from a d-dimensional Gaussian,

$$E[|PO|] = \Theta(n^{d-1})$$

square factor difference necessary for $d = 2$

On Smoothed Analysis

On Smoothed Analysis

Bounds are notoriously pessimistic (e.g. Simplex)

On Smoothed Analysis

Bounds are notoriously pessimistic (e.g. Simplex)

Method of Analysis:

On Smoothed Analysis

Bounds are notoriously pessimistic (e.g. Simplex)

Method of Analysis:

- Define a "bad" event

  ...that you can blame if your algorithm runs slowly

Bounds are notoriously pessimistic (e.g. Simplex)

Method of Analysis:

- Define a "bad" event

  ...that you can blame if your algorithm runs slowly
- Prove this event is rare

On Smoothed Analysis

Bounds are notoriously pessimistic (e.g. Simplex)

Method of Analysis:

- Define a "bad" event

  ...that you can blame if your algorithm runs slowly
- Prove this event is rare

Proposition

Randomness is your friend!

Our Approach

Our Approach

Goal

"Define" bad events using as little randomness as possible

Our Approach

Goal

"Define" bad events using as little randomness as possible

... "conserve" randomness, save for analysis

Our Approach

Goal

"Define" bad events using as little randomness as possible

... "conserve" randomness, save for analysis

Challenge

Events become convoluted (to say the least!)

Our Approach

Goal

"Define" bad events using as little randomness as possible

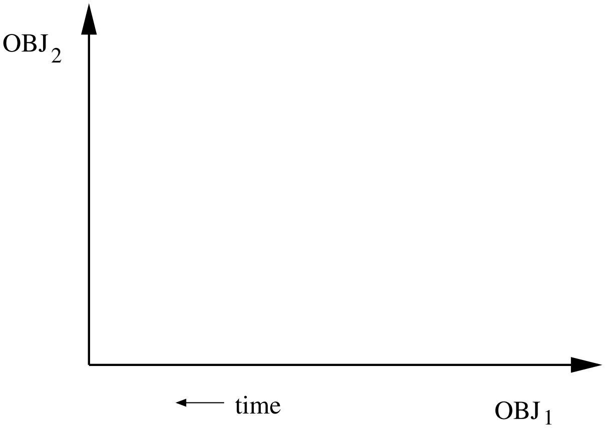
... "conserve" randomness, save for analysis

Challenge

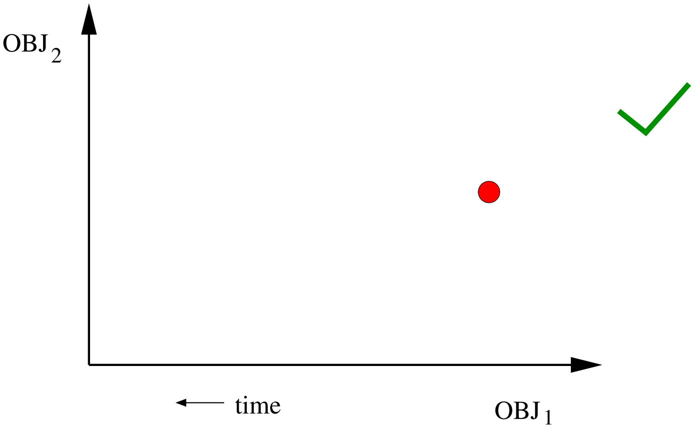
Events become convoluted (to say the least!)

We give an algorithm to define these events

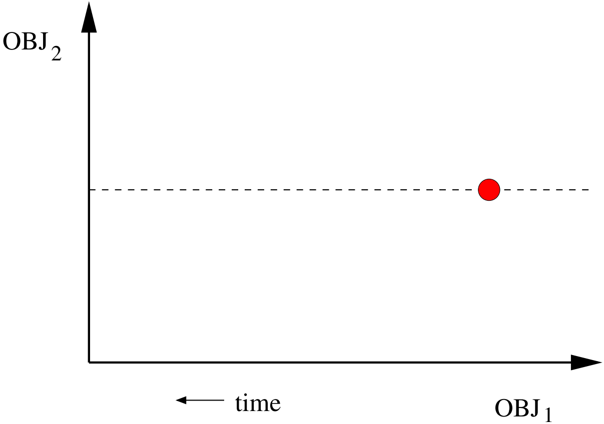
An Alternative Condition



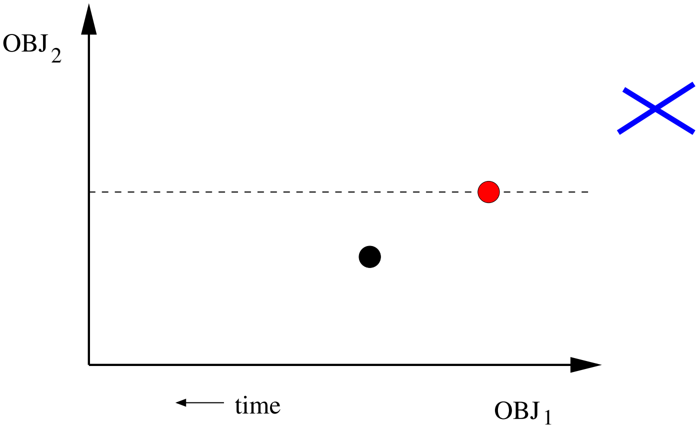
An Alternative Condition



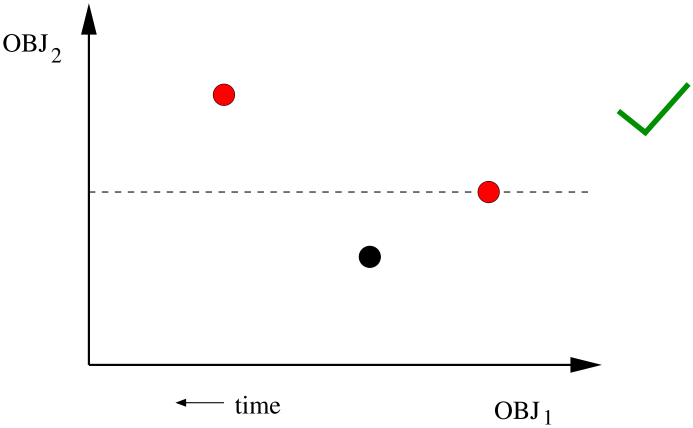
An Alternative Condition



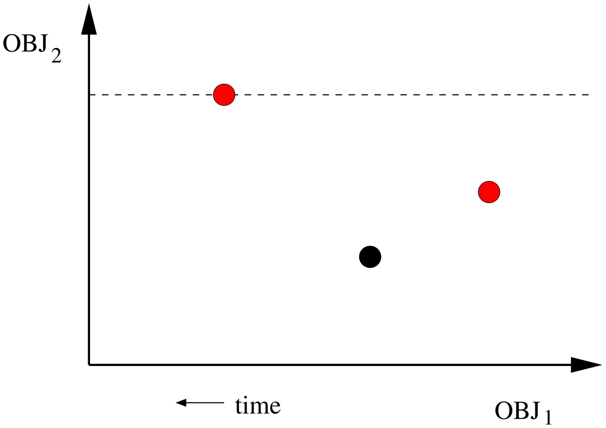
An Alternative Condition



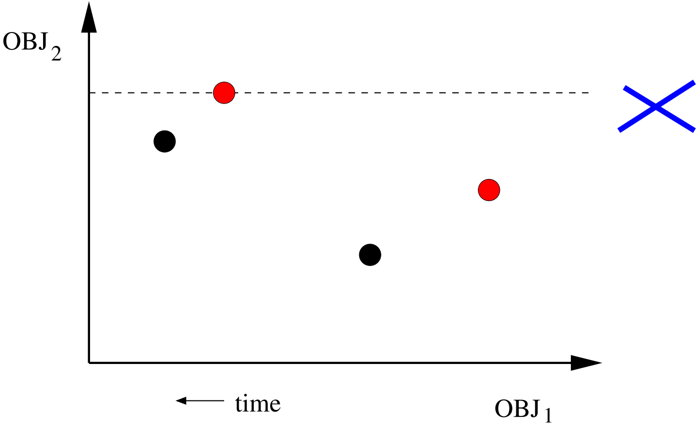
An Alternative Condition



An Alternative Condition



An Alternative Condition



Method of Proof

Goal

Count the (expected) number of Pareto optimal solutions

Method of Proof

Goal

Count the (expected) number of Pareto optimal solutions

- Define a complete family of events
  ... for each Pareto optimal solution at least one (unique) event occurs

Method of Proof

Goal

Count the (expected) number of Pareto optimal solutions

- Define a complete family of events
  ... for each Pareto optimal solution at least one (unique) event occurs
- Then bound the expected number of events

Method of Proof

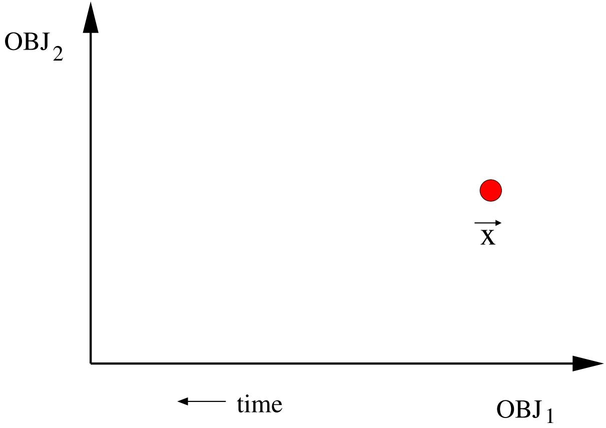
Goal

Count the (expected) number of Pareto optimal solutions

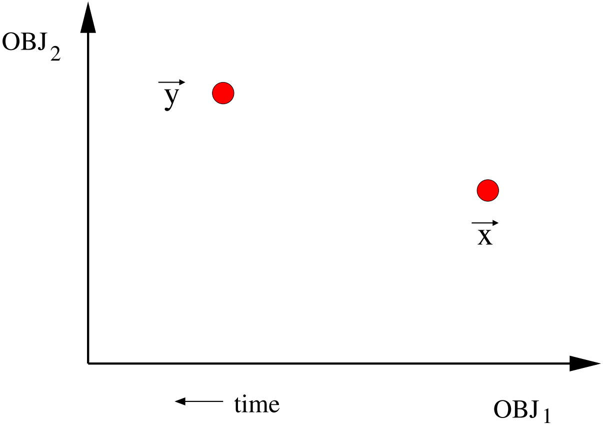
- Define a complete family of events
  ... for each Pareto optimal solution at least one (unique) event occurs
- Then bound the expected number of events

The key is in the definition

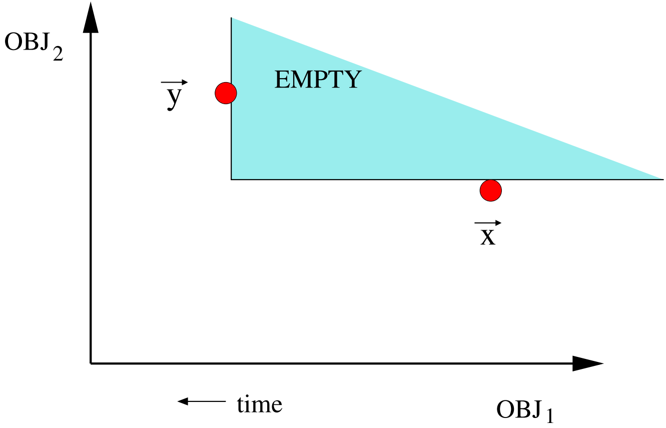
The Common Case (Beier, Röglin, Vöcking)



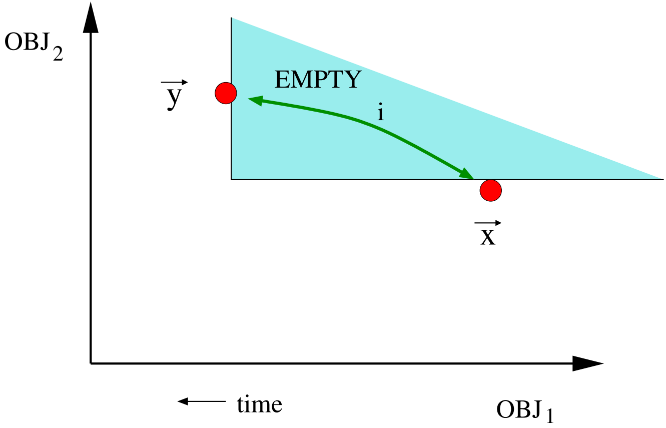
The Common Case (Beier, Röglin, Vöcking)



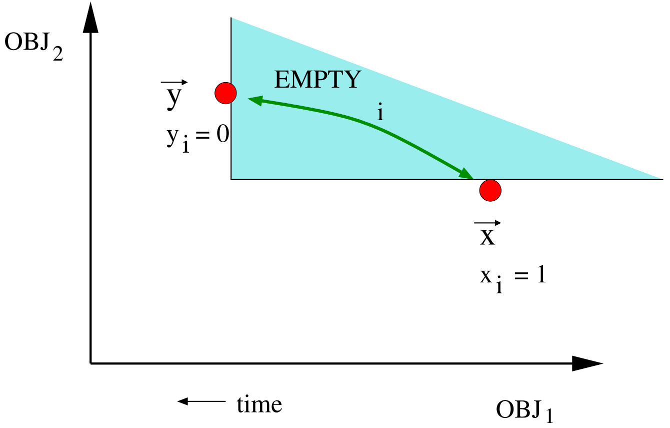
The Common Case (Beier, Röglin, Vöcking)



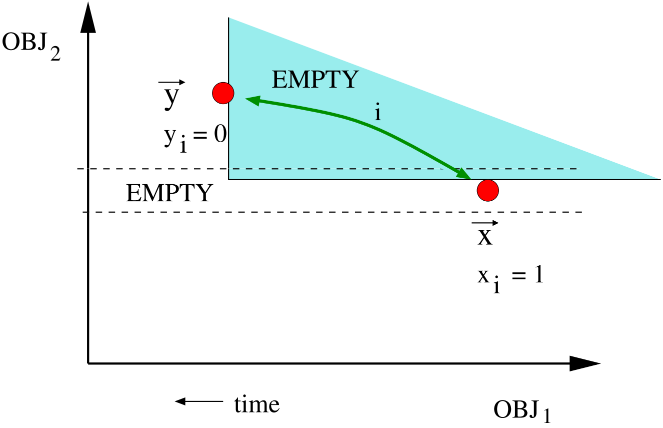
The Common Case (Beier, Röglin, Vöcking)



The Common Case (Beier, Röglin, Vöcking)



The Common Case (Beier, Röglin, Vöcking)



Complete Family of Events

Let $I = (a, b)$ be an interval of $[-n, +n]$ of width ϵ

Complete Family of Events

Let $I = (a, b)$ be an interval of $[-n, +n]$ of width ϵ

Let E be the event that there is an $\vec{x} \in PO$:

Complete Family of Events

Let $I = (a, b)$ be an interval of $[-n, +n]$ of width ϵ

Let E be the event that there is an $\vec{x} \in PO$:

- $OBJ_2(\vec{x}) \in I$

Complete Family of Events

Let $I = (a, b)$ be an interval of $[-n, +n]$ of width ϵ

Let E be the event that there is an $\vec{x} \in PO$:

- $OBJ_2(\vec{x}) \in I$
- $\vec{y}$ have the smallest $OBJ_1(\vec{y})$ among all points for which $OBJ_2(\vec{y}) > b$

Complete Family of Events

Let $I = (a, b)$ be an interval of $[-n, +n]$ of width ϵ

Let E be the event that there is an $\vec{x} \in PO$:

- $OBJ_2(\vec{x}) \in I$
- $\vec{y}$ have the smallest $OBJ_1(\vec{y})$ among all points for which $OBJ_2(\vec{y}) > b$
- $\vec{y}_i = 0, \vec{x}_i = 1$

Complete Family of Events

Let $I = (a, b)$ be an interval of $[-n, +n]$ of width ϵ

Let E be the event that there is an $\vec{x} \in PO$:

- $OBJ_2(\vec{x}) \in I$
- $\vec{y}$ have the smallest $OBJ_1(\vec{y})$ among all points for which $OBJ_2(\vec{y}) > b$
- $\vec{y}_i = 0, \vec{x}_i = 1$

Claim

$Pr[E] \leq \epsilon\phi$

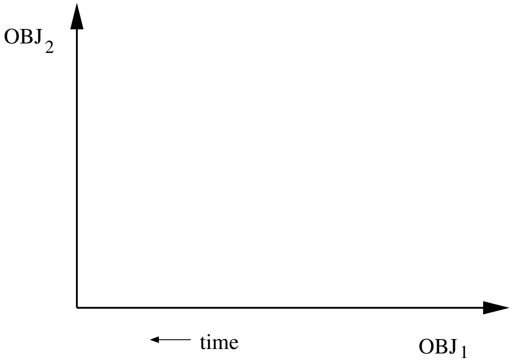
Analysis

OBJ_1 : arbitrary

OBJ_2 : [?? , ?? , ...?? ....??] $\vec{x}$

E

Z



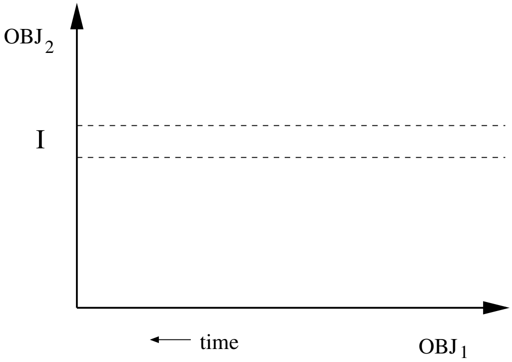
Analysis

OBJ_1 : arbitrary

OBJ_2 : [?? , ?? , ...?? ....??] $\vec{x}$

E

Z



Analysis

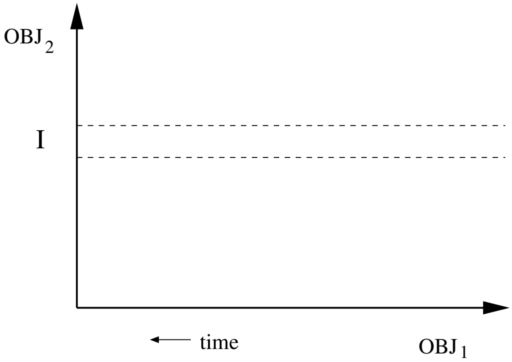
OBJ_1 : arbitrary

OBJ_2 : $[w_1, w_2, \ldots ?? \ldots w_n]\ \vec{x}$

index i

E

Z



Analysis

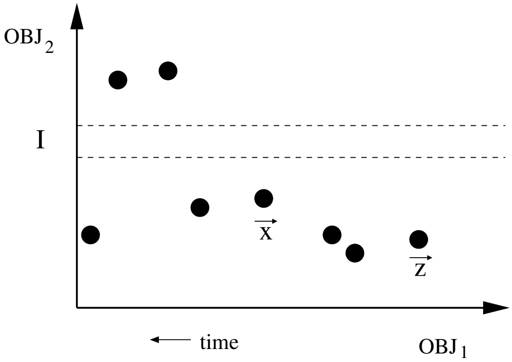
OBJ_1 : arbitrary

OBJ_2 : $[w_1, w_2, ..??....w_n]\ \vec{x}$

index i

E

Z

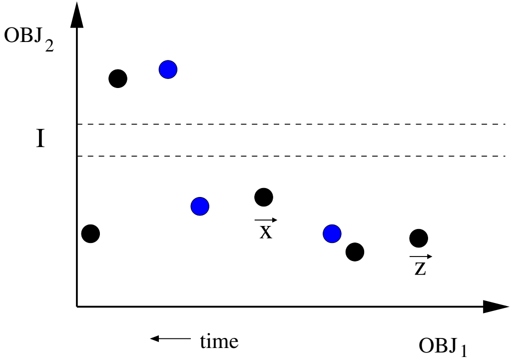


Analysis

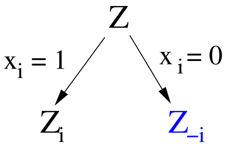
OBJ_1 : arbitrary

OBJ_2 : $[w_1, w_2, ..??....w_n]\ \vec{x}$

index i



E

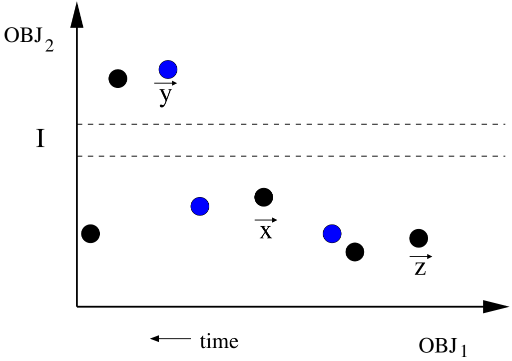


Analysis

OBJ_1: arbitrary

OBJ_2: $[w_1, w_2, .. ?? ....w_n]\ \vec{x}$

index i



E

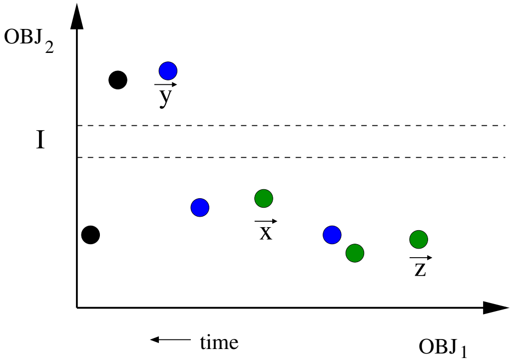
$$Z \xrightarrow{x_i = 1} Z_i \qquad Z \xrightarrow{x_i = 0} Z_{-i}$$

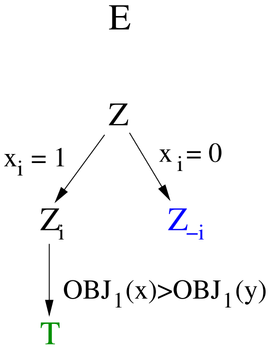
Analysis

OBJ_1: arbitrary

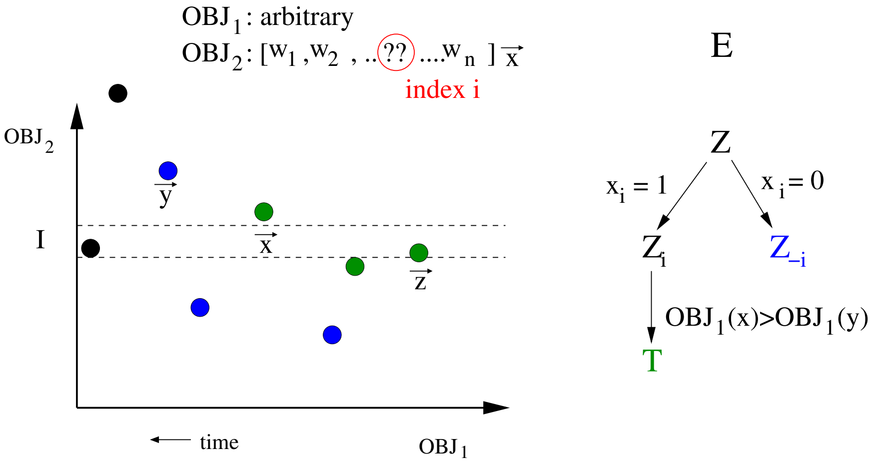
OBJ_2: $[w_1, w_2, \ldots ?? \ldots w_n]\ \vec{x}$

index i



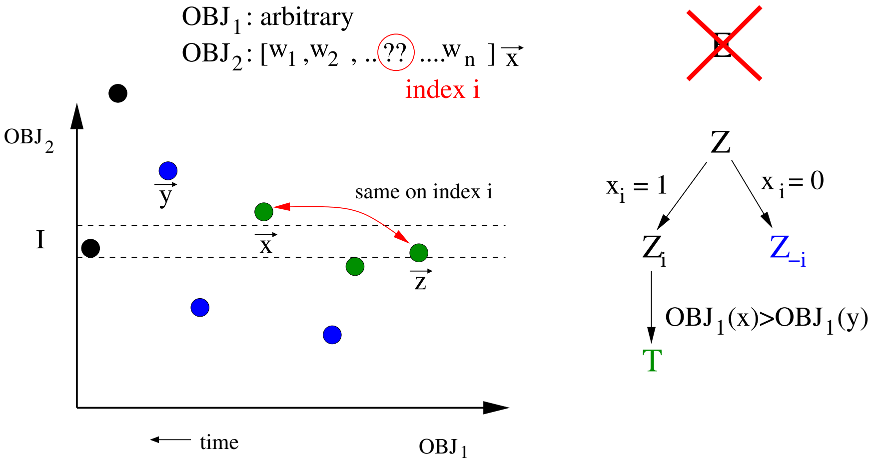


Analysis



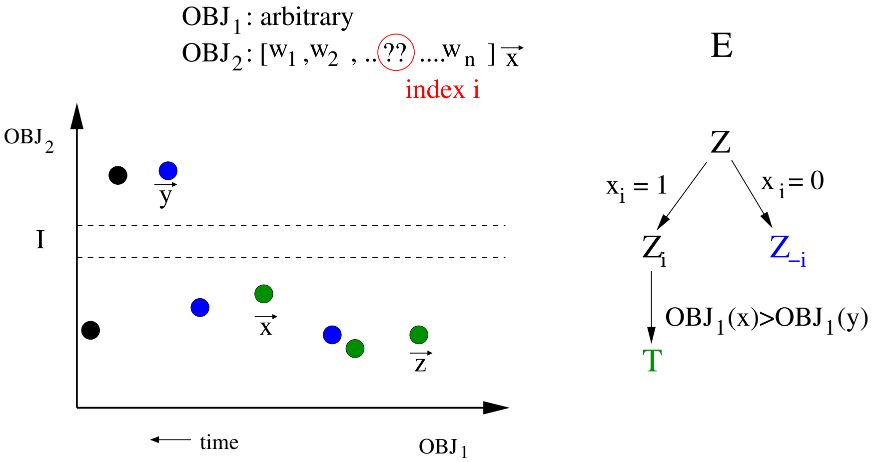
OBJ 1 : arbitrary
OBJ 2 : [w1 ,w2 , ..?? ....w n] x
index i
OBJ 2
I
y
x
z
time
OBJ 1
E
Z
x i = 1
x i = 0
Z i
Z −i
OBJ 1 (x)>OBJ 1 (y)
T

Analysis



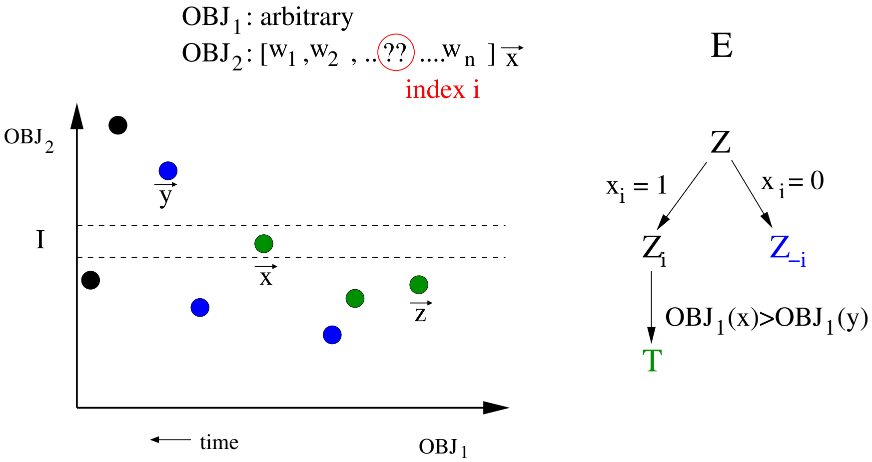
OBJ 1 : arbitrary
OBJ 2 : [w 1 ,w 2 , .. ?? ....w n] x
index i
OBJ 2
y
same on index i
I
x
z
time
OBJ 1
Z
x i = 1
x i = 0
Z i
Z −i
OBJ 1 (x)>OBJ 1 (y)
T

Analysis



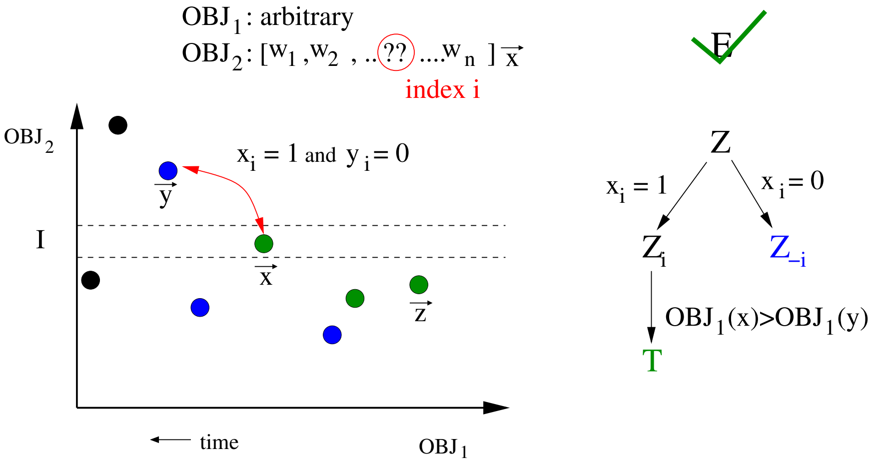
OBJ 1 : arbitrary
OBJ 2 : [w1 ,w2 , ..?? ....w n] x
index i
E
OBJ 2
y
I
x
z
time
OBJ 1
Z
xi = 1
x i = 0
Zi
Z−i
OBJ 1 (x)>OBJ 1 (y)
T

Analysis



OBJ 1 : arbitrary
OBJ 2 : [w1 ,w2 , ..?? ....w n] x
index i
OBJ 2
y
I
x
z
time
OBJ 1
E
Z
xi = 1
x i = 0
Zi
Z−i
OBJ 1(x)>OBJ 1(y)
T

Analysis



OBJ 1 : arbitrary
OBJ 2 : [w 1 ,w 2 , ..?? ....w n] x
index i
OBJ 2
x i = 1 and y i = 0
y
I
x
z
time
OBJ 1
E
Z
x i = 1
x i = 0
Z i
Z −i
OBJ 1 (x)>OBJ 1 (y)
T

A Re-interpretation

A Re-interpretation

Implicitly defined a Transcription Algorithm:

A Re-interpretation

Implicitly defined a Transcription Algorithm:

Question

Given a Pareto optimal solution x, which event should we blame?

A Re-interpretation

Implicitly defined a Transcription Algorithm:

Question

Given a Pareto optimal solution x, which event should we blame?

Blame event E: interval I, index i, x_i and y_i

A Re-interpretation

Implicitly defined a Transcription Algorithm:

Question

Given a Pareto optimal solution x, which event should we blame?

Blame event E: interval I, index i, x_i and y_i

Transcription Algorithm

- *Input: values of r.v.s (and Pareto optimal x)*
- *Output: interval I, index i, x_i and y_i*

A Re-interpretation

Suppose output is event E: interval I, index i, x_i and y_i

A Re-interpretation

Suppose output is event E: interval I, index i, x_i and y_i

Question

Which solution x caused this event to occur?

A Re-interpretation

Suppose output is event E: interval I, index i, x_i and y_i

Question

Which solution x caused this event to occur?

We do not need to know w_i to determine the identity of x!

A Re-interpretation

Suppose output is event E: interval I, index i, x_i and y_i

Question

Which solution x caused this event to occur?

We do not need to know w_i to determine the identity of x!

Reverse Algorithm

- *Find y*

A Re-interpretation

Suppose output is event E: interval I, index i, x_i and y_i

Question

Which solution x caused this event to occur?

We do not need to know w_i to determine the identity of x!

Reverse Algorithm

- *Find y*
- *Then find x*

A Re-interpretation

Reverse Algorithm: Find x (without looking at w_i)

A Re-interpretation

Reverse Algorithm: Find x (without looking at w_i)

Question

Does x fall into the interval I?

A Re-interpretation

Reverse Algorithm: Find x (without looking at w_i)

Question

Does x fall into the interval I?

Hidden random variable w_i must fall into some small range

A Re-interpretation

Reverse Algorithm: Find x (without looking at w_i)

Question

Does x fall into the interval I?

Hidden random variable w_i must fall into some small range

Proposition

We can deduce a missing input to Transcription Algorithm, from just some of the inputs and outputs

A Re-interpretation

Reverse Algorithm: Find x (without looking at w_i)

Question

Does x fall into the interval I?

Hidden random variable w_i must fall into some small range

Proposition

We can deduce a missing input to Transcription Algorithm, from just some of the inputs and outputs

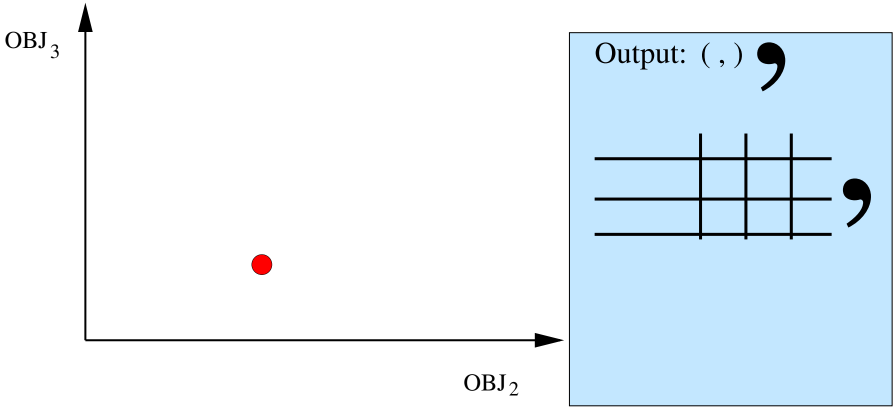
Hence each output is unlikely

Transcription for $d = 3$

OBJ_1 : adversarial

OBJ_2, OBJ_3 : linear

Z

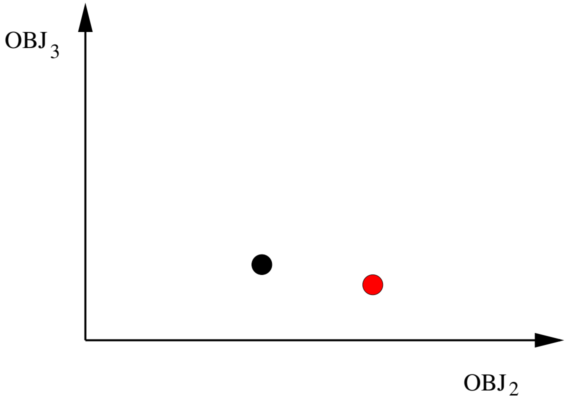


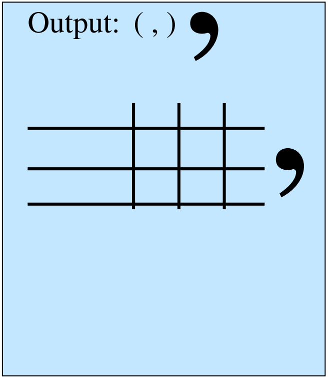
Transcription for $d = 3$

OBJ_1 : adversarial

OBJ_2, OBJ_3 : linear

Z

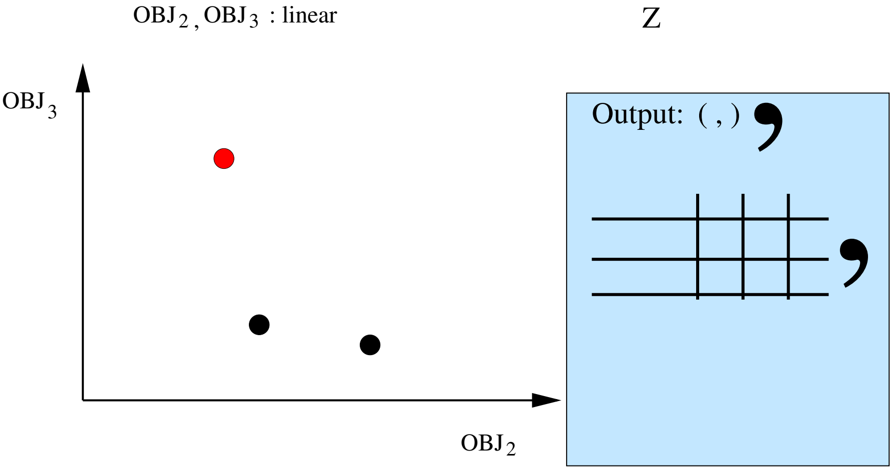




Transcription for $d = 3$

OBJ_1 : adversarial

OBJ_2, OBJ_3 : linear

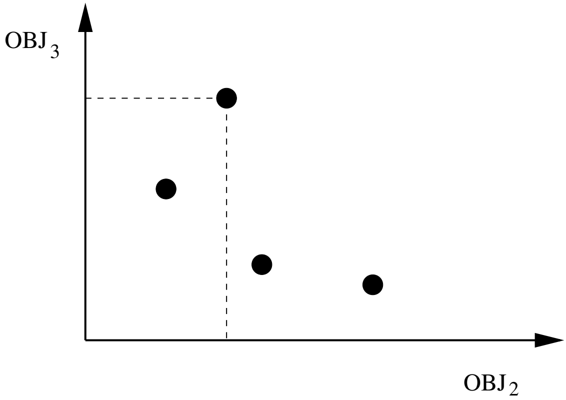


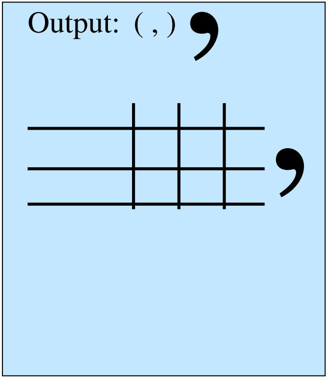
Transcription for $d = 3$

OBJ_1 : adversarial

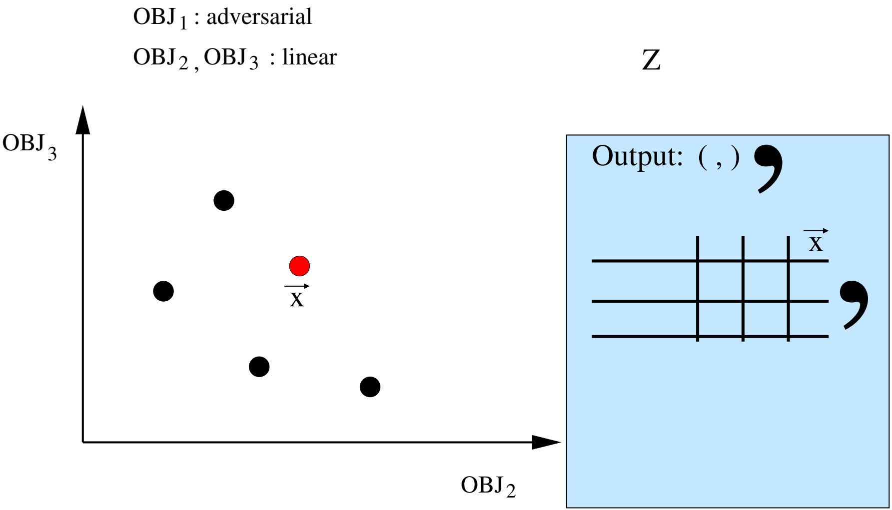
OBJ_2 , OBJ_3 : linear

Z





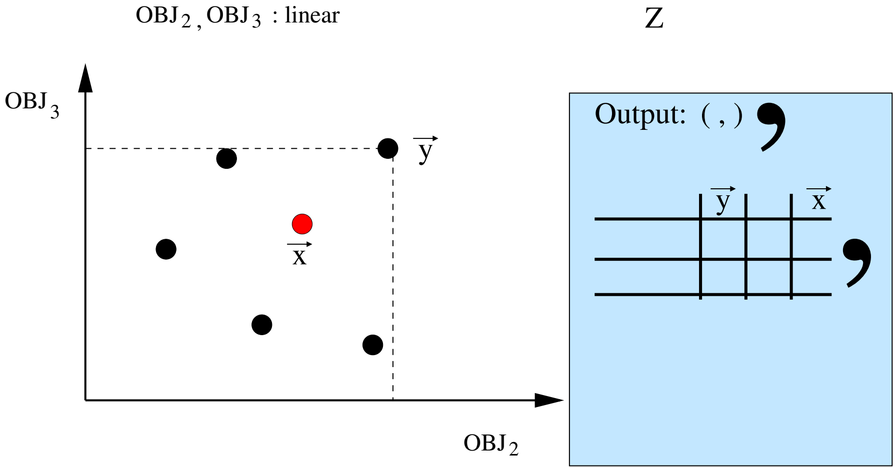
Transcription for $d = 3$



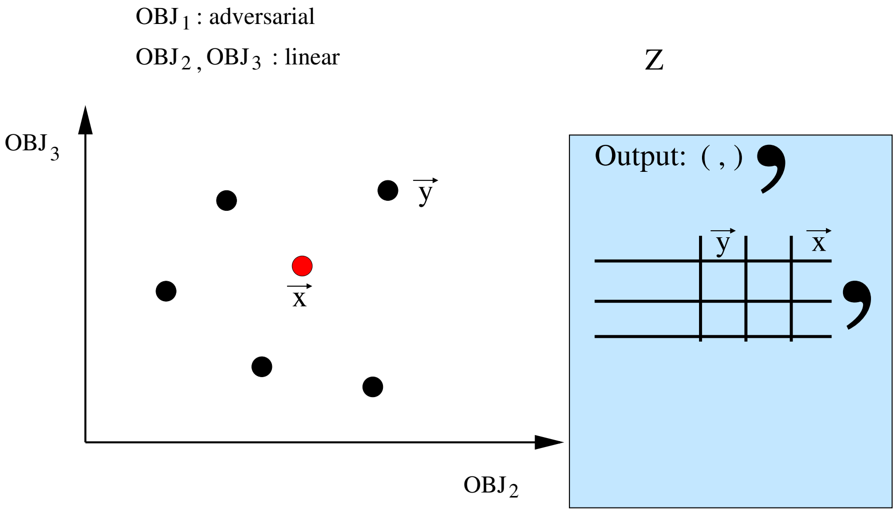
Transcription for $d = 3$

OBJ_1 : adversarial

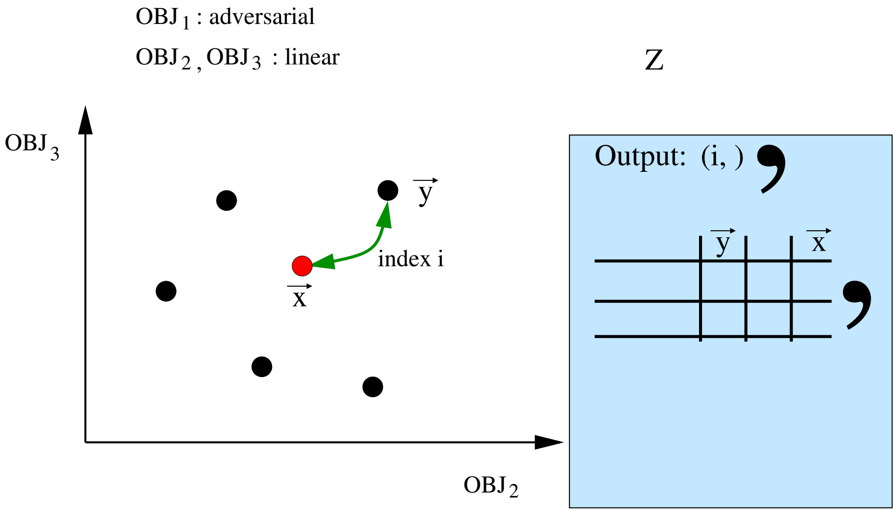
OBJ_2, OBJ_3 : linear



Transcription for $d = 3$



Transcription for $d = 3$

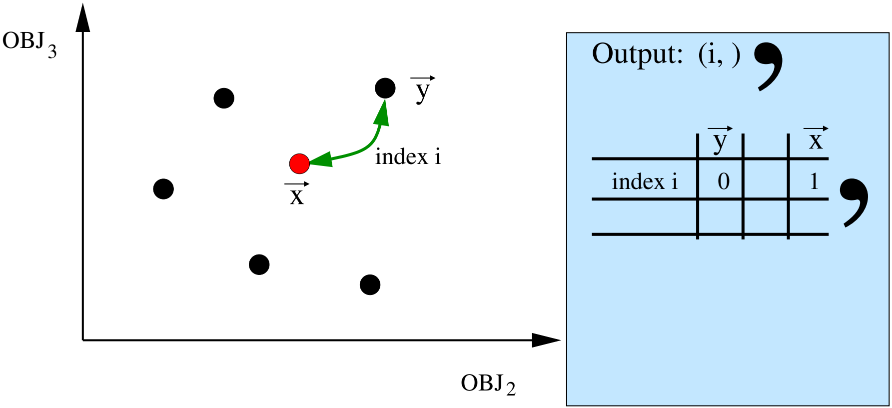


Transcription for $d = 3$

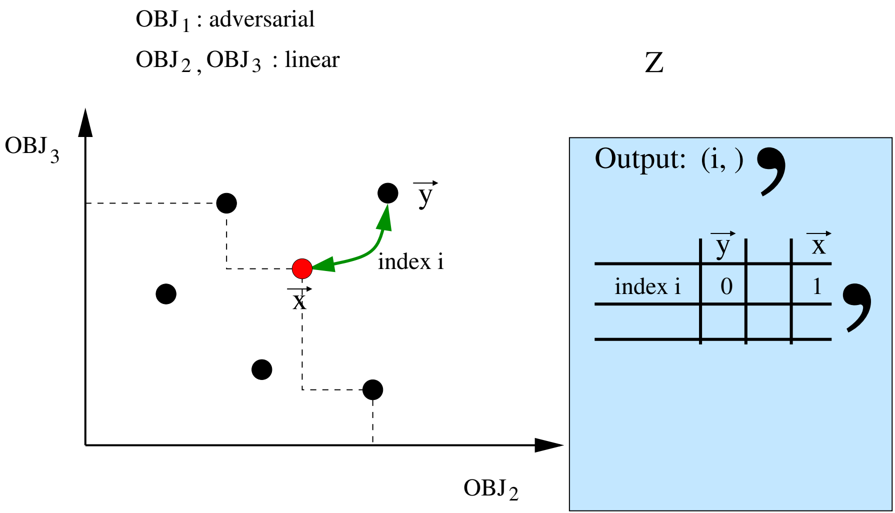
OBJ_1 : adversarial

OBJ_2 , OBJ_3 : linear

Z



Transcription for $d = 3$

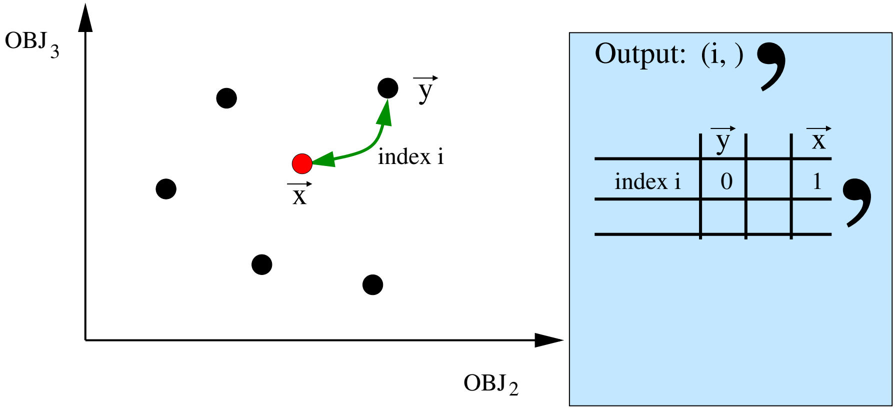


Transcription for $d = 3$

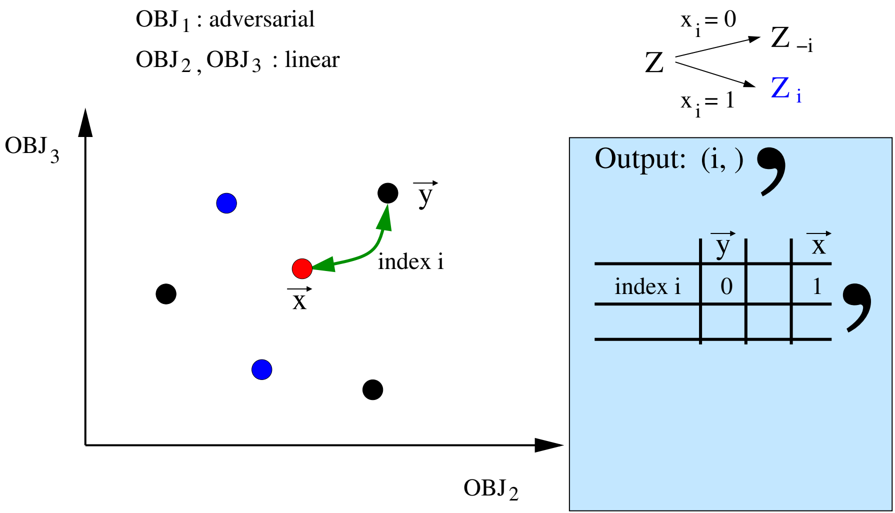
OBJ_1 : adversarial

OBJ_2 , OBJ_3 : linear

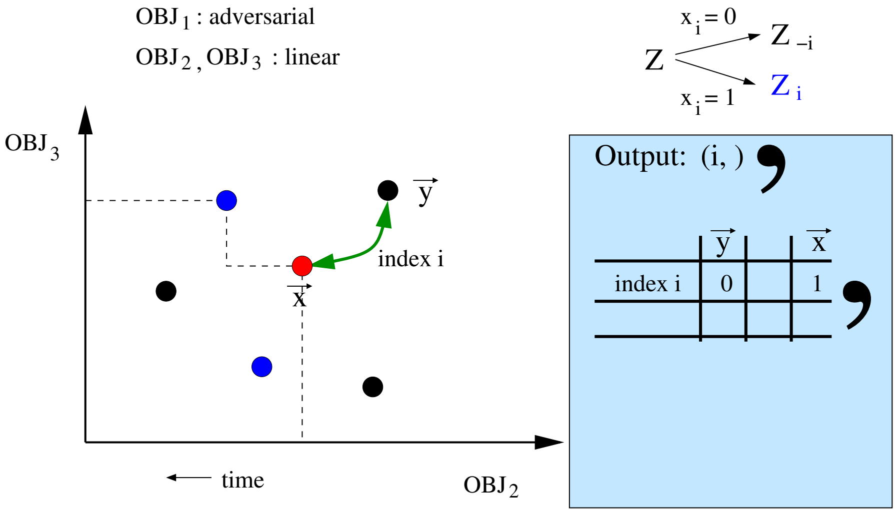
Z



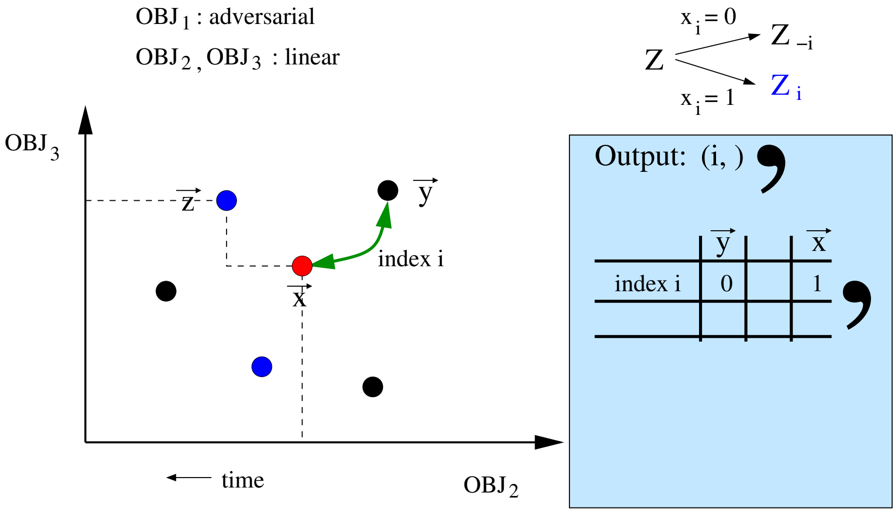
Transcription for $d = 3$



Transcription for $d = 3$



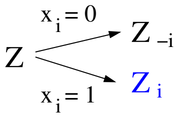
Transcription for $d = 3$

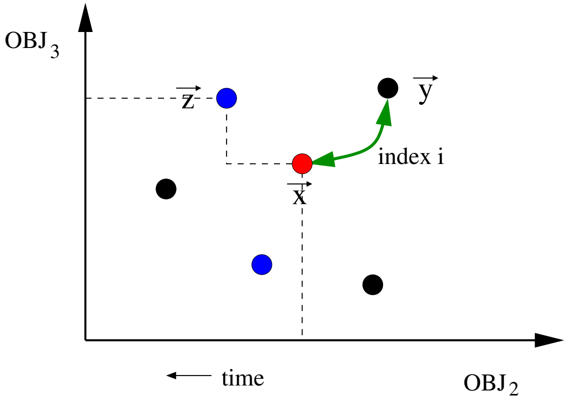


Transcription for $d = 3$

OBJ_1 : adversarial

OBJ_2, OBJ_3 : linear





Output: (i,)

	$\vec{y}$	$\vec{z}$	$\vec{x}$
index i	0	1	1

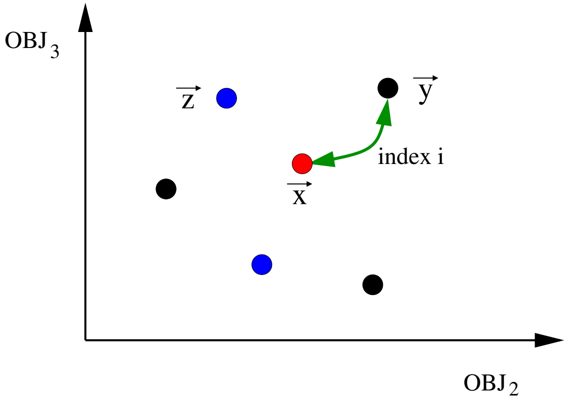
Transcription for $d = 3$

OBJ_1 : adversarial

OBJ_2, OBJ_3 : linear

$Z \xrightarrow{x_i = 0} Z_{-i}$

$Z \xrightarrow{x_i = 1} Z_i$



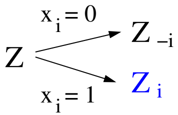
Output: (i,)

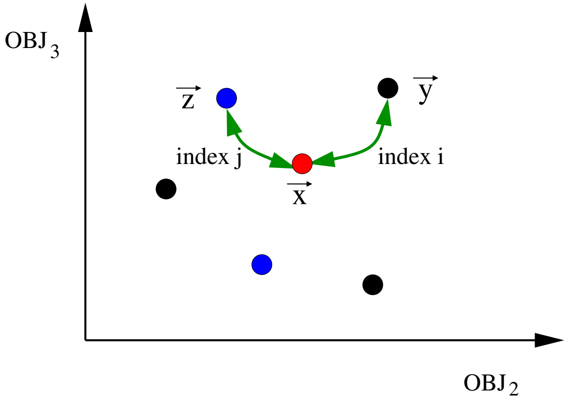
	$\vec{y}$	$\vec{z}$	$\vec{x}$
index i	0	1	1

Transcription for $d = 3$

OBJ_1 : adversarial

OBJ_2 , OBJ_3 : linear





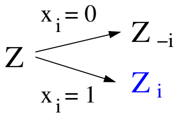
Output: (i,) ,

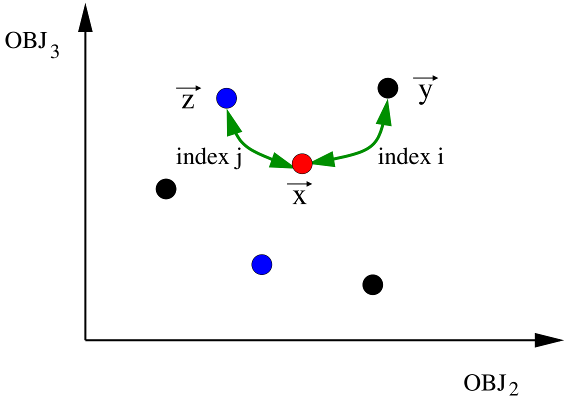
	$\vec{y}$	$\vec{z}$	$\vec{x}$
index i	0	1	1

Transcription for $d = 3$

OBJ_1 : adversarial

OBJ_2 , OBJ_3 : linear





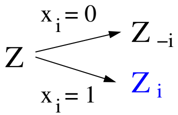
Output: (i,j)

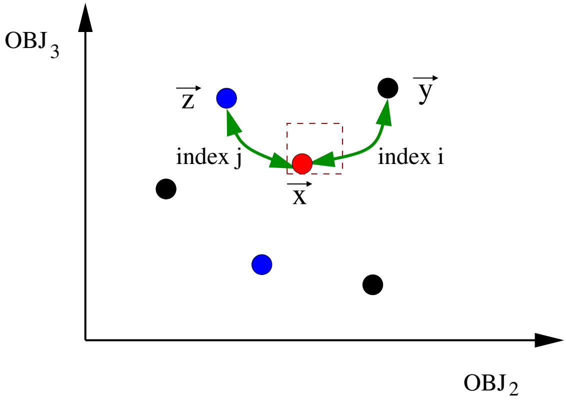
	$\vec{y}$	$\vec{z}$	$\vec{x}$
index i	0	1	1
index j	1	0	1

Transcription for $d = 3$

OBJ_1 : adversarial

OBJ_2, OBJ_3 : linear





Output: (i,j)

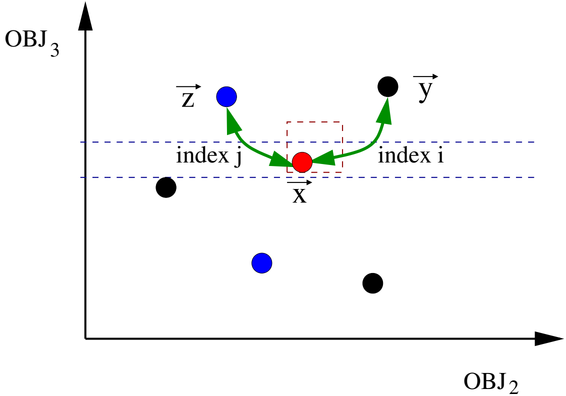
	$\vec{y}$	$\vec{z}$	$\vec{x}$
index i	0	1	1
index j	1	0	1

Transcription for $d = 3$

OBJ_1 : adversarial

OBJ_2, OBJ_3 : linear

$$Z \xrightarrow{x_i = 0} Z_{-i}$$
$$Z \xrightarrow{x_i = 1} Z_i$$



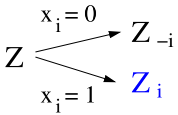
Output: (i,j)

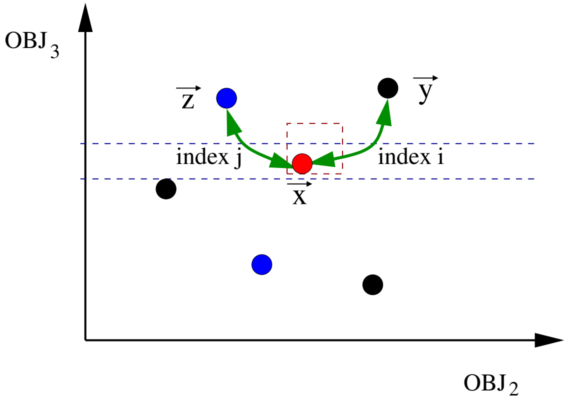
	$\vec{y}$	$\vec{z}$	$\vec{x}$
index i	0	1	1
index j	1	0	1

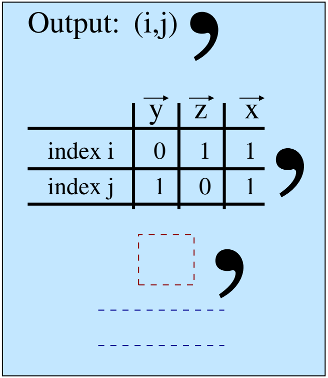
Transcription for $d = 3$

OBJ$_1$: adversarial

OBJ$_2$, OBJ$_3$: linear







	$\vec{y}$	$\vec{z}$	$\vec{x}$
index i	0	1	1
index j	1	0	1

Future Directions?

Future Directions?

Open Question

Are there additional applications of randomness "conservation" in Smoothed Analysis?

Future Directions?

Open Question

Are there additional applications of randomness "conservation" in Smoothed Analysis?

e.g. simplex algorithm

Future Directions?

Open Question

Are there additional applications of randomness "conservation" in Smoothed Analysis?

e.g. simplex algorithm

Open Question

Are there perturbation models that make sense for non-linear objectives?

Future Directions?

Open Question

Are there additional applications of randomness "conservation" in Smoothed Analysis?

e.g. simplex algorithm

Open Question

Are there perturbation models that make sense for non-linear objectives?

e.g. submodular functions

Questions?

Thanks!