Unit: Multicriteria decision analysis

Learning goals

I. General classification of strategies in Multi-objective Optimization based on the time of decision maker interaction.
II. Ways to structure the decision making process
III. What are utility functions? How can we construct rankings?
IV. Pareto optima and related definitions
V. Interpretation and Visualization of Pareto fronts

Multicriteria Decision Analysis

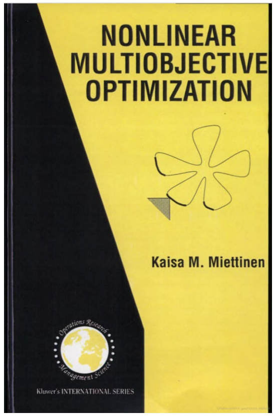
Miettinen makes the following distinction between three general classes of multicriteria decision analysis approaches:

A posteriori: First compute set of non-dominated solutions using optimization, and afterwards make a decision by viewing solutions.

A priori: First design utility function and then find single optimum with respect to the utility function. The decision making is already finished before optimization.

Progressive: Interactive methods that combine a posteriori and a priori decision making and optimization in several feedback loops.

Many classical MCDM methods are described in her book: following this structuring approach
Miettinen, Kaisa. Nonlinear multiobjective optimization. Springer, 1999.



Kaisa Miettinen,
Finnish Mathematician

A priori multicriteria decision making

Questions in MCDA

Given A finite set of alternatives $x \in \mathcal{X}$, multiple objective functions $f_1(x), \ldots, f_m(x)$ (evaluation criteria).

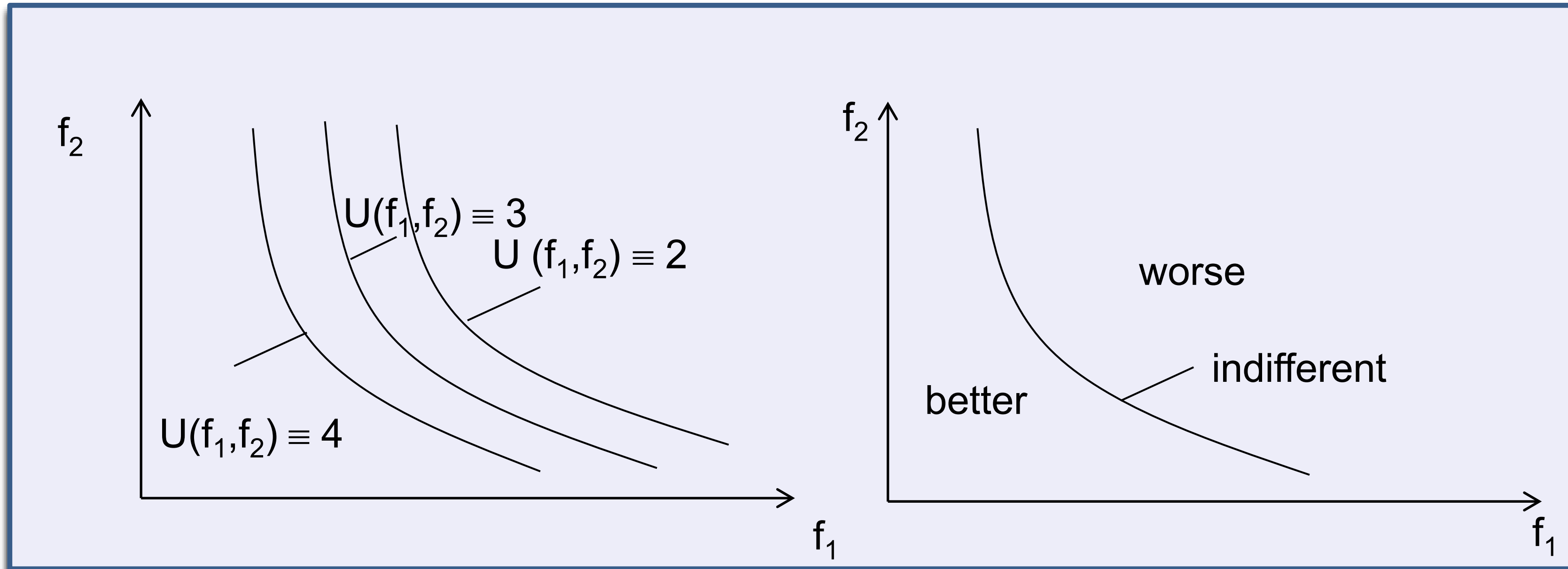
Wanted: A (partial) ranking or a preferred choice of alternatives that is compatible with a decision maker's preferences.

The information on the criteria is typically not sufficient to establish full ranking (e.g. conflicting objectives),

Additional *preference information* is asked from the decision maker (DM): *Preference elicitation process*

For instance use questionaires, or past actions and statements. Information: pairwise comparisons, weights, etc.

Utility Function, Indifference Curves



Definition: A utility function $U : \mathbb{R}^m \to \mathbb{R}$ is a mapping from objective function values to a single objective function value called utility

Definition: An *indifference curve* is a set of points in the objective space, where the utility function has the same value.

Multi-Attribute Utility Theory (MAUT)

In MAUT several criteria (attributes) are aggregated to a single utility function value Choose a functional form for the aggregation, e.g.:

Additive Utility Function:

$$U(f_1(x), ..., f_m(x)) = w_1 f_1(x)... + w_m f_m(x)$$

Keeney and Raiffa Utility Function:

$$U(f_1(x)..., f_m(x)) = (1 + f_1(x))^{w_1} \cdot ... \cdot (1 + f_m(x))^{w_m}$$

Keeney and Raiffa utility has linear indifference curves in log-log plot

Keeney, R. L. and Raiffa, H. (1976). Decisions with Multiple Objectives: Preferences and Value Tradeoffs. Wiley,New York. Reprinted, Cambridge Univ. Press, New York (1993).

Keeney, R. L. (2009). Value-focused thinking: A path to creative decision making. Harvard Univ. Press.

Two ways to define parameters of utility function

(Robust) ordinal regression: The weights or parameters of the utility function are found based on statements of decision maker (DM) on pairwise comparisons or preferred choices (e.g., past actions). They should lead to a compatible and robust ranking of alternatives.

Guided specification of parameters, desirability functions The DM specifies the parameters utility function in a guided process, e.g., first focusing on single objectives and then on their relative importance.

Ordinal regression (example)

	QTeaching max $f_1(x)$	QScience max $f_2(x)$	QCoffee max $f_3(x)$	QTechnology max $f_4(x)$
Delf TU	8	6	5	8
Leid Univ.	8	8	5	7
Mit Univ.	8	7	2	9
Star Univ.	5	5	9	5

Linear Rating: $Score(x) = w_1 * f_1(x) + ... + w_4 * f_4(x)$, all weights $\geq$ 0.

Simple question: Find weights for a utility function that ranks Star Univ. first?

More difficult question: The DM ranks $L \geq D$ and $S \geq M$. Is there a compatible utility function?

Ordinal regression (solution)

	QTeaching max $f_1(x)$	QScience max $f_2(x)$	QCoffee max $f_3(x)$	QTechnology max $f_4(x)$
Delf TU	8	6	5	8
Leid Univ.	8	8	5	7
Mit Univ.	8	7	2	9
Star Univ.	5	5	9	5

Linear Rating: $Score(x) = w_1 * f_1(x) + ... + w_4 * f_4(x)$, all weights $\geq$ 0.

$L \geq D$: $\Leftrightarrow 8w_1 + 8w_2 + 5w_3 + 7w_4 \geq 8w_1 + 6w_2 + 5w_3 + 8w_4$
$\Leftrightarrow 2w_2 - w_4 \geq 0$
$S \geq M$: $\Leftrightarrow 5w_1 + 5w_2 + 9w_3 + 5w_4 \geq 8w_1 + 7w_2 + 2w_3 + 9w_4$
$\Leftrightarrow -3w_1 - 2w_2 + 7w_3 - 4w_4 \geq 0$

One solution is $w_1 = 0, w_2 = 0.1, w_3 = 0.7, w_4 = 0.1$

Ordinal regression

For a set of pairs $(x_1^\ell, x_2^\ell) \in \mathcal{X} \times \mathcal{X}$, $\ell = 1, \ldots, q$ the DM specified x_1^ℓ better or equal to x_2^ℓ. We want to find compatible w_i, $i = 1, ..., m$ such that:

$$\sum w_i f_i(x_1^\ell) \leq \sum w_i f_i(x_2^\ell), \ell = 1, \ldots, q$$

$$\sum w_i = 1, w_i \geq 0, i = 1, \ldots, m$$

A robust solution maximizes the margins to the constraint boundaries:

$$\max \delta$$

$$\sum w_i f_i(x_1^\ell) + \delta \leq \sum w_i f_i(x_2^\ell), \ell = 1, \ldots, q$$

$$\sum w_i = 1, w_i \geq 0, i = 1, \ldots, m$$

Robust ordinal regression typically uses piecewise linear utility functions, instead of linear ones.

Further Reading

- Greco, Salvatore, Roman Słowiński, José Rui Figueira, and Vincent Mousseau. "Robust ordinal regression." In Trends in multiple criteria decision analysis, pp. 241-283. Springer US, 2010.
- Greco, Salvatore, Vincent Mousseau, and Roman Słowiński. "Ordinal regression revisited: multiple criteria ranking using a set of additive value functions." *European Journal of Operational Research* 191, no. 2 (2008): 416-436.

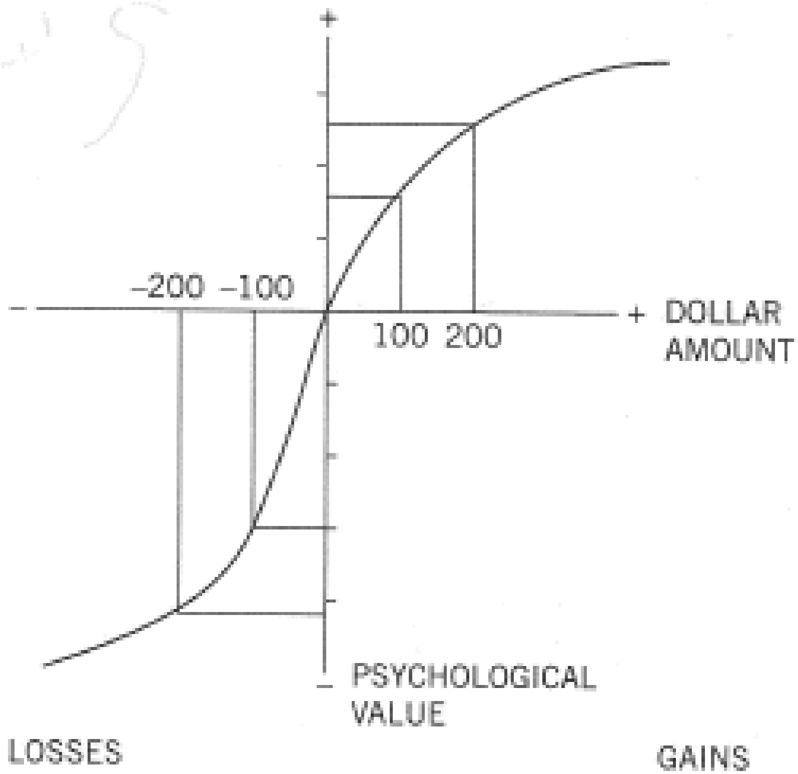
Multi-Attribute Utility Theory (MAUT)

In order rank alternatives, the decision analyst can follow these steps:

1. Determine all criteria and constraints (informal)
2. Determine decision space $\mathcal{X}$ (set of possible decisions)
3. Determine quantitative measures for criteria
4. Measure criteria values $f_1(x), ..., f_m(x)$ for $x \in \mathcal{X}$
5. Normalized scale, e.g. between 0 and 1
6. Weight objectives by importance $w_1, ..., w_m$
7. Compute weighted utility function $U(f_1(x), ..., f_m(x), w_1, ..., w_m)$ for each solution; Rank solution in $\mathcal{X}$ based on U.

Keeney, R.L. (1992). Value-focused thinking—A Path to Creative Decision making. Harvard University Press.

Client-Theory by Kahneman and Tversky



D. Kahneman: ‘Thinking fast and slow’ Penguin press. 2014.

How good are people feeling when winning a lottery?

Difference between 0 and 1000 appears bigger than difference between 100000 and 101000.

Bernoulli (1738): 'The psychological response to the change of wealth is inversely proportional to the initial amount of wealth' Degressive utility function, e.g., $\log(f(x)$ instead of linear one.

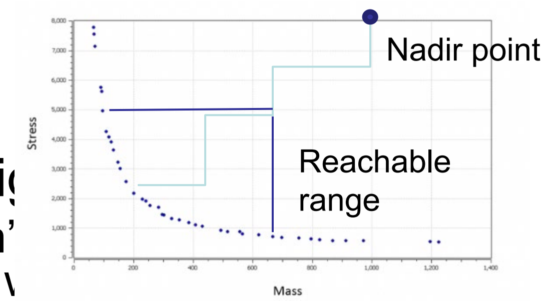
Kahnemann and Tversky: Loss is higher weighted than win by decision makers (prospect theory).

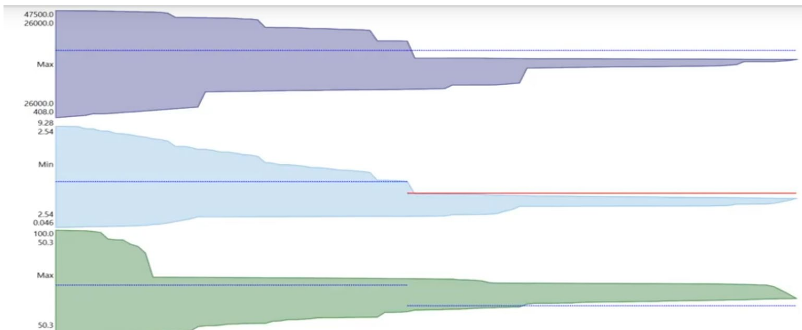
$\Rightarrow$ The initial wealth matters. (cannot be modeled by utility function)

Pareto/Nautilus Navigation

- Navigation on the Pareto front:
  - Move from point to point on the pareto front by step-wise improvements and deteriorations
  - In order to further improve a point in one objective, it needs to get worse in at least one other objective
  - Often different possible moves to improve the chosen objectives with different losses in other objectives
  - Pareto navigation is implemented in NIMBUS Package;

- Nautilus Navi
  - People don’ Navigation v only
  - Starting from a very bad and dominated point (Nadir point)
  - A point can be improved step-wise until the Pareto front is hit





https://www.youtube.com/watch?v=gjvIG8PiPBo

Eskelinen, P., Miettinen, K., Klamroth, K., & Hakanen, J. (2010). Pareto navigator for interactive nonlinear multiobjective optimization. *OR spectrum*, *32*(1), 211-227.

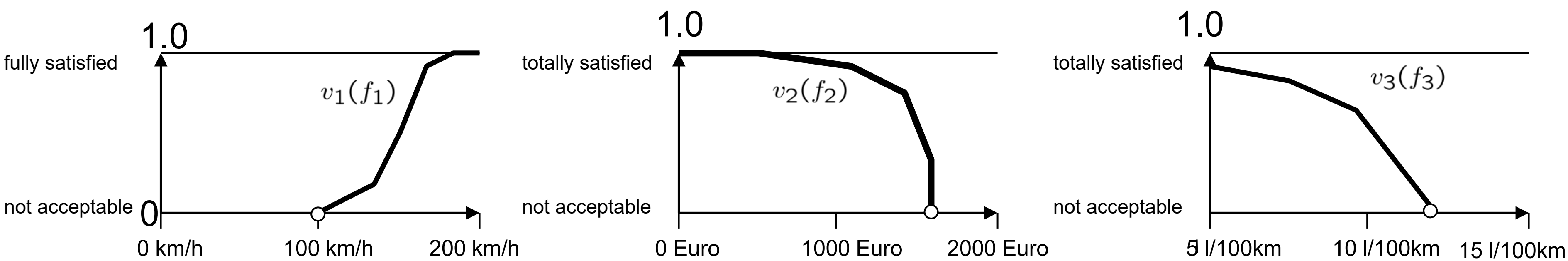
Desirability functions (DFs)

Creating a nonlinear utility by using DFs:

Step 1: Define desirability functions for each objective function value transformation

Step 2: Weight objective functions

Step 3: Aggregate multiple weighted objective functions



$w_1 = 7$ $w_2 = 8$ $w_3 = 10$

$$U(\mathbf{f}(x)) = \alpha \underbrace{\frac{1}{m}\sum_{i=1}^{m} w_i v_i(f_i(x))}_{\text{common interest}} + \beta \underbrace{\min_{i\in\{1,\ldots,m\}} w_i v_i(f_i(x))}_{\text{minority interest}},$$

(here: $m = 3$)

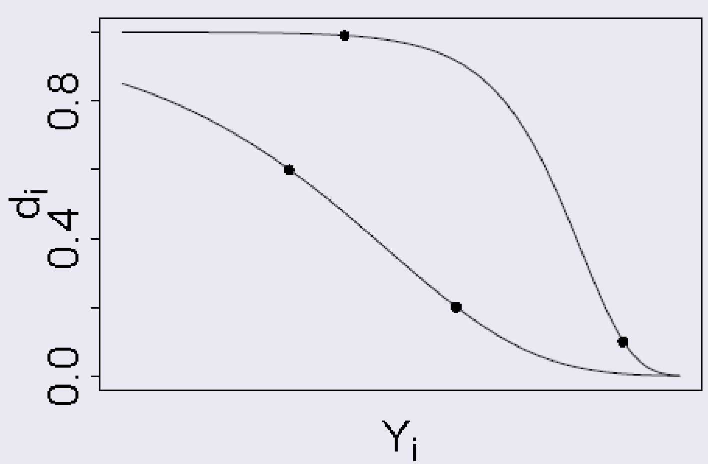
$$s.t.\ v_i(f_i(x)) > 0, i = 1, \ldots, m$$

v.d. Kuijl, Emmerich, Li: A robust multi-objective resource allocation scheme, Concurrency & Computation 22 (3), (2009)

Harrington vs. Derringer Suich type of DFs

Standardized curves controlled by few parameters.
Harrington type: never equal 0 or 1.
Derringer-Suich type: 0 if $\leq T_i$), and 1 if $\geq USL_i$

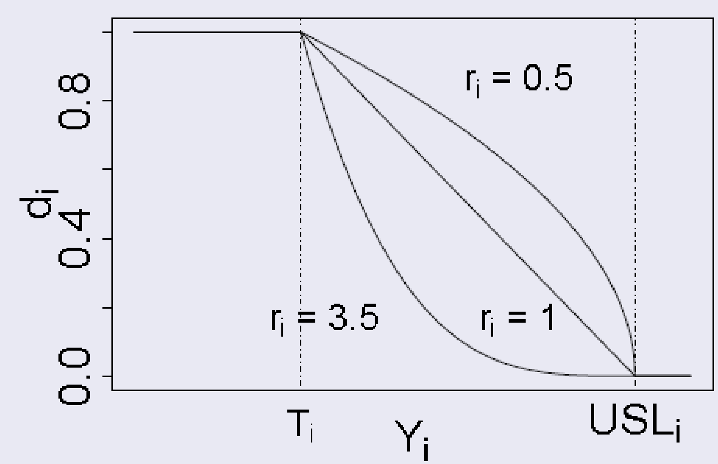
Harrington (1965)



Value pairs $(Y_i^{(1)}, d_i^{(1)})$, $(Y_i^{(2)}, d_i^{(2)})$

$$\begin{aligned} d_i(Y_i'^{(j)}) &= \exp(-\exp(-Y_i'^{(j)})) \\ Y_i'^{(j)} &= b_{0i} + b_{1i} Y_i^{(j)}, \quad j = 1, 2. \end{aligned}$$

Derringer/Suich (1980)



$$d_i(Y_i) = \begin{cases} 1, & Y_i \leq T_i \\ \left(\frac{Y_i - USL_i}{T_i - USL_i}\right)^{r_i}, & T_i < Y_i < USL_i \\ 0, & Y_i \geq USL_i \end{cases}$$

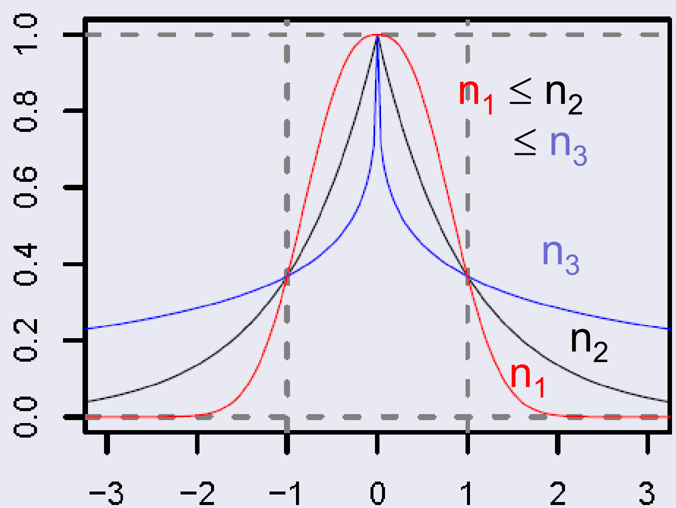
Desirability index: $D(Y_1, \ldots, Y_m) = d_1(Y_1)^{\alpha_1} * \ldots * d_m(Y_m)^{\alpha_m}$

Harrington vs. Derringer Suich type of DFs (2)

Harrington desirability functions never reach 0.

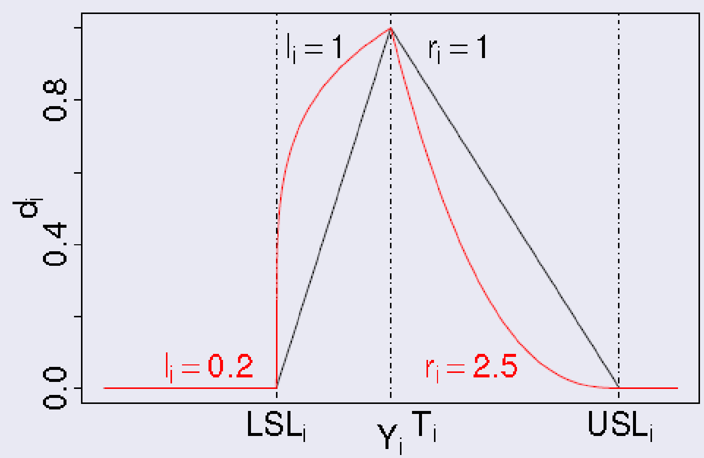
T_i is the target value; in case of Harrington it is 0, n_i and r_i control curvature (smoothness). Moreover: Lower Satisfaction Limit (LSL$_i$); Upper Satisfaction Limit (USL$_i$).

Harrington (1965)



$$d_i(Y_i') = \exp(-|Y_i'|^{n_i})$$
$$Y_i' = \frac{2Y_i - (USL_i + LSL_i)}{|USL_i - LSL_i|}$$

Derringer/Suich (1980)



$$d_i(Y_i) = \begin{cases} 0, & Y_i < LSL_i \\ \left(\frac{Y_i - LSL_i}{T_i - LSL_i}\right)^{l_i}, & LSL_i \leq Y_i \leq T_i \\ \left(\frac{Y_i - USL_i}{T_i - USL_i}\right)^{r_i}, & T_i < Y_i \leq USL_i \\ 0, & Y_i > USL_i \end{cases}$$

Desirability index: $D(Y_1, \ldots, Y_m) = d_1(Y_1)^{\alpha_1} * \ldots * d_m(Y_m)^{\alpha_m}$

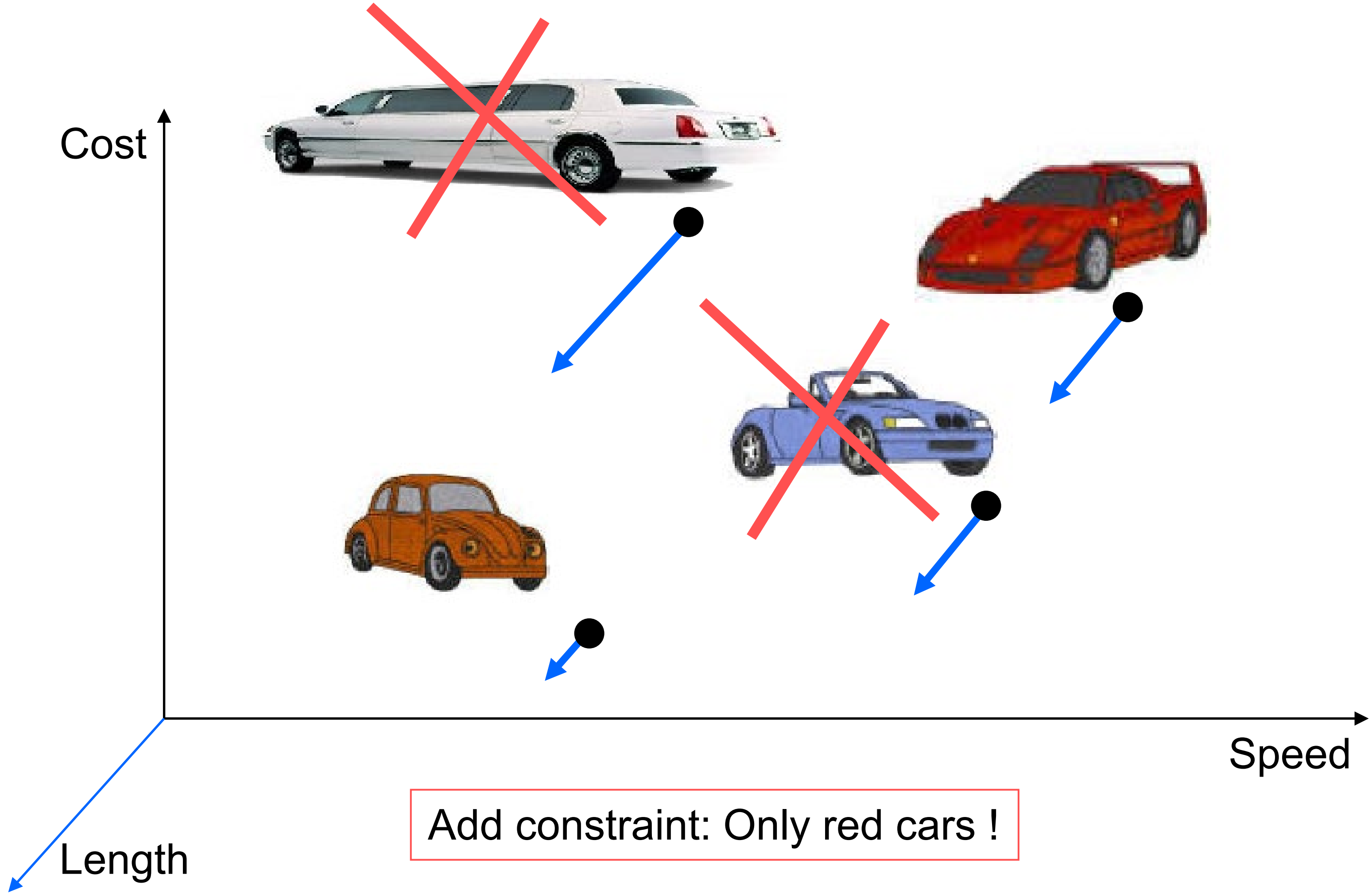
Desirability functions (DFs)

- Harrington, E. C. "The desirability function." *Industrial quality control* 21, no. 10 (1965): 494-498. (DFs type Harrington)
- Derringer, G., & Suich, R. (1980). Simultaneous optimization of several response variables. *Journal of quality technology*, *12*(4), 214-219. (DFs Type Derringer-Suich)
- Van der Kuijl, Emmerich, M., & Li, H. (2010). A robust multi-objective resource allocation scheme incorporating uncertainty and service differentiation. *Concurrency and Computation: Practice and Experience*, *22*(3), 314-328. (DFs Type v.d.Kuijl et al.)
- Trautmann, Heike, and Claus Weihs. *Pareto-Optimality and Desirability Indices*. Universität Dortmund, 2004.

A posteriori multicriteria decision making

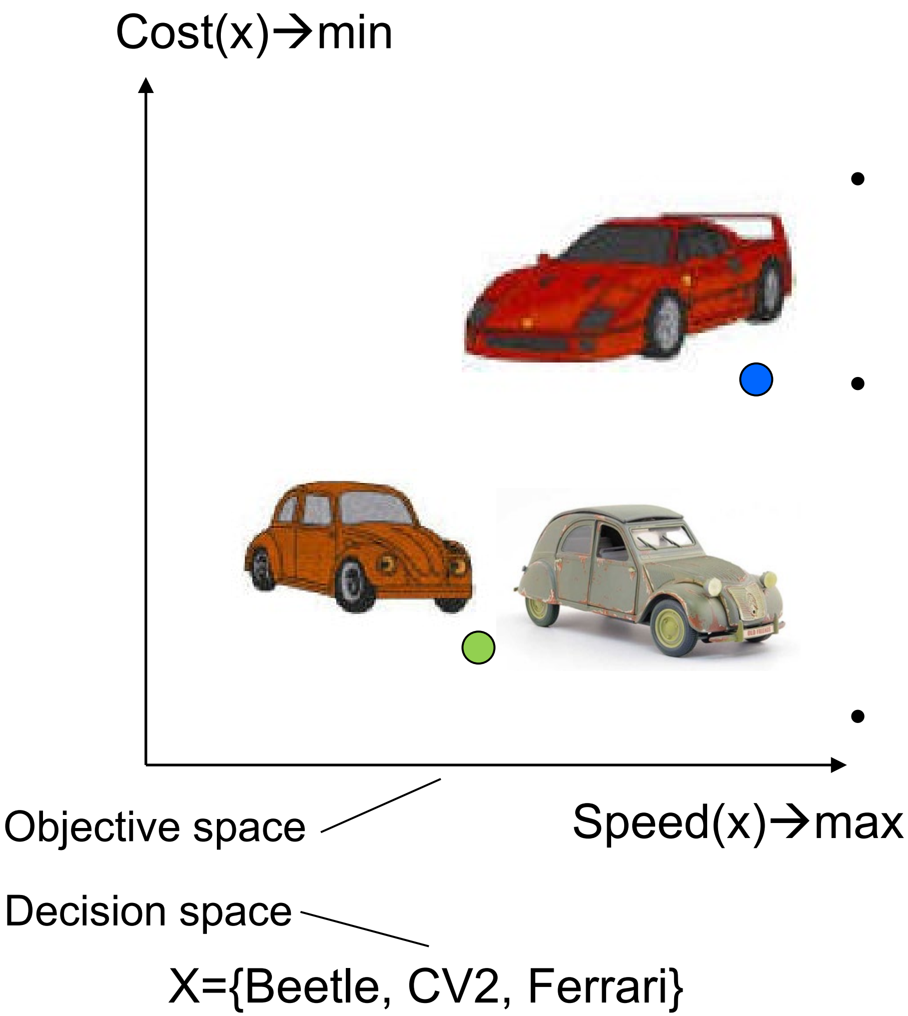
Multicriteria Problems: Car Example

Your objectives: Cost → min, Speed → max



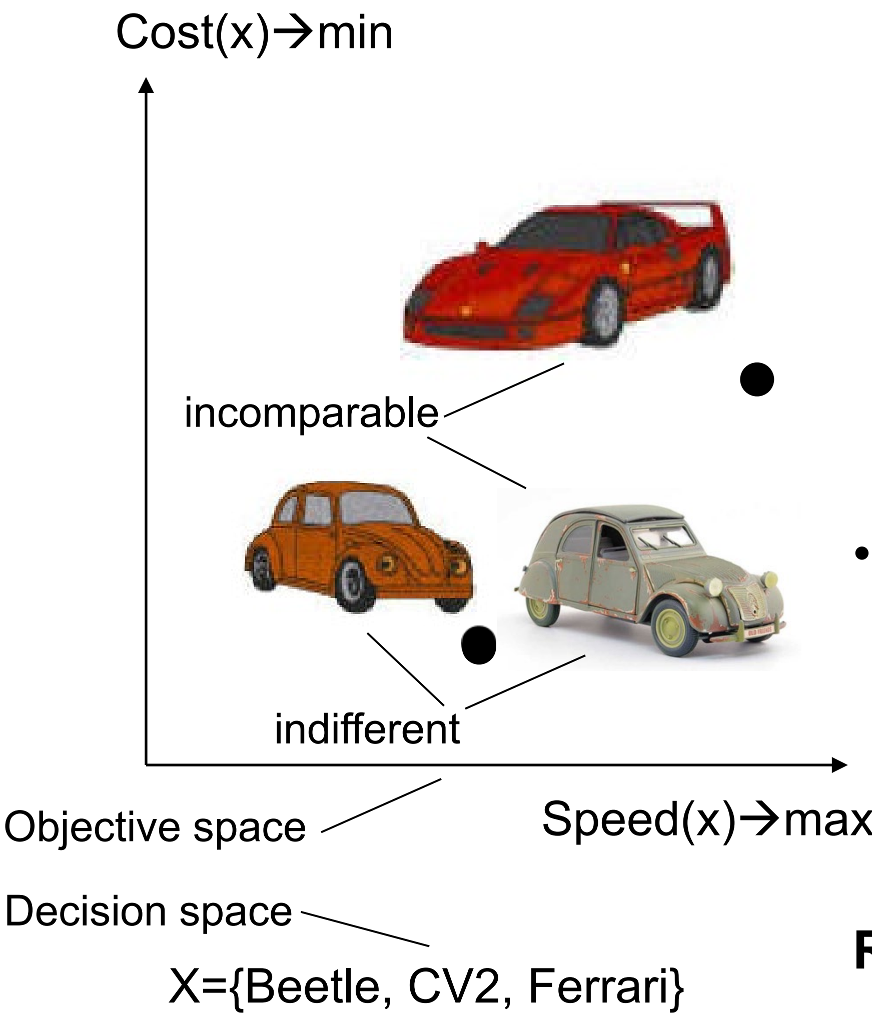
Add constraint: Only red cars !

Decision vs. Objective Space, pre-images



- **Def.:** The space of candidate solutions is called the decision space, for instance X={$Beetle$, $CV2$, $Ferrari$} in the figure.
- **Def.:** The space of objective function values (vectors) is called the objective space, for instance $\mathbb{R}^2$ in the figure.
- **Def.:** For a point in the objective space the corresponding point(s) in the decision space are called their **pre-image(s)**. In the figure, Beetle and CV2 are pre-images of the green point.
- **Remark:** Two different points in the decision space (e.g. Beetle and CV2) can map to the same point in the objective space, but two points in the objective space have never the same preimage in the decision space.

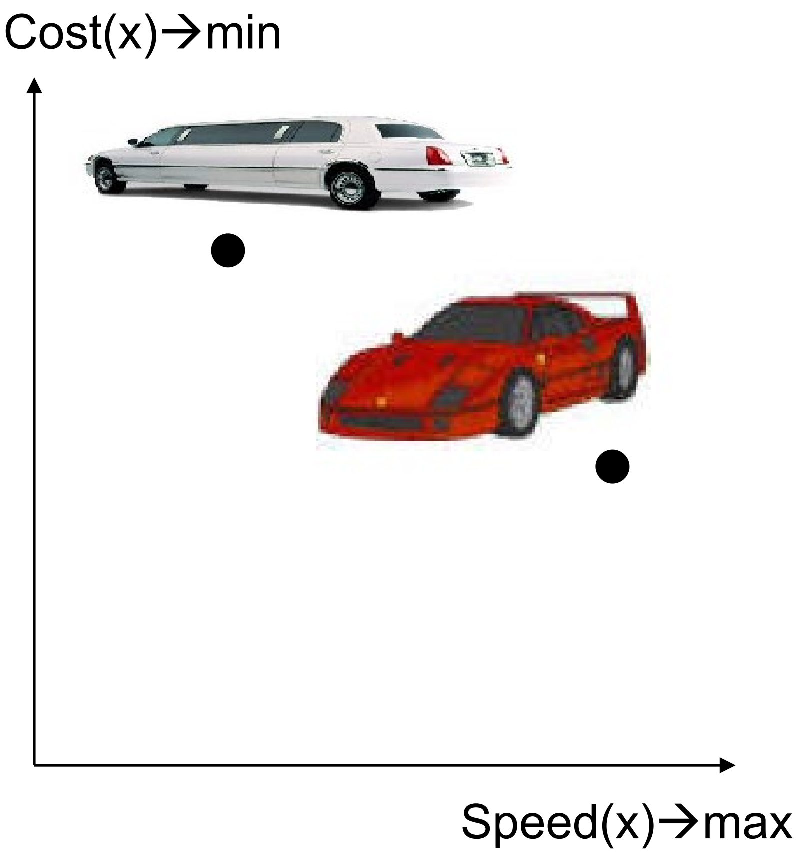
Incomparable and indifferent solutions



- **Def.:** Given two solutions and some criterion functions, the solutions are said to be **incomparable**, if and only if
  1. the first solution is better than the second solution in one or more criterion function value and
  2. the second solution is better than the first alternative in one or more other criterion function values.
- **Def.:** Given two alternatives and some criterion functions, the alternatives are said to be **indifferent** with respect to each other, if and only if they share exactly the same criterion function values.

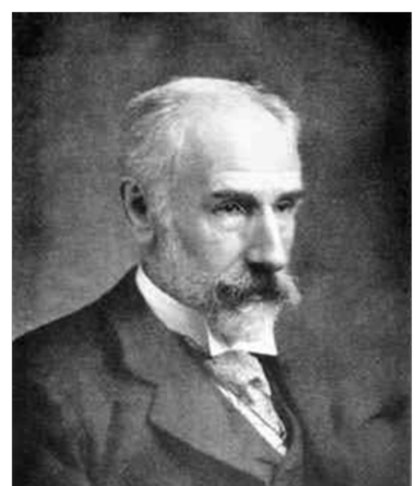
Remark: In both cases additional preference information is required to decide which solution is best.

Pareto dominance

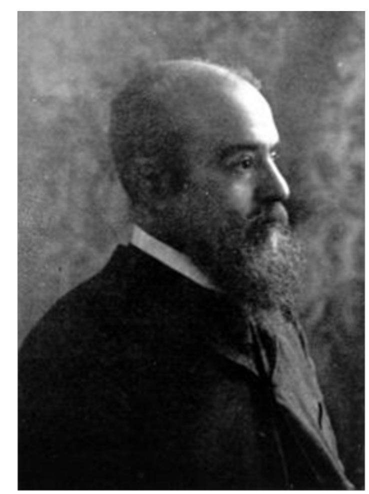


- **Def.:** Given two alternative solutions and some objective functions, the first solution ist said to Pareto dominate the second solution, if and only if
  1. the first solution is better or equal in all objective function values, and
  2. the first solution is better in at least one objective function value.

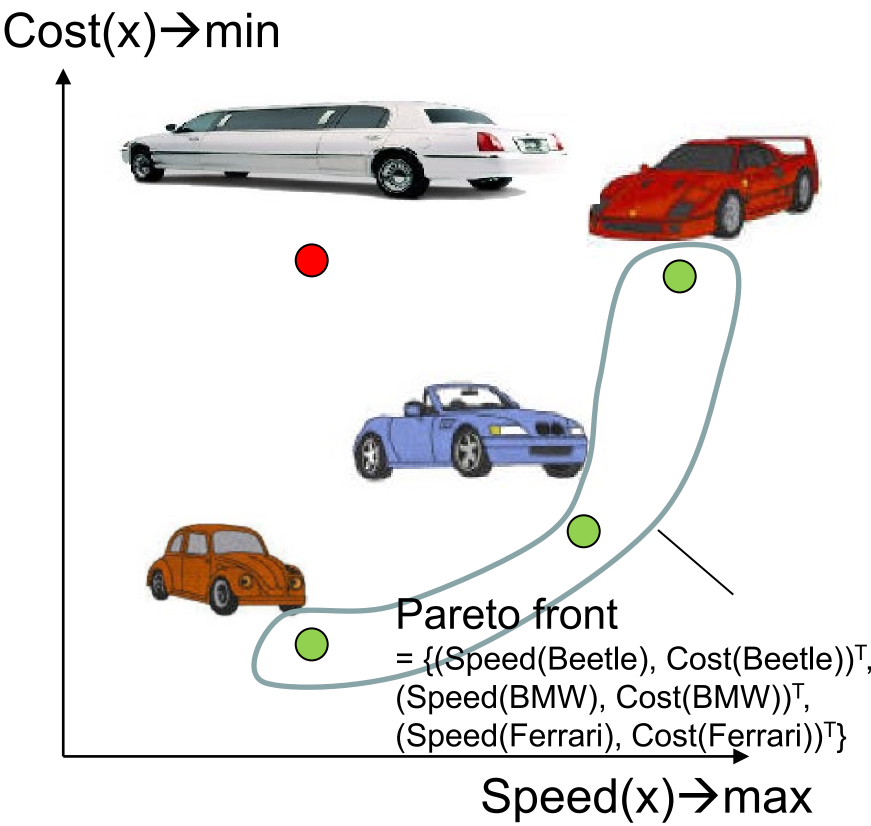
Francis Y. Edgeworth
Irish Economist
1845-1926



Vilfredo Pareto
Italian Economist
1848-1923



Pareto Dominance, Pareto front



X={Beetle, Limosine, BMW, Ferrari}

X_E={Beetle, BMW, Ferrari}

Efficient set

Given: *A decision space $\mathcal{X}$ comprising all (feasible) decision alternatives, a number of criterion functions*

$$f_i : \mathcal{X} \rightarrow \mathbb{R}\ ,\ i = 1, \ldots, m$$

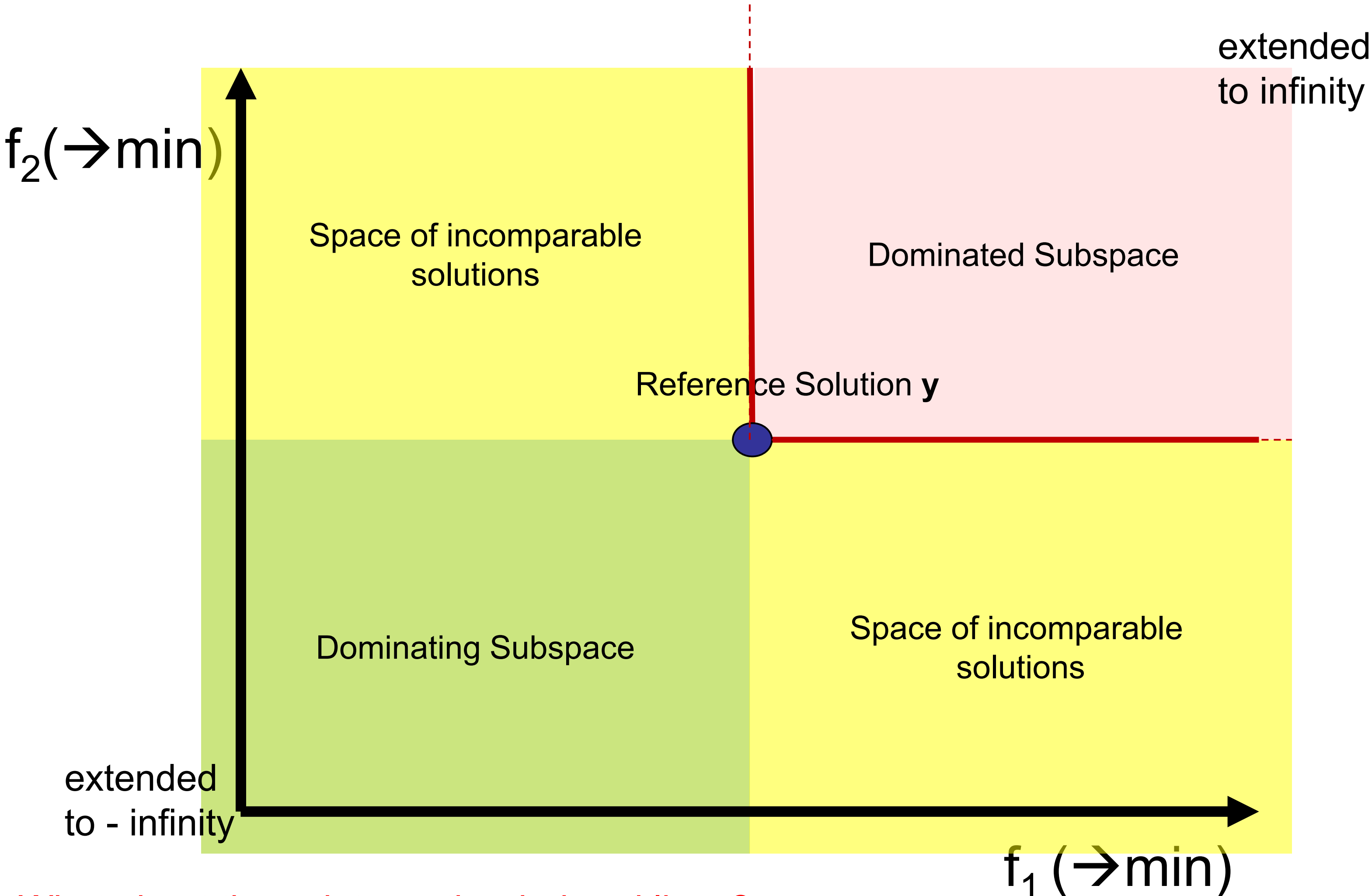
Def.: *A decision alternative x in X dominates a solution x' in X, iff it is not worse in each objective function value, and better in at least one objective function value.*

Def.: *If a solution $x \in \mathcal{X}$ is not dominated by any other solution in X , then it is called Edgeworth-Pareto optimal (or Pareto optimal) (in X).*

Def.: *The set of all Pareto optimal solutions in X is called efficient set.*

Def.: *The set of all Pareto optimal function vectors of solutions in the efficient set is called the Pareto front.*

Fundamental Concepts: Dominance diagram

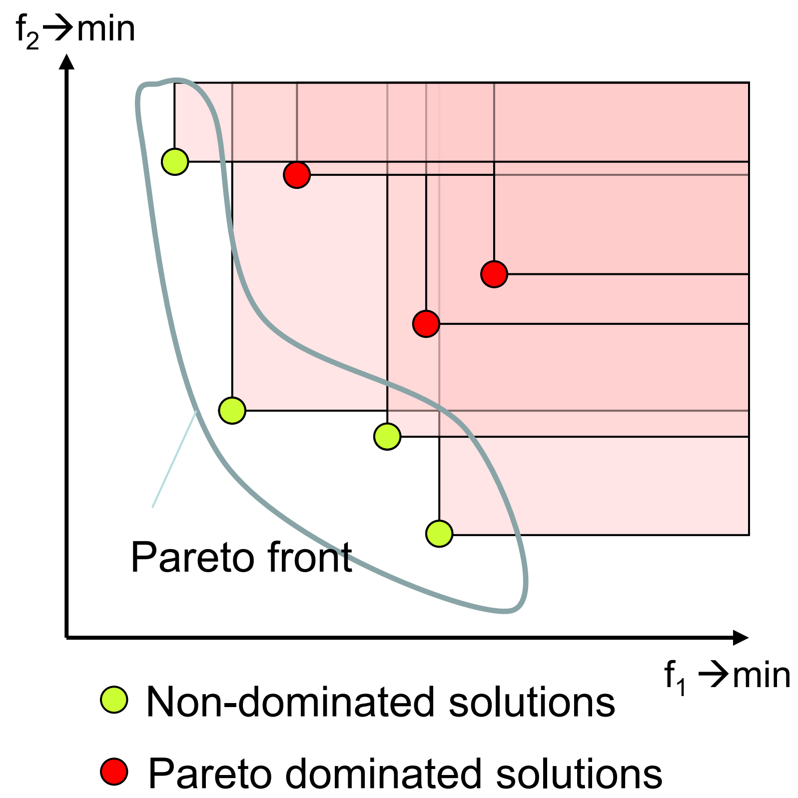


What about the points on the dark red lines?
How would this diagram look for maximization of f_1 and minimization of f_2?

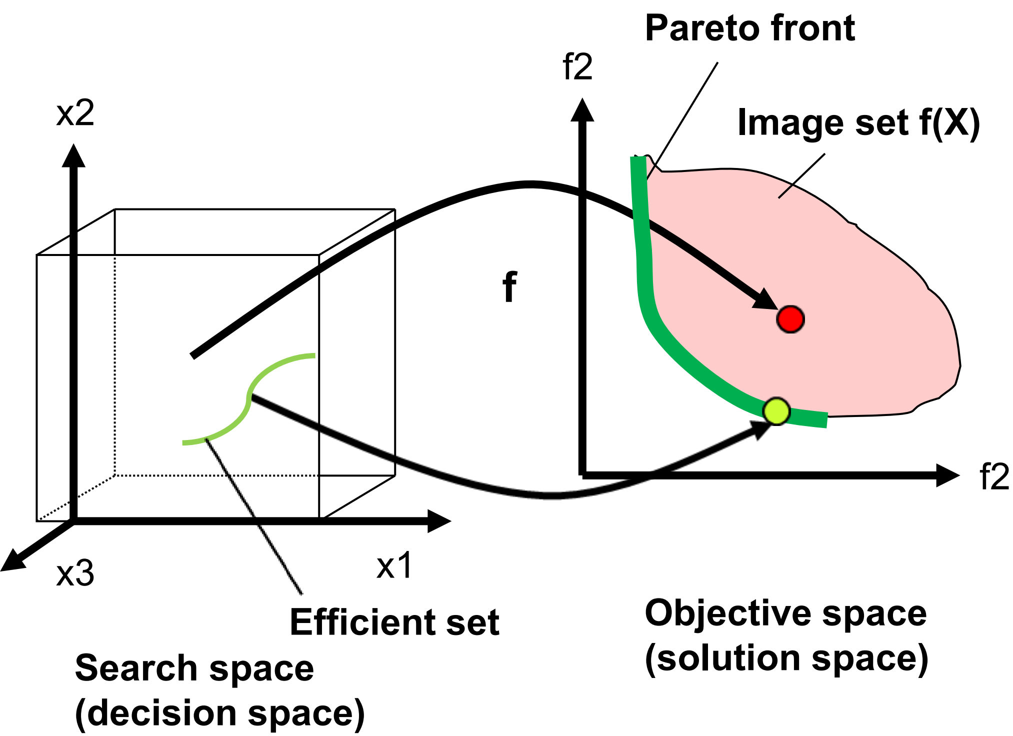
Construction of Pareto front in 2-D

Geometrical construction for separating dominated and non dominated points in 2-D:

1. Indicate for each point the dominated subspace by shading
2. The covered subspace consists of dominated points within the set, that is points that are dominated by at least one other point
3. The outer corner points on the lower left boundary form the Pareto front of the point set.



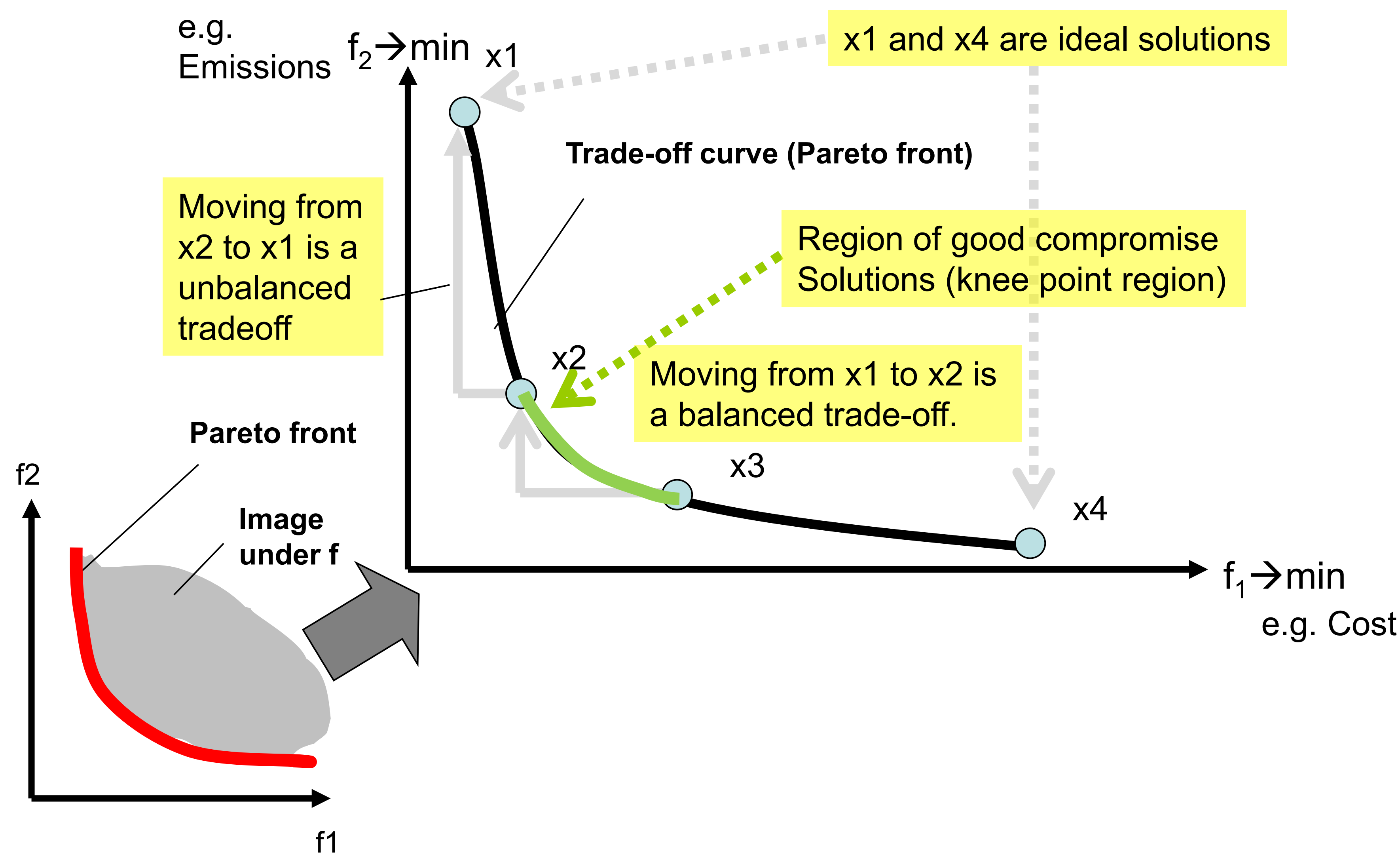
Construction for 2-D continuous functions



- For $f=(f_1, \ldots, f_m)$ the **image set f**(X) is defined as the set of all $(f_1(x), \ldots, f_m(x))$ for $x \in X$.
- The non-dominated solutions in **f**(X) are located at the lower left boundary and form the Pareto front.
- Note, that for unbounded or non-closed sets **f**(X) the Pareto front does not always exist.
- Theorem: The Pareto front of a m-objective problem has at most m-1 dimensions. Why?
- The set of all preimages of points in the Pareto front is the efficient set.

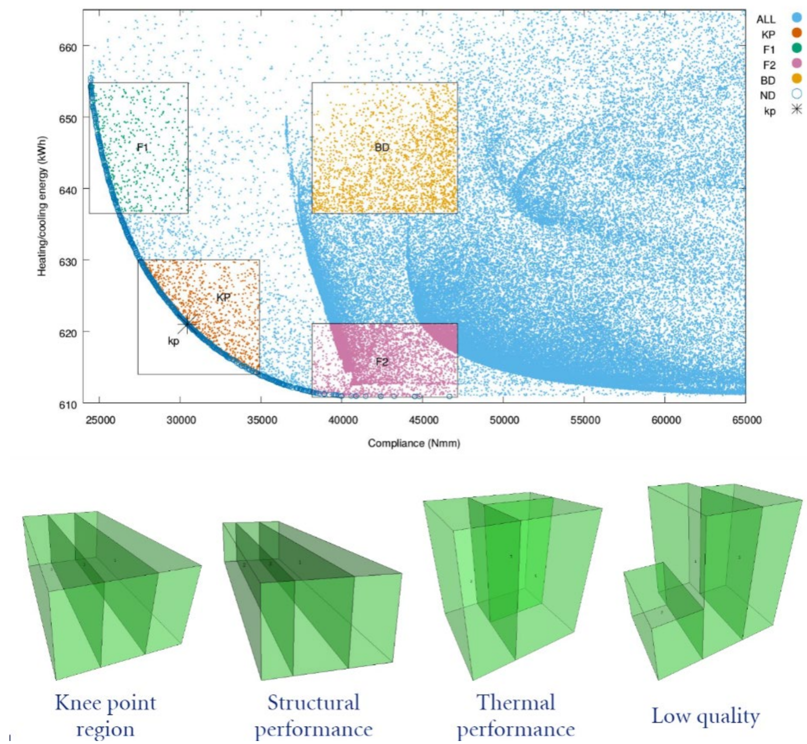
Interpreting and visualizing Pareto fronts

Thinking about trade-off, knee points



Innovization – Design principles from multiobjective optimization

- By looking at changing designs across the Pareto front, designers can study design principle
- How does the design change when moving across the Pareto front?
- Innovization: Finding design principles by multicriteria optimization



Deb et al.:. "Innovization: Innovating design principles through optimization" GECCO. ACM, 2006

van der Blom, Koen, et al. "Analysing optimisation data for multicriteria building spatial design." *International Conference on Evolutionary Multi-Criterion Optimization*. Springer, Cham, 2019.

A 3-D Pareto front

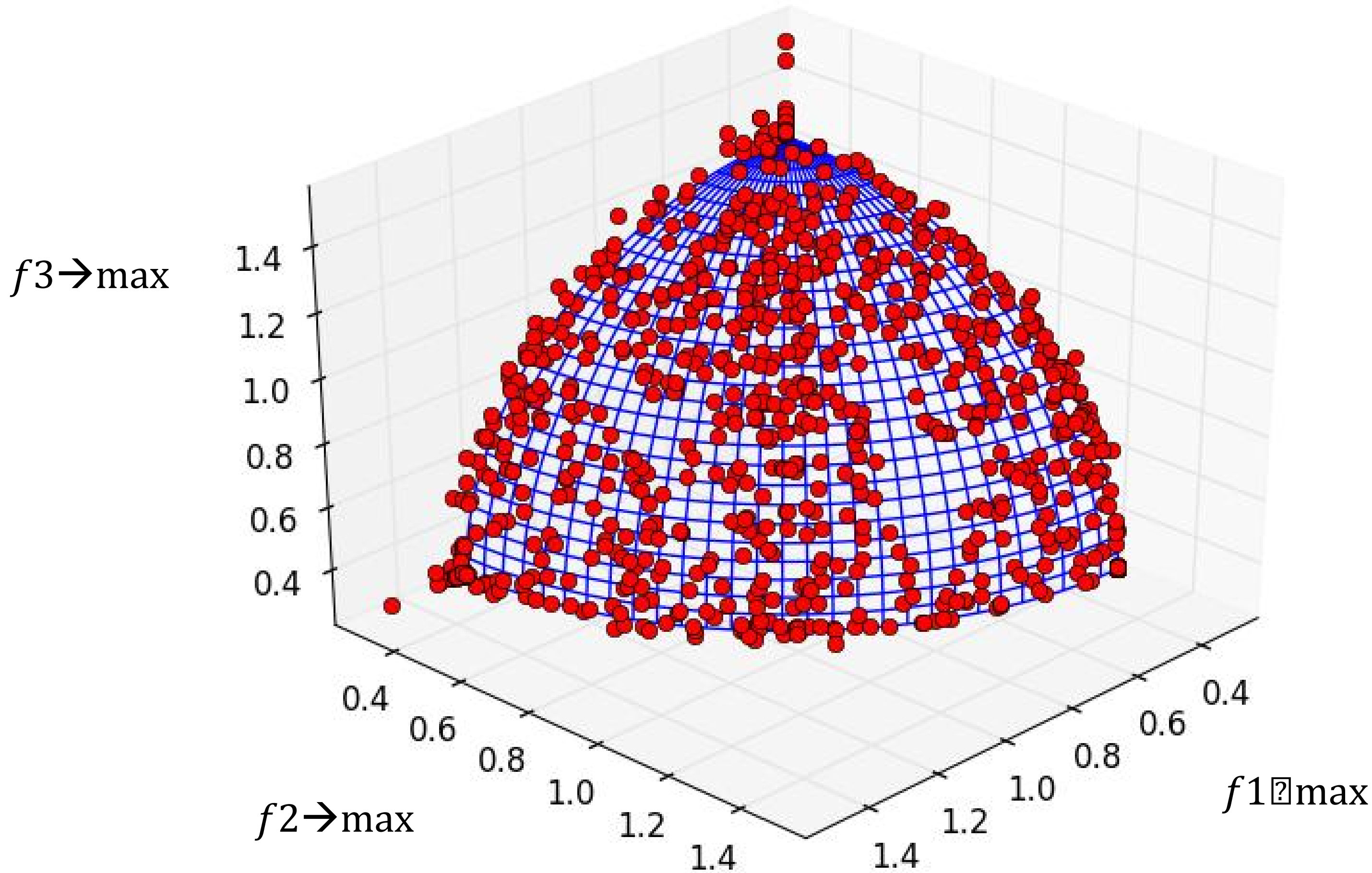
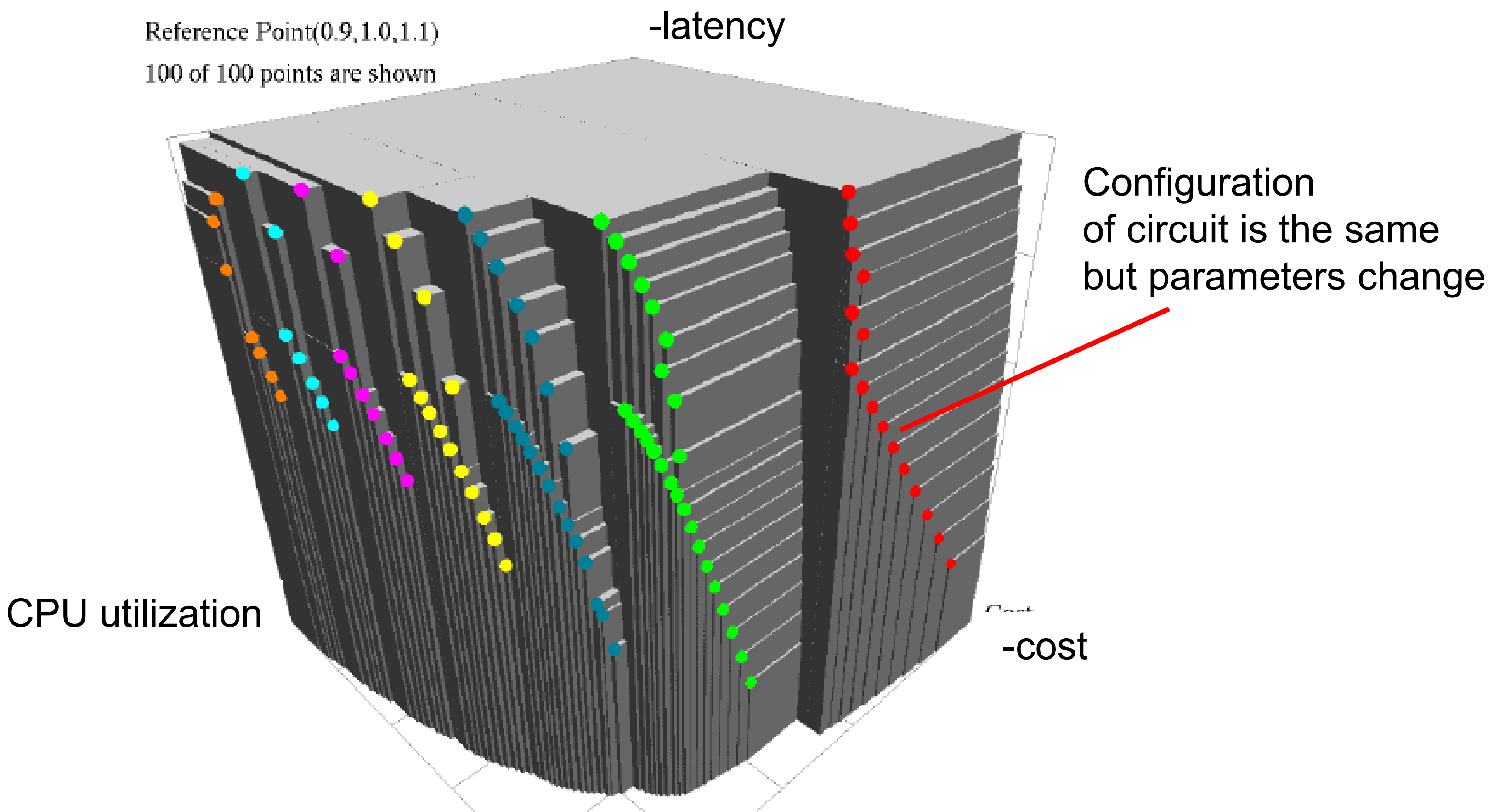


Figure: Approximation to 3-D Pareto Front
Here, minimization is the goal.

Example Pareto front: embedded systems design



Li, R., Etemaadi, R., Emmerich, M. T., & Chaudron, M. R. (2011, June). An evolutionary multiobjective optimization approach to component-based software architecture design. In *Evolutionary Computation (CEC), 2011 IEEE Congress on* (pp. 432-439). IEEE.

Many-objective optimization

- In more than 3 dimensions it is a convention to replace the term **Multi-objective Optimization** by the term **Many-objective Optimization**
- **Why is this regarded as a separate field?**
  - Other means of visualizing the Pareto front
  - More solutions tend to get non-dominated. Why?
  - Gedankenexperiment (left)
  - Articulation of preferences becomes more important

Gedankenexperiment

Let S denote a search space. And $f_1 : S \rightarrow \mathbb{R}$ ... $f_n : S \rightarrow \mathbb{R}$denote n objective functions, and S_E denote the efficient set.

What happens to the efficient set if we add another objective $f_{n+1} : S \rightarrow \mathbb{R}$ (constraint)?

Think about the answer …

Summary: Take home messages

1. In multioobjective optimization a priori, a posteriori, and progressive methods are distinguished, depending on when the DM interacts.
2. Multicriteria decision analysis structures decision process
3. Utilitiy functions capture user preferences: Linear weighting, Keeney Raiffa, Derringer-Suich type and Harrington Desirability functions.
4. In multicriteria optimization and decision analysis, two solutions can be incomparable, indifferent, or one solution dominates the other.
5. Pareto dominance and (Edgeworth-)Pareto optimality characterizes Pareto optima in multiobjective optimization.
6. Pareto fronts can be used to reason about trade-offs and also to find new design principles. They reveal the nature of the conflict(s).
7. Pareto fronts can be interpreted as trade-off curves (2-D) or as a trade-off surfaces (3-D).
8. More than 3-D → Many-objective optimization.
9. Next lesson: Using tools for Multi/Many-objective optimization in design optimization.