University of Central Lancashire
UCLan

CO3519
Artificial Intelligence

Decision support
Overall utility (or cost)
Visualization of decision making

Where opportunity creates success

Module overview

Upon successful completion of this module, a student will be able to:

1) Explain the theoretical underpinnings of algorithms and techniques specific to artificial intelligence;
2) Critically evaluate the principles and algorithms of artificial intelligence;
3) Analyse and evaluate the theoretical foundations of artificial intelligence and computing;
4) Implement artificial intelligence algorithms.

uncertainty quantification

optimization

reasoning and learning

agents and decisions

game theory

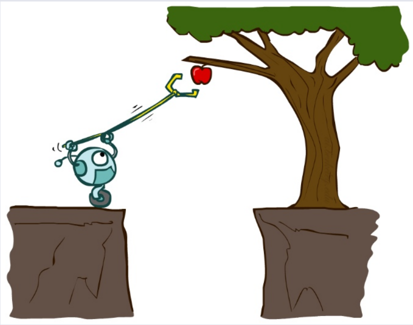
knowledge representation

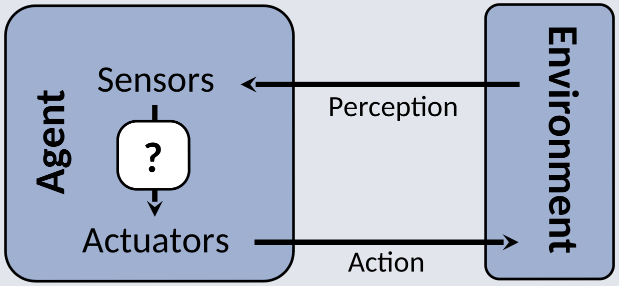
modelling

Agents and decisions

On the field of **agents and decisions**, we will:

- Review common definitions of agency and knowledge-based intelligent agents;
- Discuss the use of AI in assisting human decision making;
- Consider philosophical issues pertaining to the field, such as explainable AI and epistemic opacity.





agents and decisions

game theory

modelling

knowledge representation

reasoning and learning

Decision support

CO3519

2nd November 2021

Example decision-support scenario

1.2.3. *Cost functions with more than two parameters*

The `multivar` notebook defines a cost function for a hypothetical industrial operation; at planning and design stage, you have direct control over the following parameters:

- The investment i, done a single time, in units of £.
- The amount of goods p to be produced, in units of £/year.
- The depreciation period d (how long it is meant to operate), in units of years.

In the investment-decision example from multivar, we specified three parameters i, p, and d. However, in the underlying model, the amount of goods p that can be produced *without external manufacturers* is given by a function $p = g(i)$ of the value i. This might be **simplified to a two-parameter problem**.

Example decision-support scenario

1.2.3. *Cost functions with more than two parameters*

The `multivar` notebook defines a cost function for a hypothetical industrial operation; at planning and design stage, you have direct control over the following parameters:

- The investment i, done a single time, in units of £.
- The amount of goods p to be produced, in units of £/year.
- The depreciation period d (how long it is meant to operate), in units of years.

```
total investment        i = GBP 100000
production volume       p = GBP 1000000 per year
depreciation period     d = 7 years

operating cost: GBP 102915.03 per year
depreciation:   GBP 14285.71 per year
prod.cost:      GBP 900000.0 per year
sales contrib.: GBP -1000000.0 per year
==
total deficit:  GBP 17200.74 per year
```

A single-objective cost function was given. However, the evaluation specifies multiple contributions to it.

It may make sense to **distinguish two objectives**: Minimize proper costs, and maximize sales income.

Example decision-support scenario

In the *pareto-front* Jupyter Notebook, a version of this problem is given that expresses it with **two parameters and two minimization objectives**.

Two parameters (m = 2):

- investment $i = x_0$
- depreciation period d = x_1

Two optimization criteria (n = 2):

- expenses y_0
- contribution from sales y_1 (that is, –1 × the income from sales)

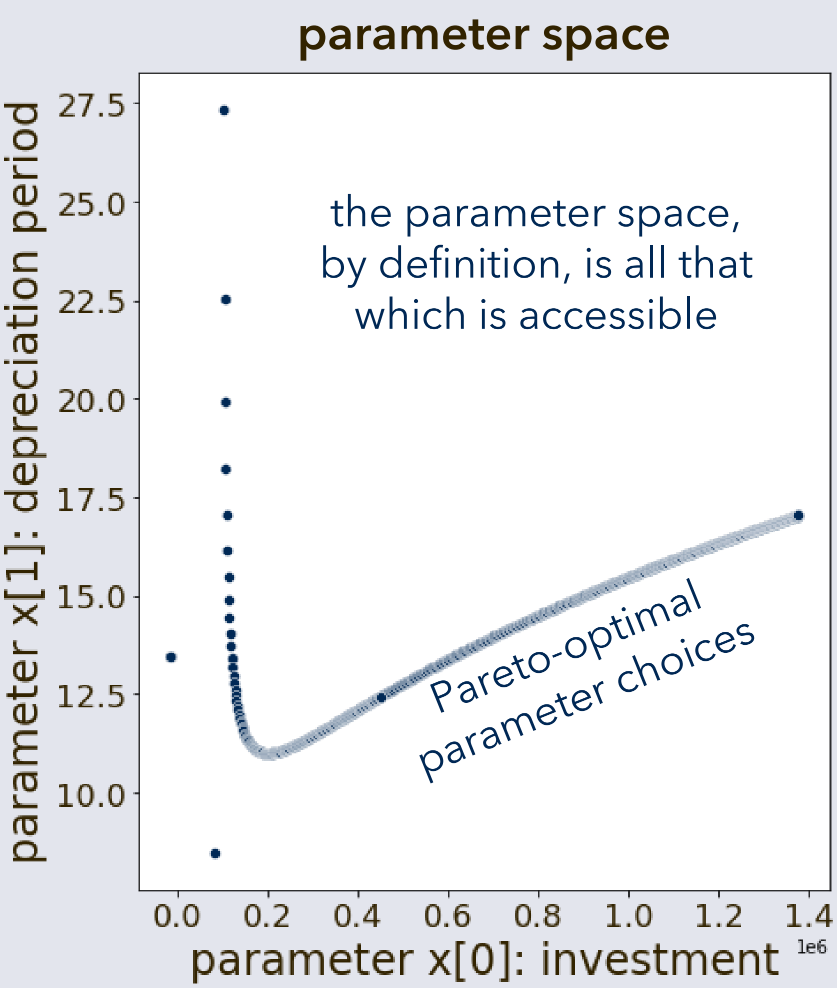
The Jupyter Notebook contains code for constructing a Pareto front.

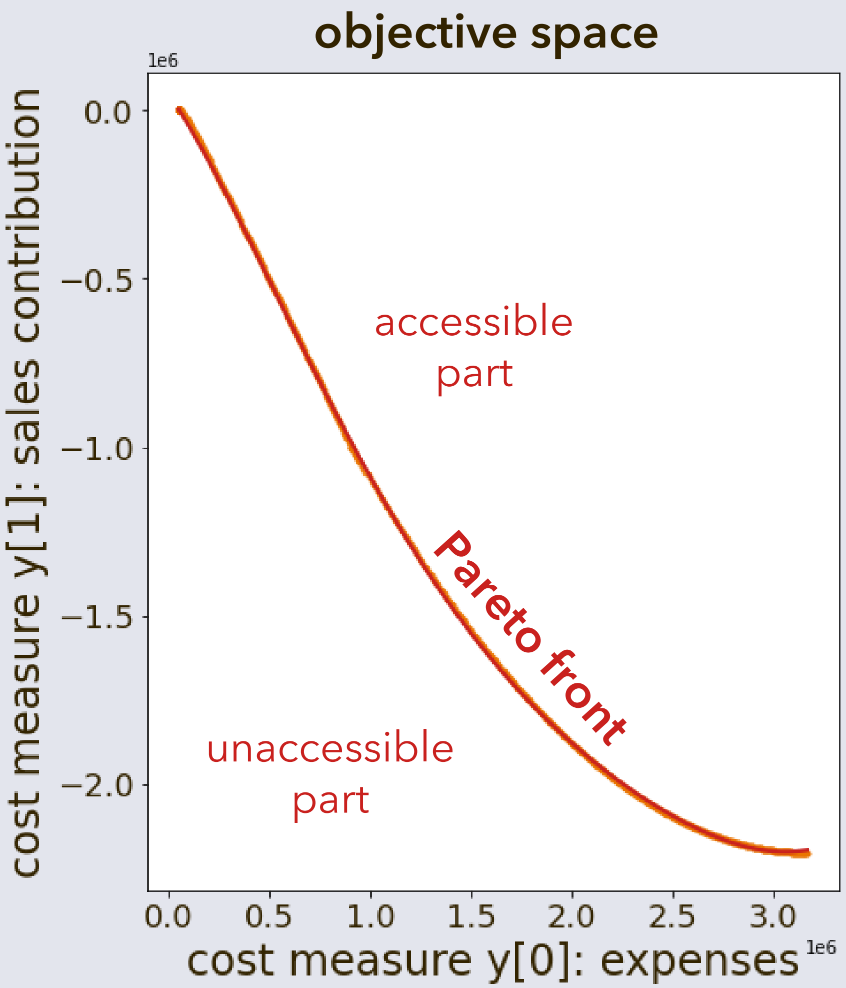
```
total investment         i = GBP 100000
production volume        p = GBP 1000000 per year
depreciation period      d = 7 years

operating cost: GBP 102915.03 per year
depreciation:   GBP 14285.71 per year
prod.cost:      GBP 900000.0 per year
sales contrib.: GBP -1000000.0 per year
==
total deficit:  GBP 17200.74 per year
```

A single-objective cost function was given. However, the evaluation specifies multiple contributions to it.

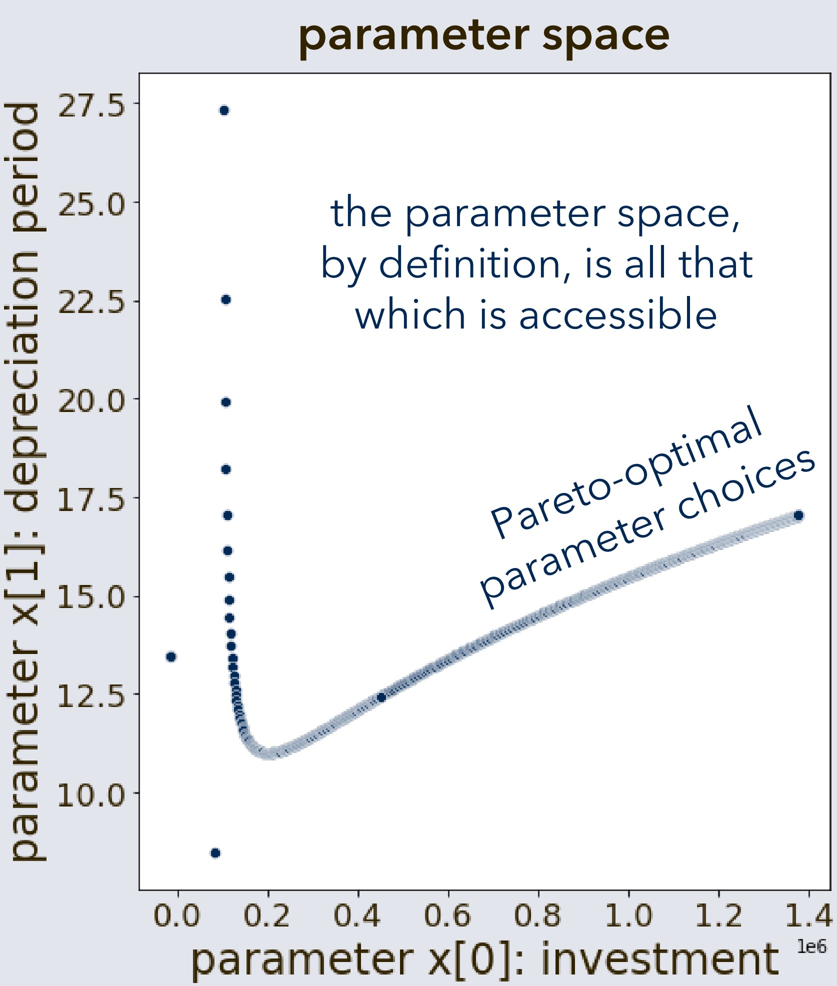
It may make sense to distinguish two objectives: Minimize proper costs, and maximize sales income.

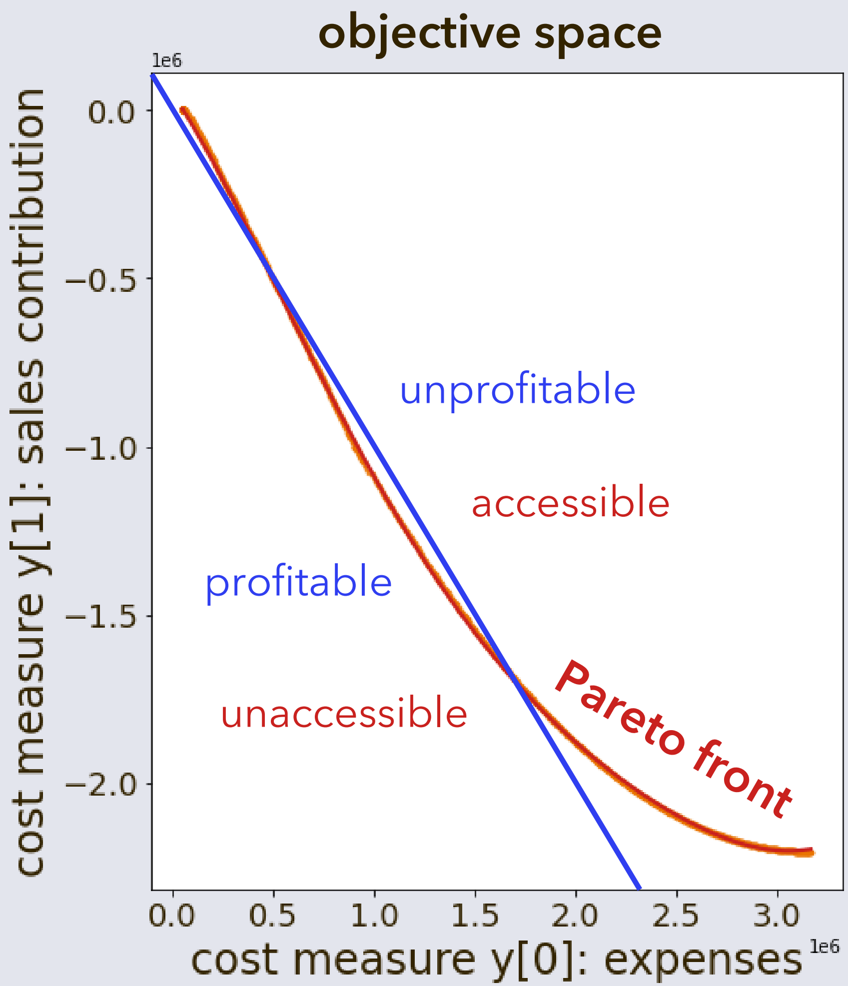
Example decision-support scenario



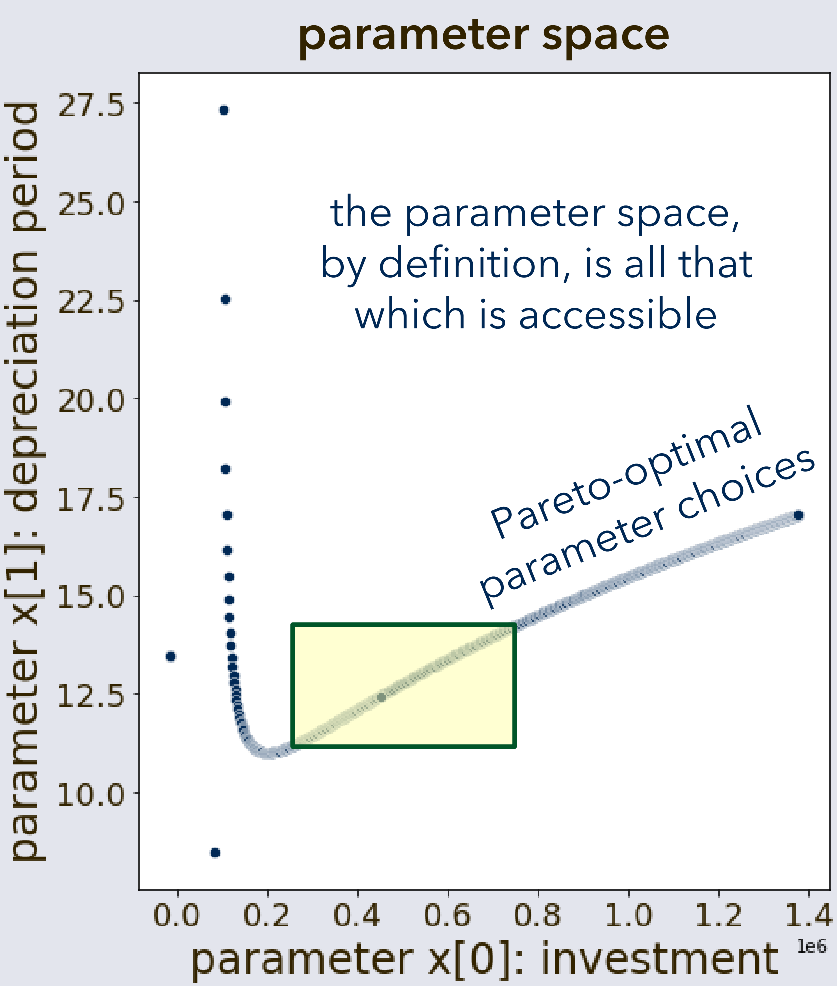


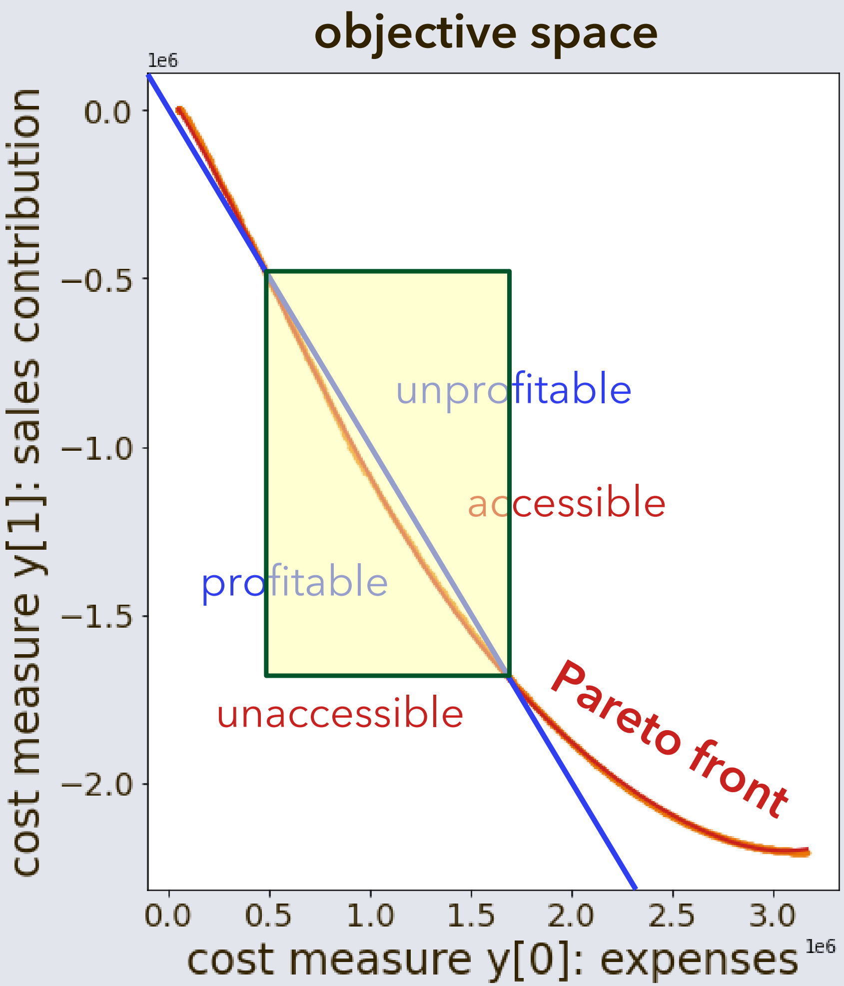
Example decision-support scenario



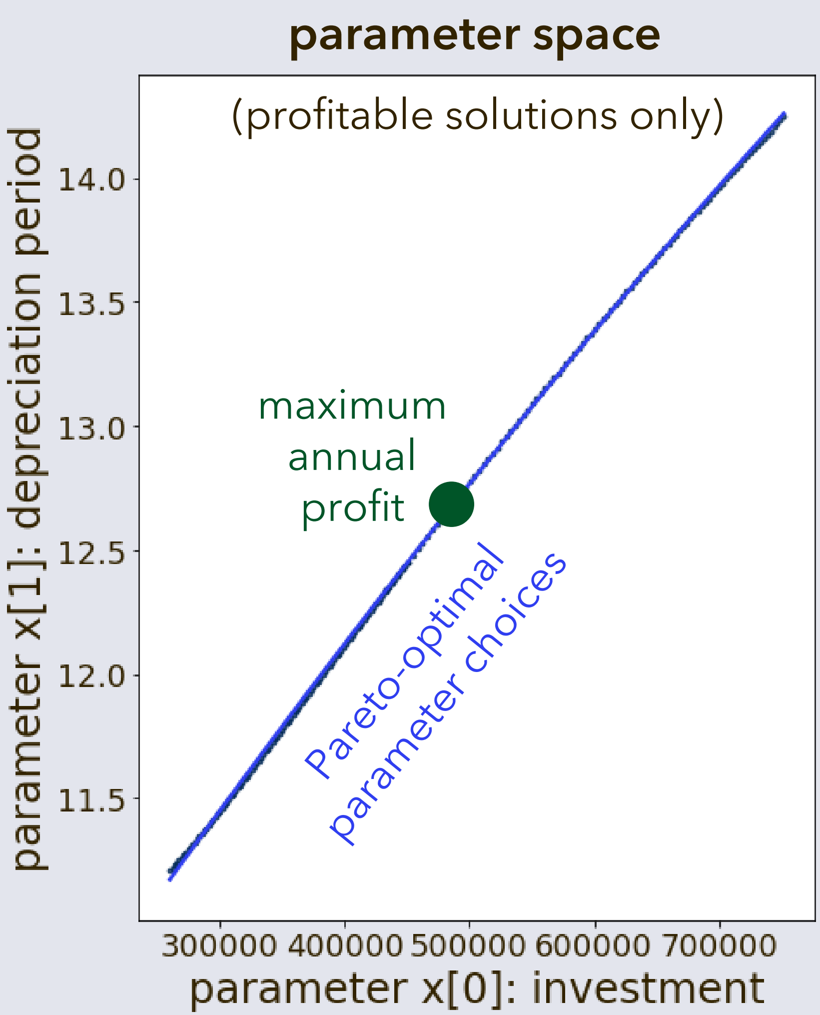


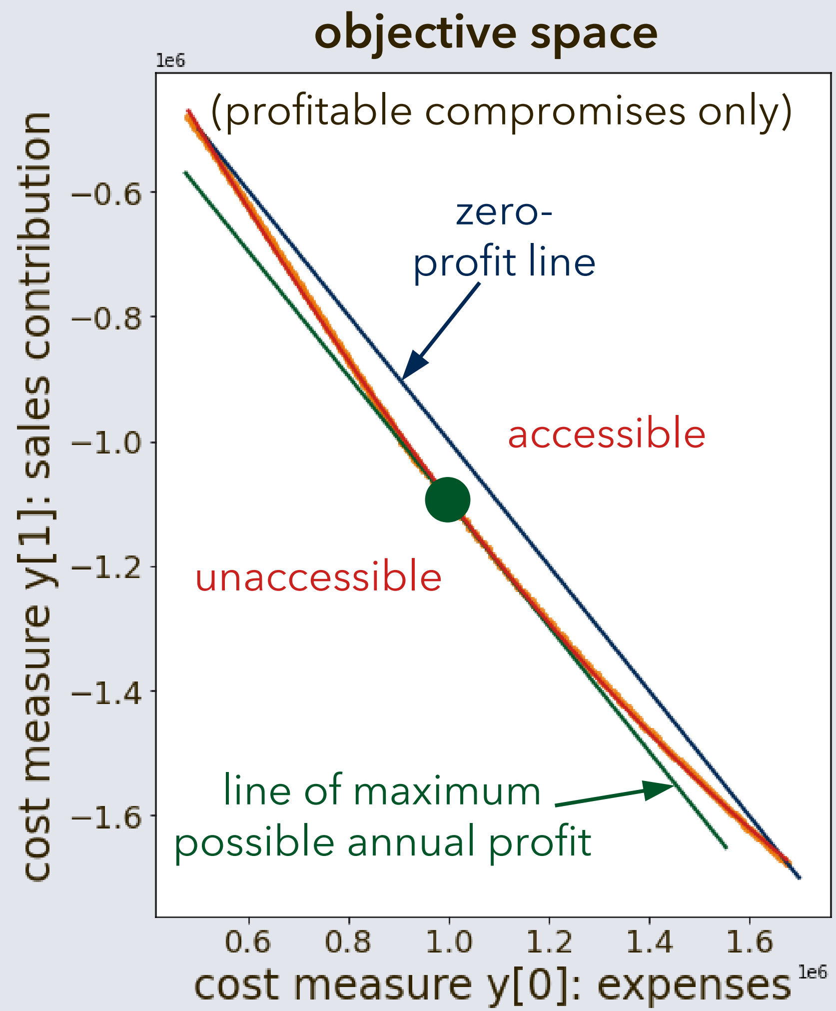
Example decision-support scenario





Example decision-support scenario





Decision support systems

Example: European guidelines on **business decision support systems (BDSS)** for manufacturing relying on AI infrastructures based on materials modelling:[1]

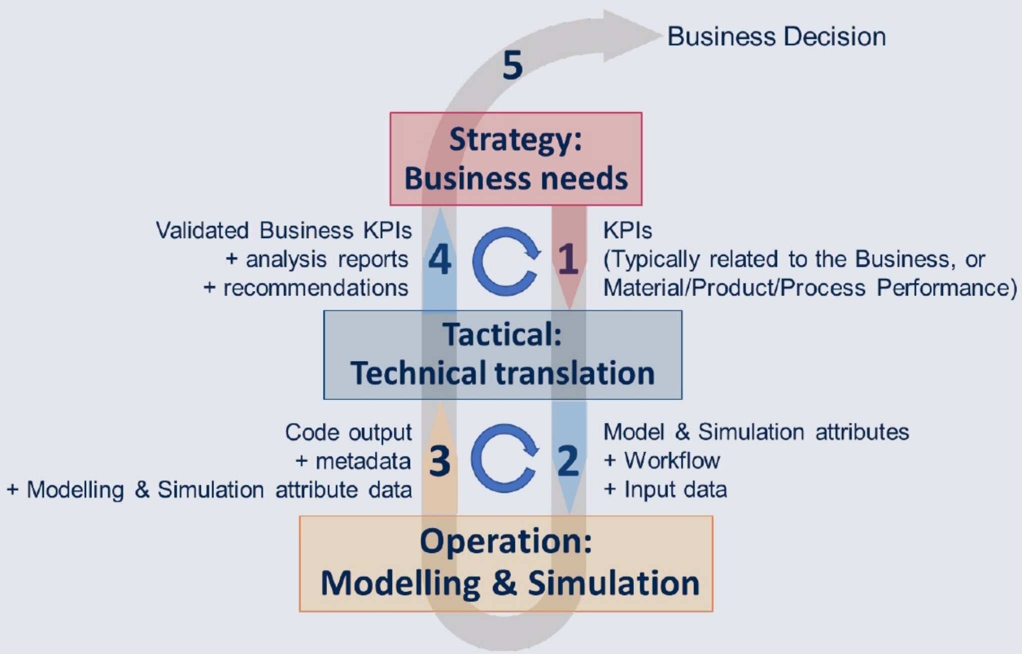


Figure 4: BDSS generic workflow between level of business entities or stakeholders

[1]D. Dykeman *et al.*, *Guideline for Business Decision Support Systems (BDSS) for Materials Modelling*, EMMC ASBL, **2020**.

Reality-to-model "translation" in decision support

Example: European guidelines on **business decision support systems (BDSS)** for manufacturing relying on AI infrastructures based on materials modelling:[1]

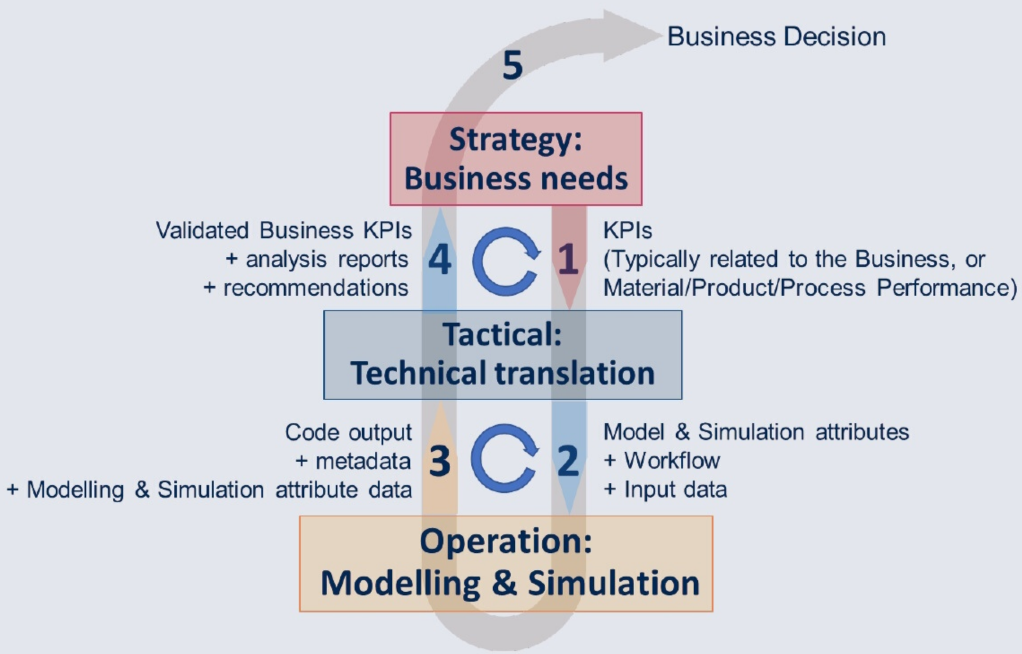
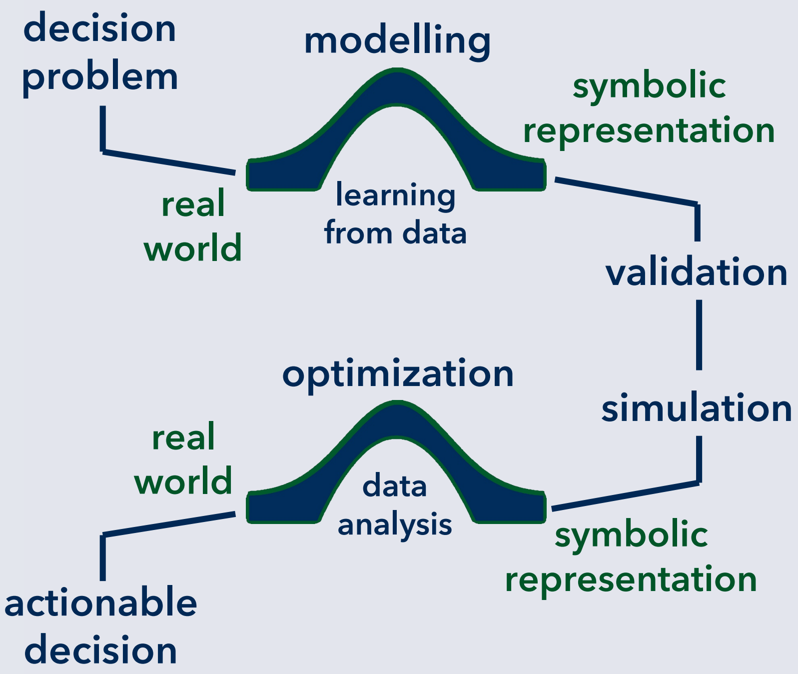


Figure 4: BDSS generic workflow between level of business entities or stakeholders



[1]D. Dykeman *et al.*, *Guideline for Business Decision Support Systems (BDSS) for Materials Modelling*, EMMC ASBL, **2020**. [2]P. Klein *et al.*, *Translation in Materials Modelling: Process and Progress*, **2021**.

Overall utility (or cost)

CO3519

2nd November 2021

Overall utility or cost measure

When to reduce optimization criteria to a single utility or cost measure:

1) If multiple criteria are **not found to be in a genuine conflict** with each other, or the cases where they would come into conflict are not so relevant and can be neglected: Combine them, or select one of them.

   Example: Be a friend of A, and also of B; also, A and B are good friends.

Overall utility or cost measure

When to reduce optimization criteria to a single utility or cost measure:

1) If multiple criteria are **not found to be in a genuine conflict** with each other, or the cases where they would come into conflict are not so relevant and can be neglected: Combine them, or select one of them.

   Example: Be a friend of A, and also of B; also, A and B are good friends.

In such a case, strategies may include:

- Neglecting all criteria with the exception of one, *e.g.*, $f(\mathbf{x}) = y_1$, where $\mathbf{y} = [y_0, y_1]$; since $\boldsymbol{y_0}$ **and** $\boldsymbol{y_1}$ **are correlated**, y_0 is accounted for by y_1.
- Combining the two criteria, *e.g.*, by a **linear combination** such as $f(\mathbf{x}) = 0.4\, y_0 + 0.6\, y_1$, such that y_0 would contribute 40% and y_1 60%.

Overall utility or cost measure

When to reduce optimization criteria to a single utility or cost measure:

1) If multiple criteria are **not found to be in a genuine conflict** with each other, or the cases where they would come into conflict are not so relevant and can be neglected: Combine them, or select one of them.

   Example: Be a friend of A, and also of B; also, A and B are good friends.

2) At the **moment of decision making**, there can only be one criterion.

   Example: Be friends with A, and with B; but A and B hate each other.
   At some point, the decision between A and B needs to be made.

   Select a combination of the criteria based on your priorities, or analyse the Pareto front as a whole to find a solution that is particularly resilient.

Overall utility or cost measure

When to reduce optimization criteria to a single utility or cost measure:

1) If multiple criteria are **not found to be in a genuine conflict** with each other, or the cases where they would come into conflict are not so relevant and can be neglected: Combine them, or select one of them.

   Example: Be a friend of A, and also of B; also, A and B are good friends.

2) At the **moment of decision making**, there can only be one criterion.

   Example: Be friends with A, and with B; but A and B hate each other.
   At some point, the decision between A and B needs to be made.

3) If you are a follower of **utilitarianism** in the British tradition, giving a measure for the "maximum overall good" a moral interpretation.

Linear combinations of optimization criteria

In a **linear combination**, multiple objectives $y_0, y_1, \ldots, y_{n-1}$ are fused to construct a single objective

$$y = c_0 y_0 + c_1 y_1 + \ldots + c_{n-1} y_{n-1},$$

where $c_0, c_1, \ldots, c_{n-1}$ are constant coefficients.

- For $y = 0.3\, y_0 + 0.4\, y_1 + 0.3\, y_2$, the objective y_1 contributes 40% to the overall cost or utility function; the other criteria each contribute 30%.
- Multiplying the coefficients *all* by the same value has no effect on the optimization outcome; only the ratio between them is relevant.

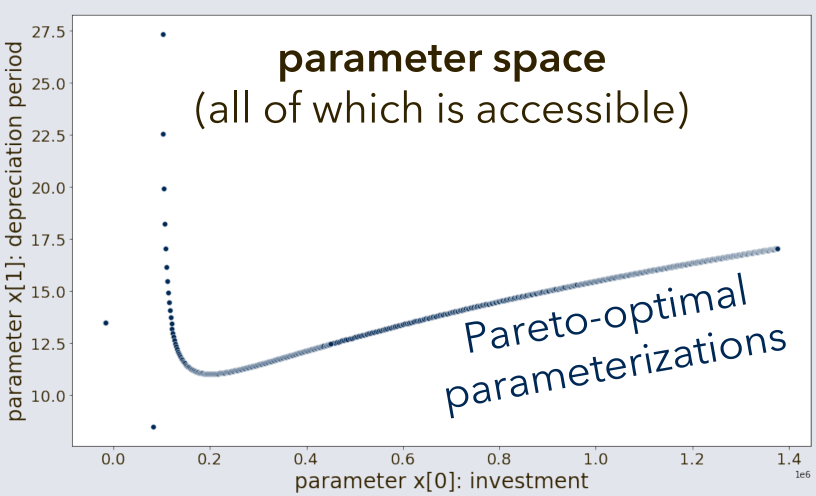
  With $y = 3\, y_0 + 4\, y_1 + 3\, y_2$, the contribution of y_1 is still 40%, *etc.*

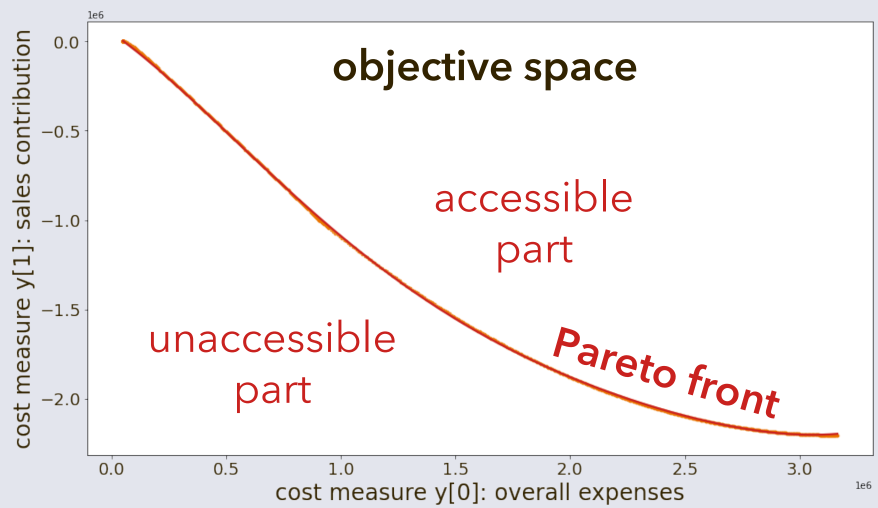
Linear combinations of optimization criteria

In a **linear combination**, multiple objectives y_0, y_1, …, y_{n-1} are fused to construct a single objective

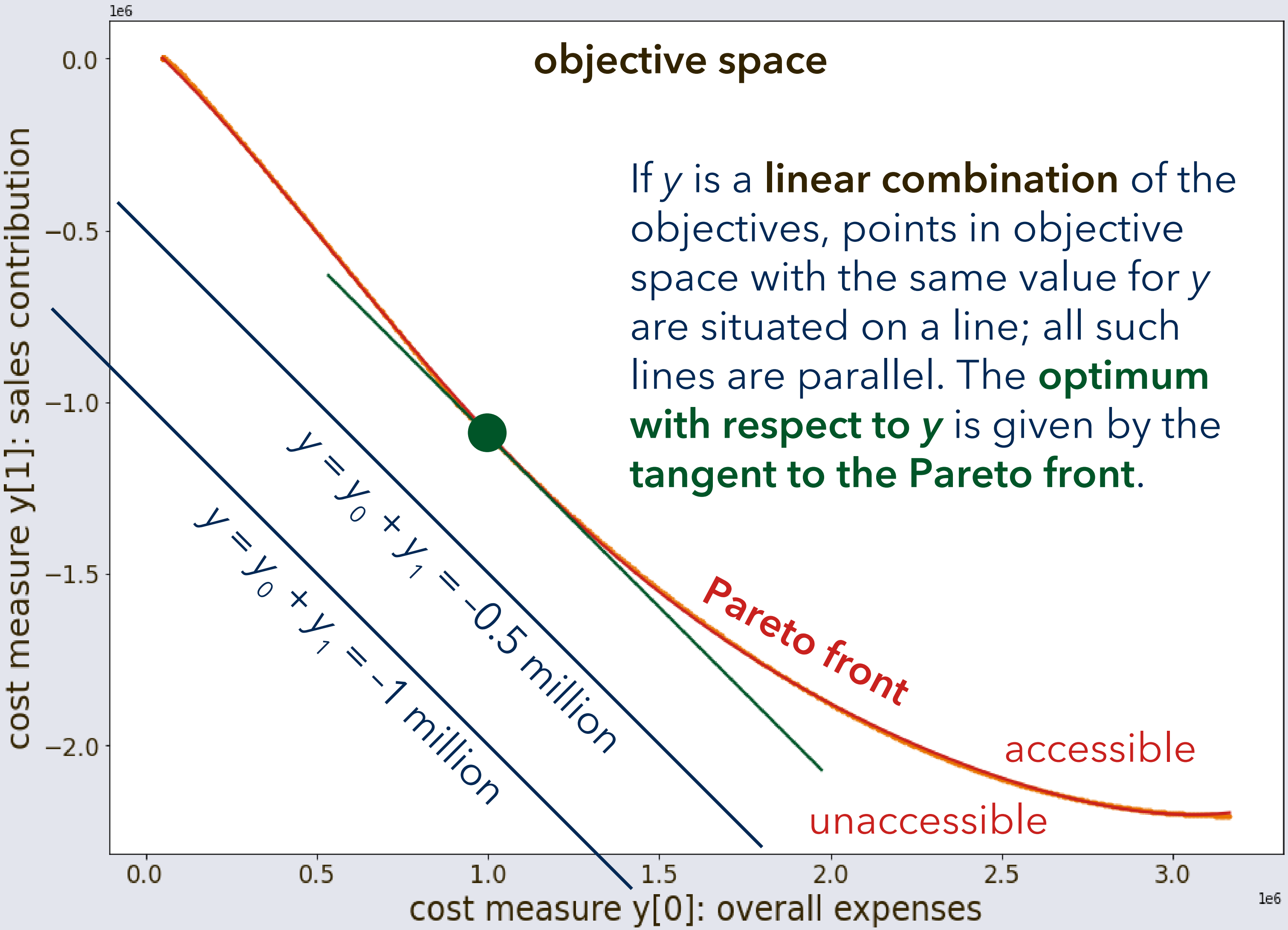
$$y = c_0 y_0 + c_1 y_1 + \ldots + c_{n-1} y_{n-1},$$

where c_0, c_1, …, c_{n-1} are constant coefficients. Points in objective space with the same value for y are then all situated on a line; all these lines are parallel.



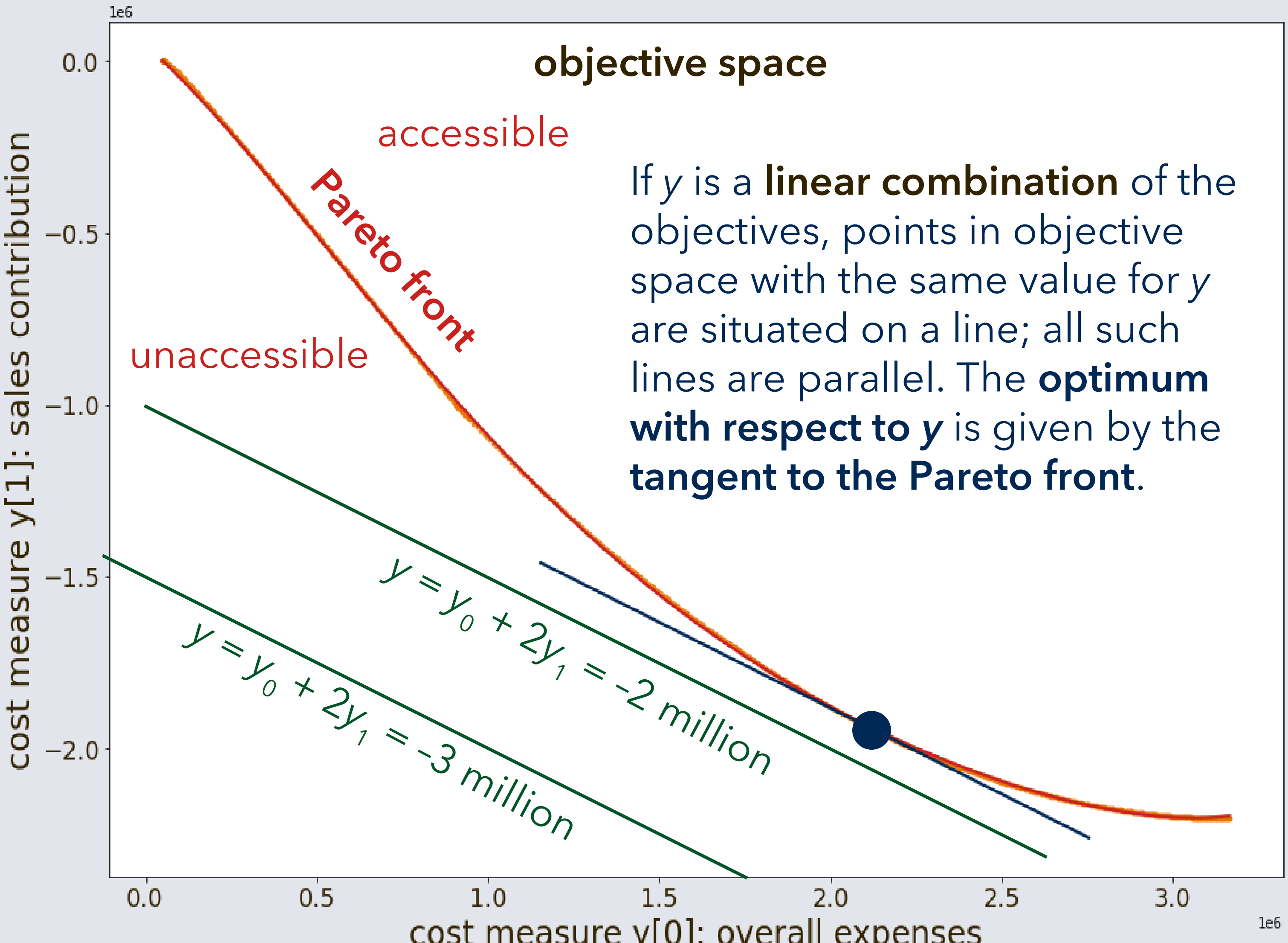


Linear combinations of optimization criteria



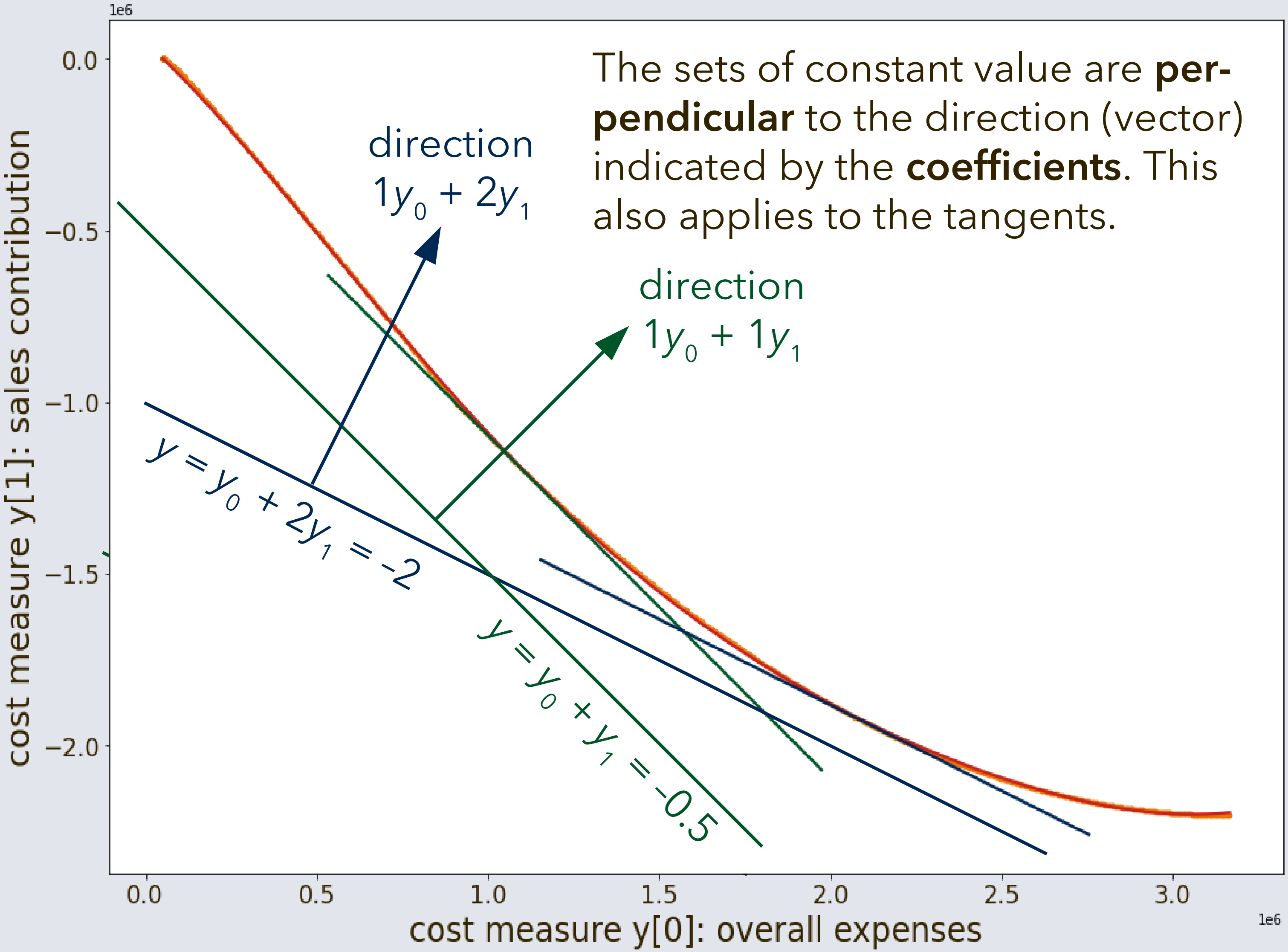
If y is a **linear combination** of the objectives, points in objective space with the same value for y are situated on a line; all such lines are parallel. The **optimum with respect to y** is given by the **tangent to the Pareto front**.

Linear combinations of optimization criteria

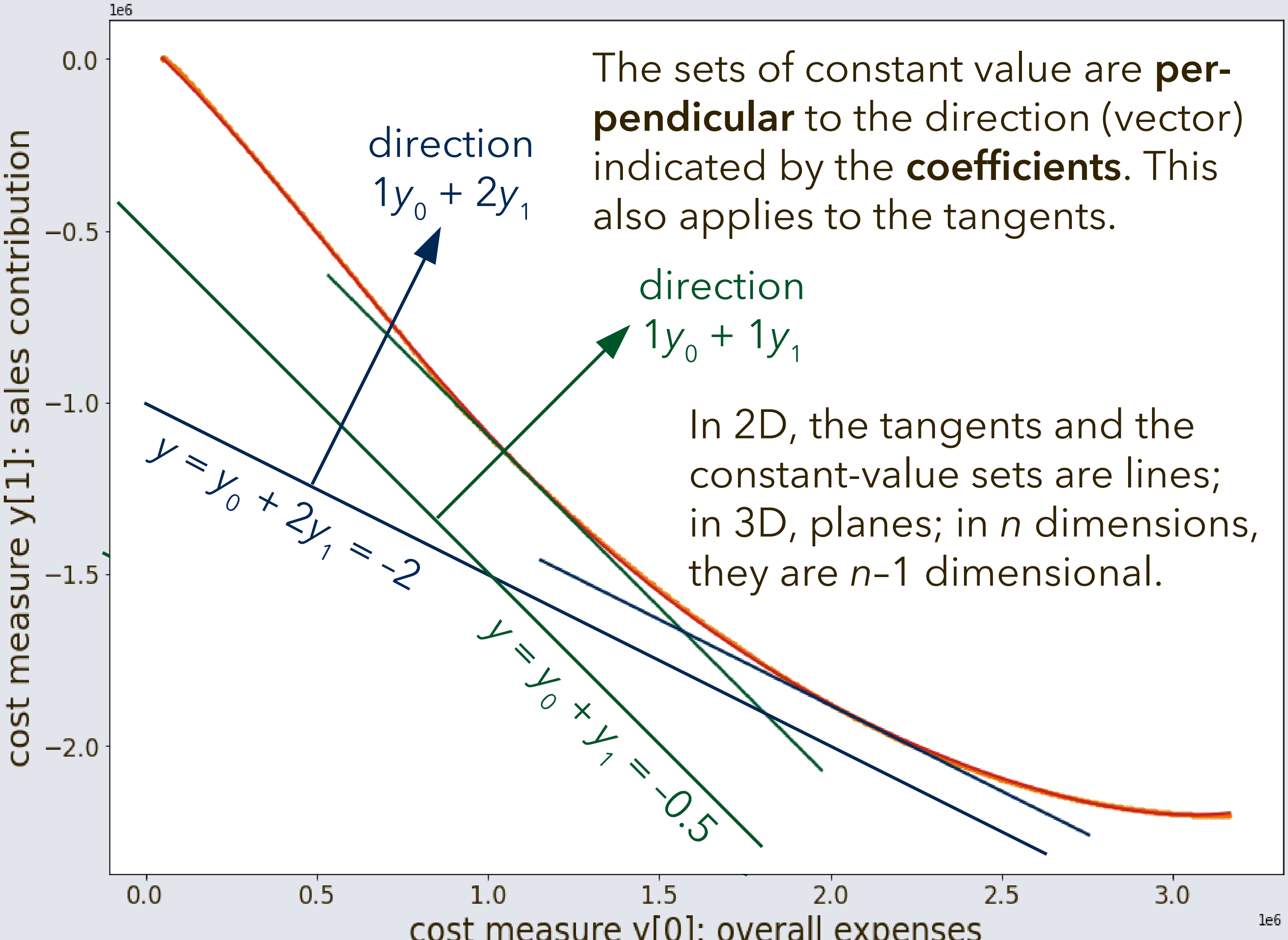


If y is a **linear combination** of the objectives, points in objective space with the same value for y are situated on a line; all such lines are parallel. The **optimum with respect to *y*** is given by the **tangent to the Pareto front**.

Linear combinations of optimization criteria



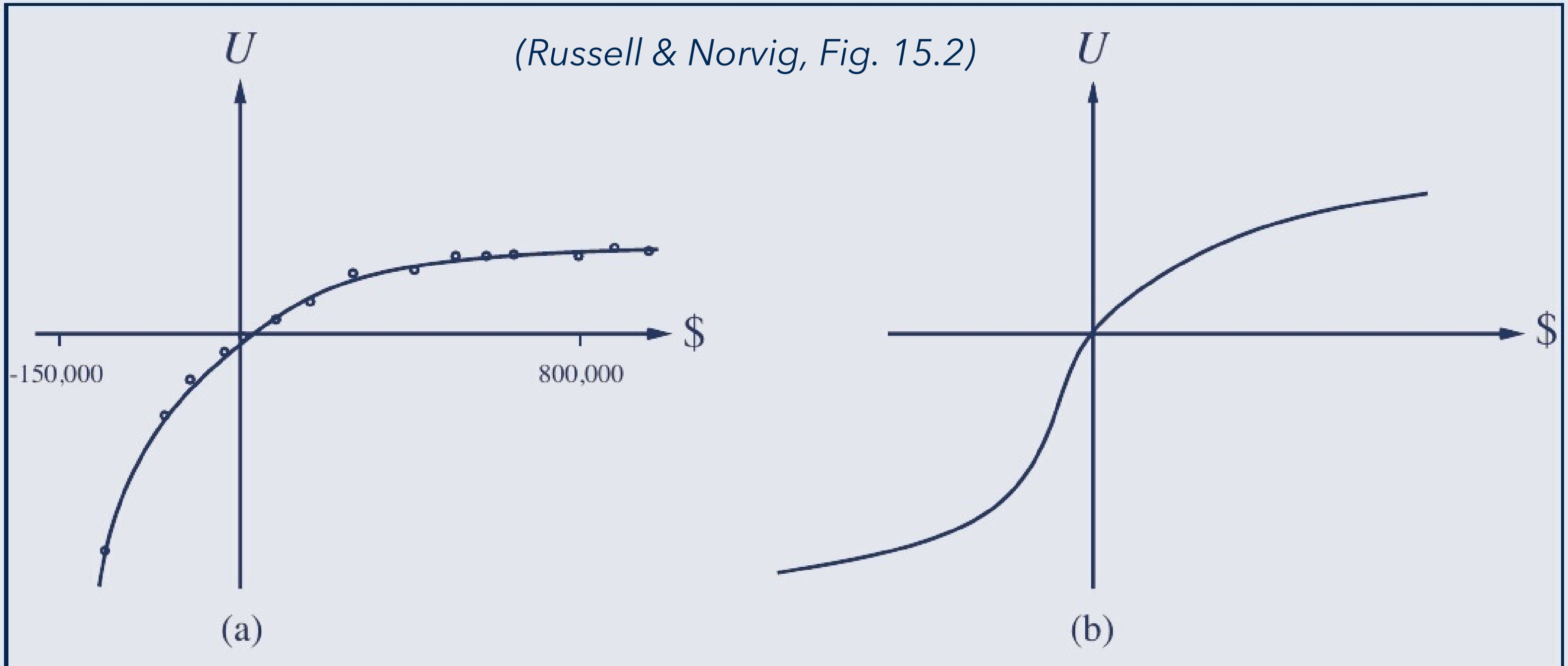
Linear combinations of optimization criteria



The sets of constant value are **perpendicular** to the direction (vector) indicated by the **coefficients**. This also applies to the tangents.

In 2D, the tangents and the constant-value sets are lines; in 3D, planes; in *n* dimensions, they are *n*–1 dimensional.

Nonlinear contributions to utility



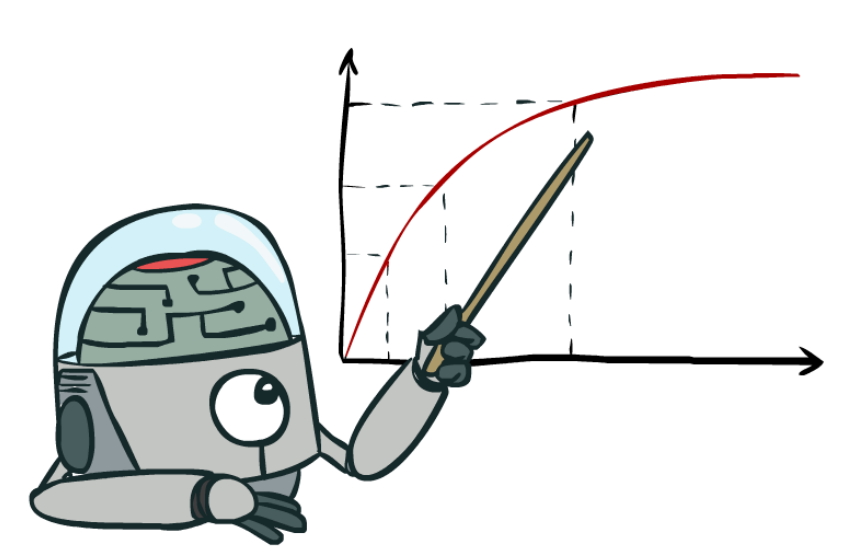
The utility of money. (a) Empirical data for Mr. Beard over a limited range. (b) A typical curve for the full range.

To a billionaire, £10,000 may not even be worth the effort of scheduling an appointment; another objective such as tranquility may be more valuable …

Prerequisites for decision making by optimization

At decision-making stage, if numerical methods are to be applied at all, the problem needs to have been reduced to **single-objective optimization**.

Moreover, the **cost function** (or utility function) **needs to be known**, and any unknown quantities occurring in it must have been eliminated.

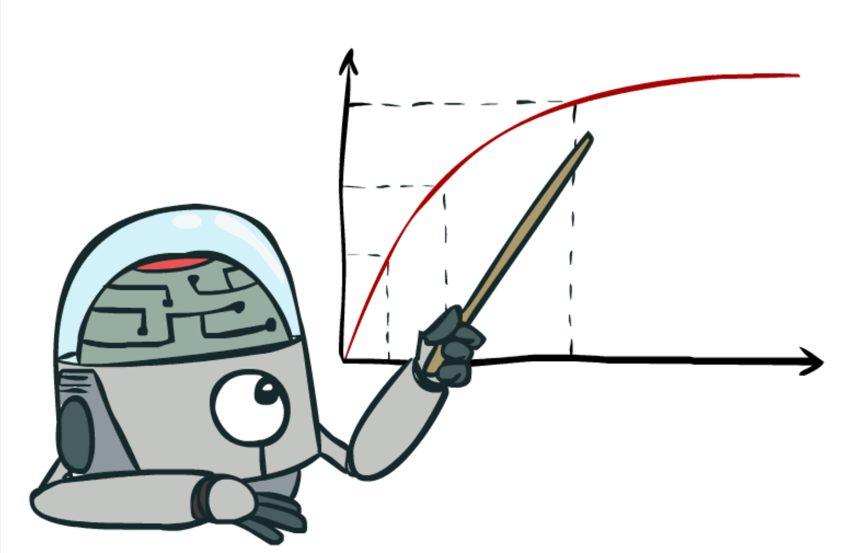


Decision making and the unknown

At decision-making stage, if numerical methods are to be applied at all, the problem needs to have been reduced to **single-objective optimization**.

Moreover, the **cost function** (or utility function) **needs to be known**, and any unknown quantities occurring in it must have been eliminated.

In reality, however, the outcome of a decision often depends on **unknown quantities** such as the weather, unpredictable social developments, or actions of an competitor; in a game, future moves of the opponent.



Approach: Treat the unknown quantities as **random variables**.

Expected utility (or cost)

The outcome of a decision often depends on **unknown quantities**:

$$\mathbf{y} = f(\mathbf{x}, \mathbf{r}) = f(x_0, x_1, \ldots, x_{m-1}, r_0, r_1, \ldots, r_{l-1}), \text{ where the } r_i \text{ are unknown;}$$

These quantities are neither in the decision maker's direct control (*i.e.*, they are not parameters), nor are they influenced indirectly by the decision.

For purposes of decision making and decision support, unknown quantities can be treated as **random variables** whenever not the value itself, but at least a probability distribution $P(\mathbf{r})$ can be reasonably assumed. Then an **expected utility** function (or expected cost function) is obtained by averaging:

$$\mathbf{y} = f(\mathbf{x}) = \text{expected value } E[\, f(\mathbf{x}, \mathbf{r}) \,] = \sum_{\mathbf{r}} P(\mathbf{r}) f(\mathbf{x}, \mathbf{r}).$$

Expected utility (or cost)

The outcome of a decision often depends on **unknown quantities**:

$$\mathbf{y} = f(\mathbf{x}, \mathbf{r}) = f(x_0, x_1, \ldots, x_{m-1}, r_0, r_1, \ldots, r_{l-1}), \text{ where the } r_i \text{ are unknown;}$$

These quantities are neither in the decision maker's direct control (*i.e.*, they are not parameters), nor are they influenced indirectly by the decision.

For purposes of decision making and decision support, unknown quantities can be treated as **random variables** whenever not the value itself, but at least a probability distribution $P(\mathbf{r})$ can be reasonably assumed. Then an **expected utility** function (or expected cost function) is obtained by averaging:

$$\mathbf{y} = f(\mathbf{x}) = \text{expected value } \mathrm{E}[\, f(\mathbf{x}, \mathbf{r}) \,] = \textstyle\sum_{\mathrm{r}} P(\mathbf{r}) f(\mathbf{x}, \mathbf{r}).$$

There, the sum of all probabilities must add up to one: $\sum_{\mathrm{r}} P(\mathbf{r}) = 1.$

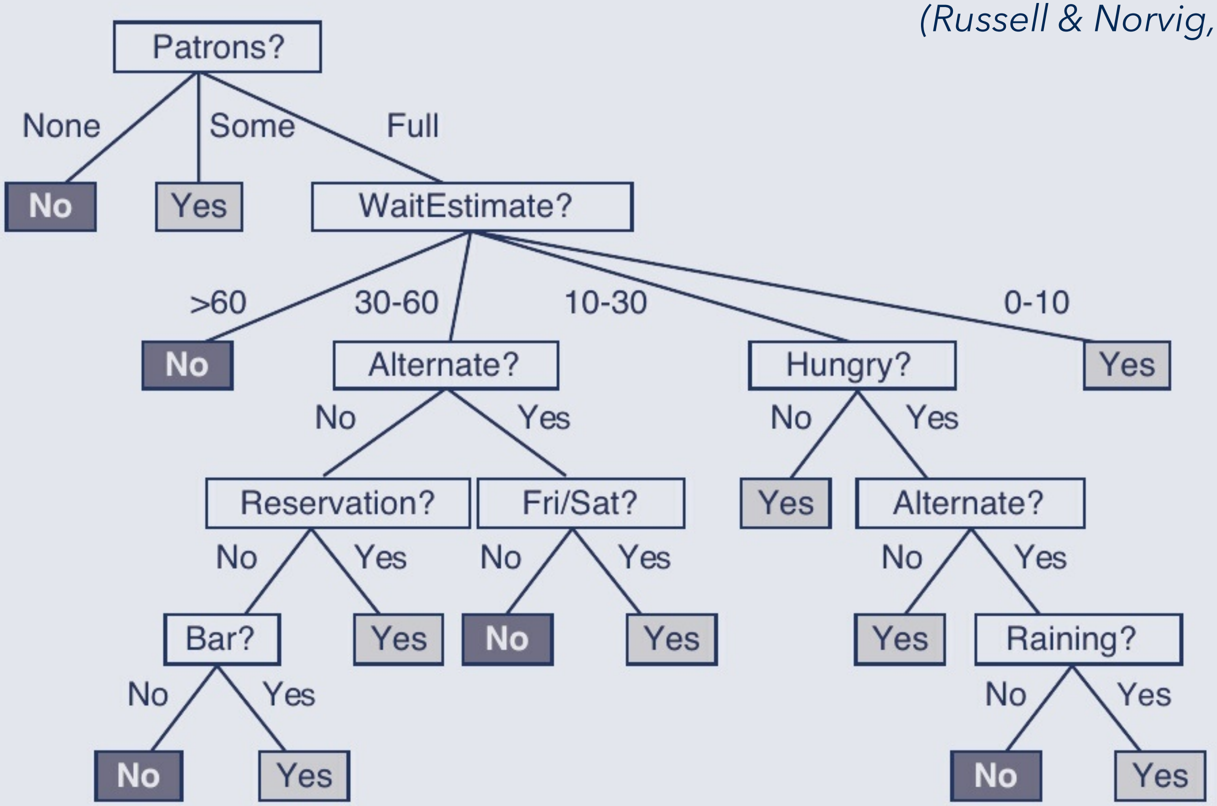
Visualization of decision making

CO3519

2nd November 2021

Decision trees

(Russell & Norvig, Fig. 19.3)



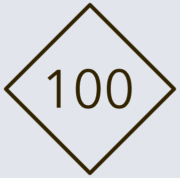
A decision tree for deciding whether to wait for a table.

Decision trees with utility nodes[1]

Barber[1] employs a more expressive notation for decision trees, where three kinds of elements can occur in any order from top to bottom:

rectangles represent **decisions (parameters)**

circles or ellipses represent **unknown quantities**



diamonds contain **contributions to utility**

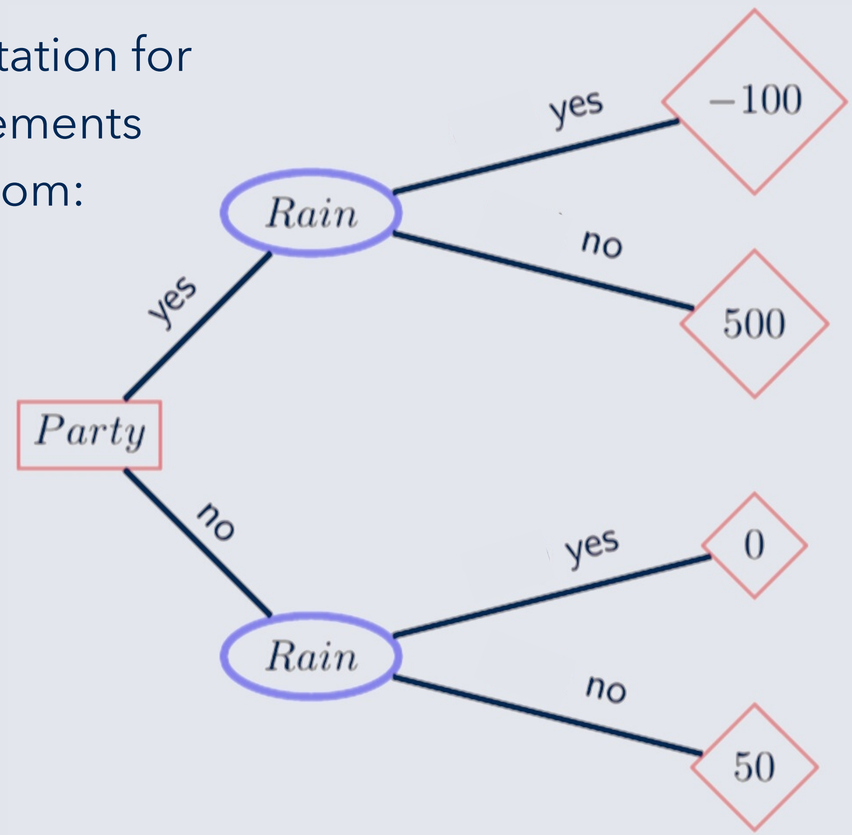


Figure 7.1: A decision tree containing chance nodes (denoted with ovals), decision nodes (denoted with rectangles) and utility nodes (denoted with diamonds).

[1] D. Barber, *Bayesian Reasoning and Machine Learning*, Cambridge Univ. Press, **2012**.

Decision trees with random variables[1]

Barber[1] employs a more expressive notation for decision trees, where three kinds of elements can occur in any order from top to bottom:

Party — rectangles represent decisions (parameters)

Rain — circles or ellipses can represent **random variables**

100 — diamonds contain contributions to utility

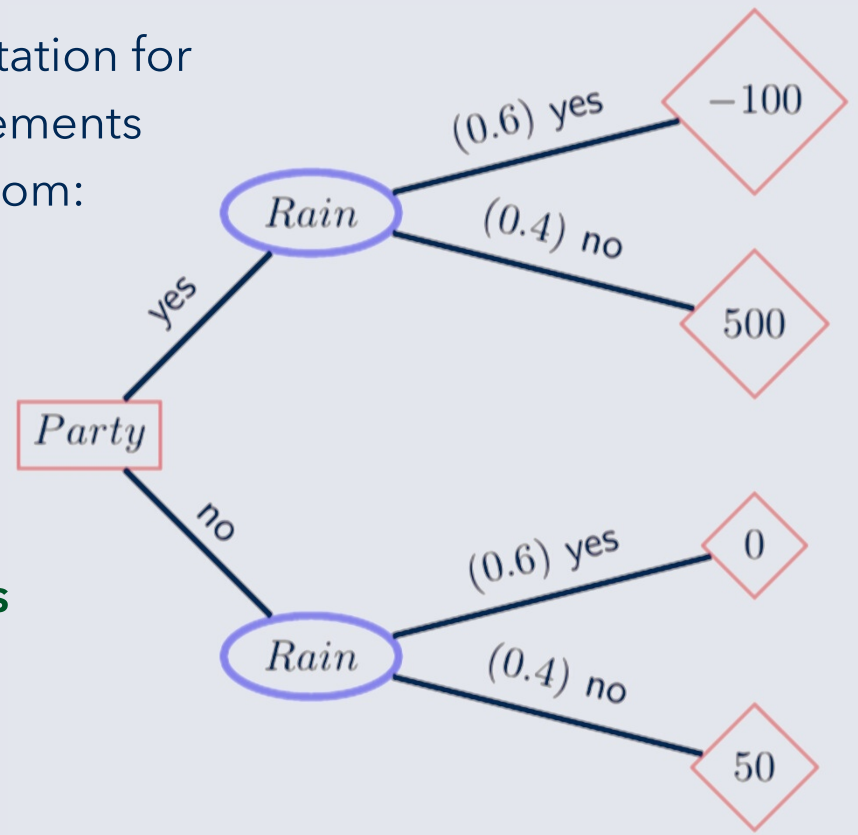


Figure 7.1: A decision tree containing chance nodes (denoted with ovals), decision nodes (denoted with rectangles) and utility nodes (denoted with diamonds).

[1] D. Barber, *Bayesian Reasoning and Machine Learning*, Cambridge Univ. Press, **2012**.

Decision trees with random variables[1]

Evaluation of the example decision tree:

If I go to the party, it will rain (60% chance) or not rain (40%); if it rains, utility is –100, otherwise 500. The **expected utility** is:

yes $\leftarrow$ $0.6 \cdot (-100) + 0.4 \cdot 500 = 140.$

If I do not go, it will rain (60% chance) or not rain (40%); if it rains, utility is 0, otherwise 50. The **expected utility** is:

no $\leftarrow$ $0.6 \cdot 0 + 0.4 \cdot 50 = 20.$

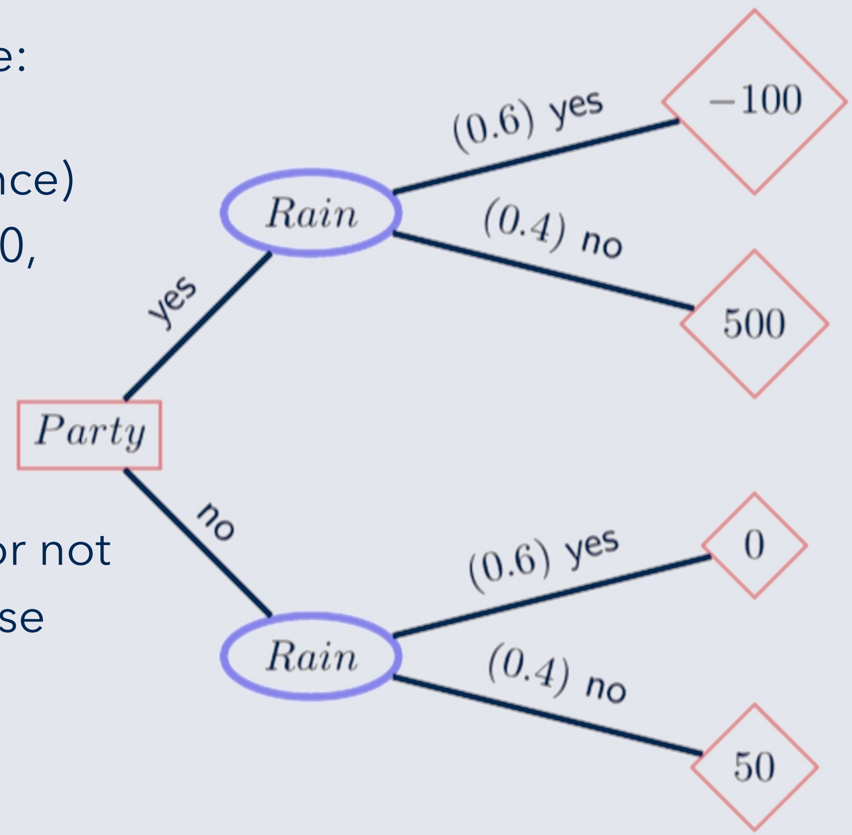


Figure 7.1: A decision tree containing chance nodes (denoted with ovals), decision nodes (denoted with rectangles) and utility nodes (denoted with diamonds).

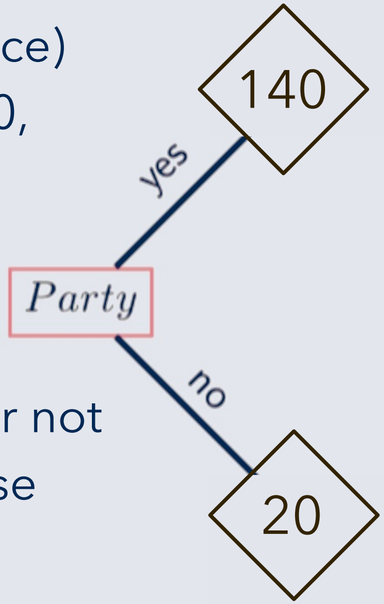
[1]D. Barber, *Bayesian Reasoning and Machine Learning*, Cambridge Univ. Press, **2012**.

Decision trees with random variables[1]

Evaluation of the example decision tree:

If I go to the party, it will rain (60% chance) or not rain (40%); if it rains, utility is –100, otherwise 500. The **expected utility** is:

yes $\leftarrow$ $0.6 \cdot (-100) + 0.4 \cdot 500 = 140.$



If I do not go, it will rain (60% chance) or not rain (40%); if it rains, utility is 0, otherwise 50. The **expected utility** is:

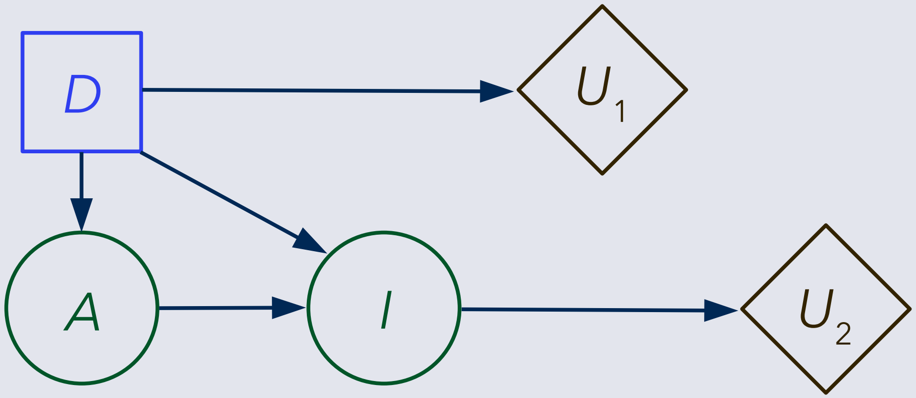
no $\leftarrow$ $0.6 \cdot 0 + 0.4 \cdot 50 = 20.$

Since utility must be maximized, I should go.

[1]D. Barber, *Bayesian Reasoning and Machine Learning*, Cambridge Univ. Press, **2012**.

Influence diagrams

Influence diagrams (also: decision networks) visualize how different quantities are connected to each other in a decision-making process.



(Example based on Barber, Fig. 7.6)

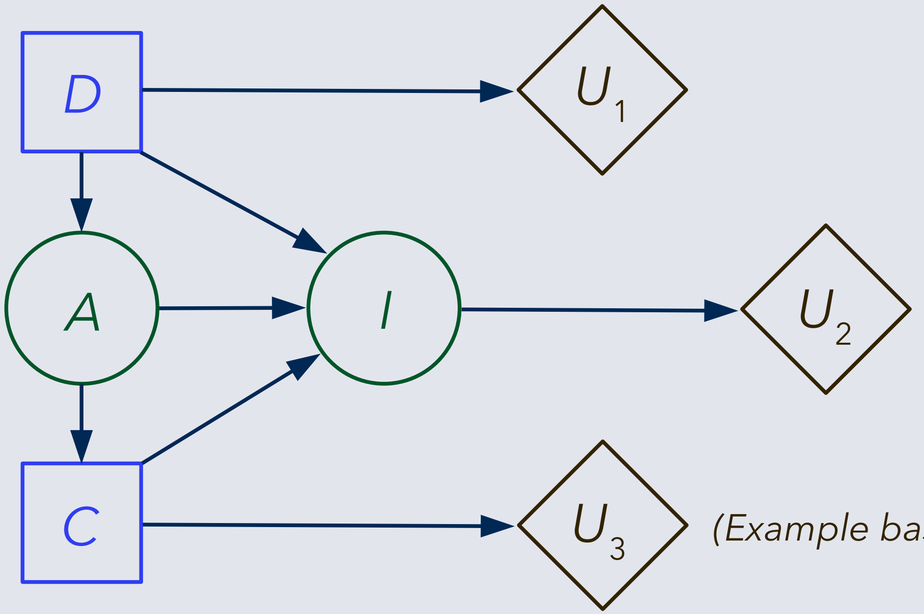
D: Should I work on a doctorate?

A: Academic recognition measure
I: Life income

U_1, U_2: Contributions to utility.

Influence diagrams

Influence diagrams (also: decision networks) visualize how different quantities are connected to each other in a decision-making process.



D: Should I work on a doctorate?
C: Should I found a consultancy?

A: Academic recognition measure
I: Life income

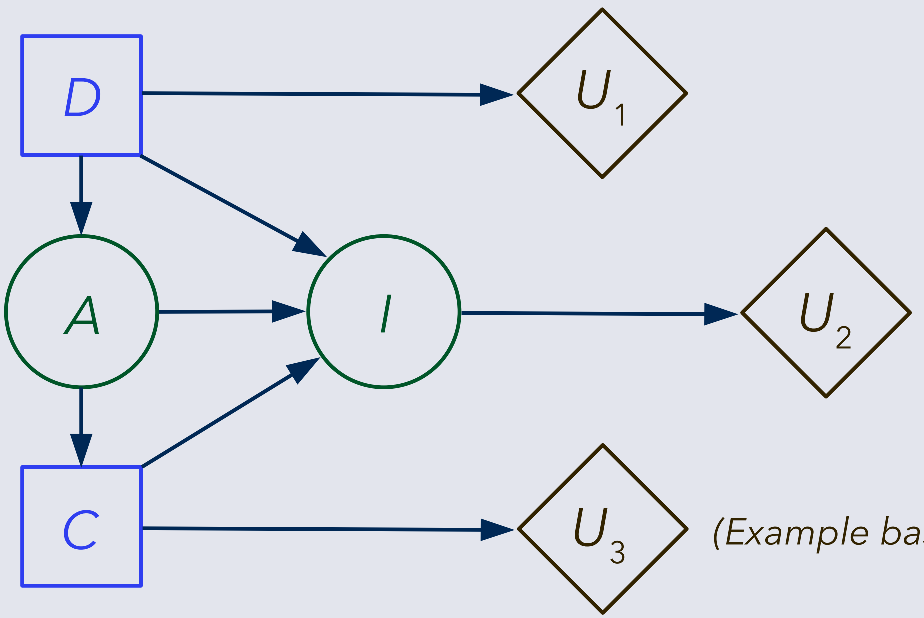
U_1, U_2, U_3: Contributions to utility.

(Example based on Barber, Fig. 7.6)

Decision trees are only applicable to qualitative (discrete) decision making, such as yes/no choices. Influence diagrams are more general: They are also suitable for **quantitative decision making** based on continuous optimization.

Influence diagrams

Influence diagrams (also: decision networks) visualize how different quantities are connected to each other in a decision-making process.



D: Should I work on a doctorate?
C: Should I found a consultancy?

A: Academic recognition measure
I: Life income

U_1, U_2, U_3: Contributions to utility.

(Example based on Barber, Fig. 7.6)

Observation: Whereas a decision tree alone is enough to make a decision, **an influence diagram visualizes a process** by which quantities are evaluated. For the diagram to represent a valid process, it **must not contain any cycles**.

University of Central Lancashire
UCLan

CO3519
Artificial Intelligence

Decision support
Overall utility (or cost)
Visualization of decision making

Where opportunity creates success