BDMA - Decision Modeling

Jose Antonio Lorencio Abril

Fall 2023

CentraleSupélec

université
PARIS-SACLAY

Professor: Brice Mayag

Student e-mail: jose-antonio.lorencio-abril@student-cs.fr

This is a summary of the course *Decision Modeling* taught at the Université Paris Saclay - CentraleSupélec by Professor Brice Mayag in the academic year 23/24. Most of the content of this document is adapted from the course notes by Mayag, [1], so I won't be citing it all the time. Other references will be provided when used.

Contents

1 Introduction

1.1 Models

There are many definition of what a model is, depending on the perspective used. For example, a model can be understood as a '*standard or example for imitation or comparison*' or as a '*person employed to wear clothing ot pose with a product for purposes of display and advertising*'. Nonetheless, in the context of decision modeling, which is the one in which we are interested, the definition of model is:

Definition 1.1. A **model** is a *simplified representation of a system or phenomenon, with any hypotheses required to describe the system or explain the phenomenon, often mathematically.*

Models are useful to enhance our understanding of the world to improve our decision making, and they enable us to elaborate a scientific methodology to solve a problem in a duplicable way and with the aim of reducing bias in mind.

A model is said to be **deterministic** if the outcomes are precisely determined through known relationships among states and events. This kind of models always produce the same output when given the same input. On the other hand, a model is **probabilistic** (or stochastic) when all the data that it tries to explain is not known with certainty.

For example, the Newtonian model for gravity is deterministic, while a prediction model for college acceptance is probabilistic.

Deterministic models are used in domains such as Multi-Attribute Decision Making or Linear Programming, among others; while probabilistic models are used in queuing problems, simulations, etc.

We can also classify the data that is used to define a model into **qualitative** and **quantitative**. The former refers to data that is expressed in terms of words, while the latter is data easily expressed using numbers. An example of qualitative data is the hair color of people in class, and for quantitative data is the height of people in class.

We can be more precise in our wording, and call the models that we are talking about **formal models**, which refer to those models that provide a precise statement of the components of the model and their relationships, usually by means of mathematical equations. This make them easy to communicate precisely and the ability to give replicable results. However, being formal does not mean being true. A model can fail to represent the reality that it tries to describe.

1.2 Decision Theory and Decision Analysis

Definition 1.2. A **decision** is a choice that is made about something after thinking about several possibilities.

Decisions appear in many domains, including Mathematics, Economics, Computer Sciences, Psychology,...

Definition 1.3. **Decision Analysis** consists in trying to provide answers to questions raised by actors involved in a decision process using a model.

B. Roy

In the previous definition, a **Decision Process** refers to a strategy of intervention, such as aid, communication or justification, among others. There are many ways to provide decision aid and no single way to compare methods. This, together with the fact that different models may lkead to different recommendations, makes it hard to assess when a Decision Analysis model is 'good' or, more appropriately, 'suitable'.

Therefore, we cannot compare decision making to solving a well-defined problem, as the former is highly dependent on opinions, interests and, more generally, different human factors involved. In every decision process, there are several possible interventions, among which we can find imagining compromises, communicating, coordinating, controlling, motivating or conducting change.

There are many different models used in Decision Analysis nowadays, with the advantages of:

- Providing a clear language that can be leveraged as a communication tool
- Capturing the essence of a situation, acting a structuration tool
- Answering '*what-if*' questions, serving as a exploration tool

On the other hand, their drawbacks are their high complexity and opaqueness. In addition, in many situations people could argue that such models are not necessary because they already know how to take decisions and they would over-complicate the process; or would ask for higher-level explanations or ideas that are not suitable for formalization; or would rather rely on their intuition.

1.3 Main Steps of Developing a Decision Model

1. **Formulation**: translate the problem scenario into a mathematical model.

   This involves the definition of the problem and the development of a decision model, i.e., the definition of the **variables** or measurable quantities that vary, and the **parameters** or measurable quantities inherent to the problem.

2. **Solution**: solve the mathematical expressions from the formulation.

   This process involves the development of the solutions by correctly manipulating the model to arrive at the best solution, and the testing of the solution, to check that it works as expected and meets the expectations.

3. **Interpretation**: discover the implications of the results.

   This is usually done by conducting **sensitivity analysis**, i.e., testing the different outcomes obtained under a variety of states; and **implementing results**, enacting the solutions and monitoring the performance.

The outlined process is very high level, and there are many possible problems that can arise:

- Defining the problem: we can find conflicting viewpoints that impact differently the stakeholders.
- Model development: it is not always easy to find the formal model that describes the problem at hand, and it is usual to make adaptations.
- Acquering data: can be hard in some scenarios, as well as checking its validity and correctness.
- Developing a solution: we can find many limitations, such as only finding one answer, finding approximate answers, prohibitive computing times,...
- Implementation: it is crucial that the solution is feasible to be implemented, both from a managerial point of view and from the user perspective.

1.4 Decision's Algorithm & Transparency

Decisions made by algorithms can be opaque because of technical and social reasons, in addition to being made purposely opaque to protect intellectual property.

Definition 1.4. An **algorithm** is a sequence of instructions, typically used to solve a class of problems or perform a computation.
It must be:

- Finite: it must eventually solve the problem.
- Well-defined: the series of step must be precise and understandable.
- Effective: it must solve all cases of the problem for which it was defined.

Usually, we find contradictory objectives when developing an algorithm, because simpler algorithms are usually time intensive, while algorithms that are very efficient are very complex are hard to understand.

Definition 1.5. Algorithmic Transparency is the principle that the factors that influence the decisions made by algorithms should be transparent to the people who use, regulate and are affected by systems that employ those algorithms.

This concept is opennes about the purpose, structure and underlying actions of the algorithms used to search for, process and deliver information. Two important properties of transsparency are:

- **Explanability**: systems and institutions that use algorithm decision making are encouraged to produce explanations regarding both the procedures followed by the algorithm and the specific decisions that are made. This is specially relevant in public policy contexts.
- **Accountability**: institutions should be held responsible for decisions made by the algorithms they use, even if it is not feasible to explain in detail how the algorithms produce their results.

2 Preferences as binary relations

2.1 Definitions

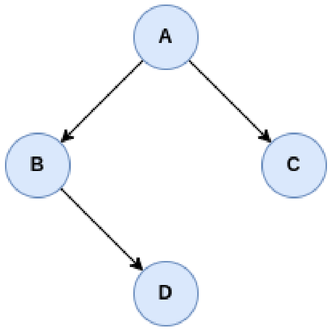
Definition 2.1. Given a set X, a **binary relation**, R, is a subset of ordered pairs of elements in X:

$$R \subseteq X \times X.$$

We can write $(x, y) \in R$ or, equivalently, xRy.

Relations can be expressed as directed graphs. For instance, a relation R of a set X can be represented as the graph $G_R = (N_X, E_R)$, where N_X are the nodes, representing each element in X and E_R are the edges, representing each pair in R. The edges are constructed in such a way that $e = (x, y) \in E_R \iff xRy$.

Example 2.1. Let $X = \{a, b, c, d\}$ and $R = \{(a, b), (a, c), (b, d)\}$, then we can represent this by G_R as:



Another way to represent relations is using matrices. We can construct a matrix M_R by

$$M_R = (m_{xy})_{(x,y) \in X},$$

where

$$m_{xy} = \begin{cases} 1 & if\ xRy \\ 0 & if\ not\,(xRy) \end{cases}.$$

Example 2.2. The previous example can be represented with the following matrix:

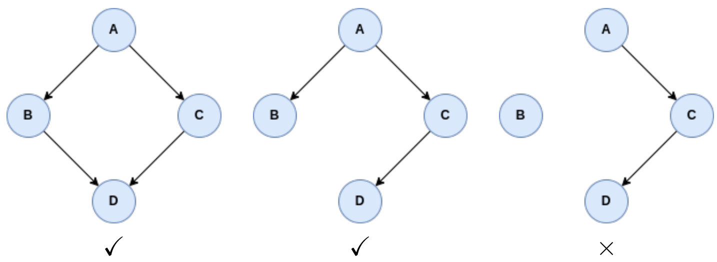
$$M_R = \begin{array}{c} \\ a \\ b \\ c \\ d \end{array} \begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}.$$

Depending on how a relation is constructed, it can possess different properties. Some interesting properties are defined as follows:

Definition 2.2. A binary relation R on a set X is said to be:

- **Reflexive** if $xRx, \forall x \in X$.
- **Irreflexive** if $not\,(xRx)\,, \forall x \in X$.
- **Complete** if for every $x, y \in X$, we have xRy or yRx (or both).
- **Weakly complete** if for every $x, y \in X, x \neq y$, we have xRy or yRx (or both).
- **Symmetric** if $[xRy \implies yRx]\,, \forall x, y \in X$.
- **Asymmetric** if $[xRy \implies not\,(yRx)]\,, \forall x, y \in X$.
- **Antisymmetric** if $[xRy \wedge yRx \implies x = y]\,, \forall x, y \in X$.
- **Transitive** if $[xRy \wedge yRz \implies xRz]\,, \forall x, y, z \in X$.
- **Negatively transitive** if $[not\,(xRy) \wedge not\,(yRz) \implies not\,(xRz)]\,, \forall x, y, z \in X$.
- **Semi-transitive** if $[xRy \wedge yRz \implies xRt \vee tRz]\,, \forall x, y, z, t \in X$.

Example 2.3. A semi-transitive relation example.



In addition, we can define paths and cycles on relations, analogously as how it is done in graph theory:

Definition 2.3. A **path** from $x \in X$ to $y \in X$ exists if there are $x_1, ..., x_n \in X$ such that

$$x = x_1 R x_2 R \ldots R x_{n-1} R x_n = y.$$

A path is called a **cycle** if the it goes from x to x.

For every relation, we can extract two subrelations, as its symmetric part and its asymmetric part:

Definition 2.4. Given a binary relation R on X, we can define its **symmetric part**, I, as

$$xIy \iff [xRy \wedge yRx]\,,$$

and its **asymmetric part**, P, as

$$xPy \iff [xRy \wedge not\,(yRx)]\,.$$

The symmetric part is denoted by I because we can understand this as the all the indiferent pairs of R. In others words, if we understand a relation as a preference over the elements in X, then xRy would mean x *is at least as convenient as* y. Therefore, if we have xRy and yRx, we could think of them as equally convenient, so the decision between them is indiferent. On the other hand, the asymmetric part is denoted by P, from preference, following a similar reasoning.

When we have two different relations, R and R', on the same set, X. We can define their concatenation:

Definition 2.5. Let R and R' be two relations on X. We define their **concatenation** as

$$xR \bullet R'y \iff \exists z \in X : [xRz \land zR'y] .$$

The following proposition establishes different relationships between the concepts we have seen so far:

Proposition 2.1. *Let R be a binary relation on X. Then:*

- *R transitive $\implies$ $R \bullet R \subset R$.*
- *R asymmetric $\implies$ R irreflexive.*
- *R complete $\iff$ R reflexive and weakly complete.*
- *R asymmetric and negative transitive $\implies$ R transitive.*
- *R complete and transitive $\implies$ R negative transitive.*

Proof. Let's go one by one:

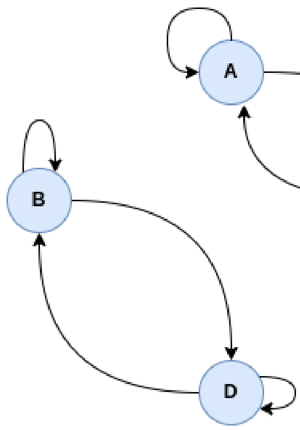
- By definition, we have $xR \bullet Ry \iff \exists z : xRz \land zRy$. Since R is transitive, then it must be xRy. Therefore $R \bullet R \subset R$.
- If R is not irreflexive, then there exists $x \in X$ such that xRx, but this is a symmetric relationship, so R cannot be assymetric.
- Trivial.
- By reduction ad absurdum, seeking a contradiction, let's assume that R is not transitive. This means that there exist $x, y, z \in X$ such that $xRy \land yRz$ but $not\,(xRz)$. By hypothesis, R is asymmetric, so $xRy \implies not\,(yRx)$. If we combine these two facts, and use the hypothesis that R is negative transitive, we find that
$$not\,(yRx) \land not\,(xRz) \overset{neg\ trans}{\implies} not\,(yRz)\,\#$$
This is a contradiction, because we assumed that yRz. Therefore, R must be transitive.
- By reduction ad absurdum, seeking a contradiction, let's assume that R is not negative transitive. This means that there exist $x, y, z \in X$ such that $not\,(xRy) \land not\,(yRz)$, but xRz. By hypothesis, R is complete, so $not\,(yRz) \implies zRy$. If we combine the two facts, xRz and zRy, then the transitivity of R gives us $xRy\#$, which is a contradiction with our assumption, $not\,(xRy)$. Therefore, R must be negative transitive.

$\square$

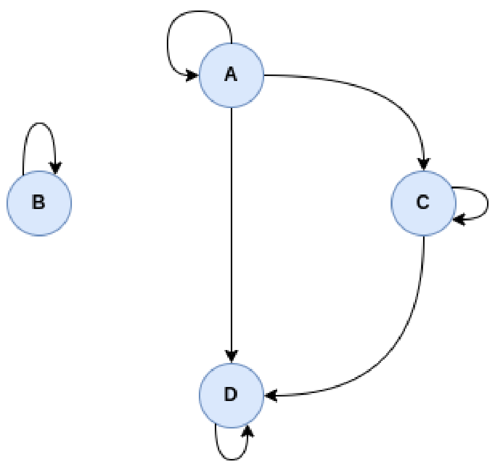
There are some relations that fulfill several of the properties that we have seen, and that hold special characteristics. For instance:

Definition 2.6. Different types of relations
An **equivalence relation** is a relation which is reflexive, symmetric and transitive.
A **preorder** is a relation which is reflexive and transitive.
A **weak order** or a **complete preorder** is a relation which is complete and transitive.
A **total order** or **linear order** is a relation which is complete, antisymmetric and transitive.

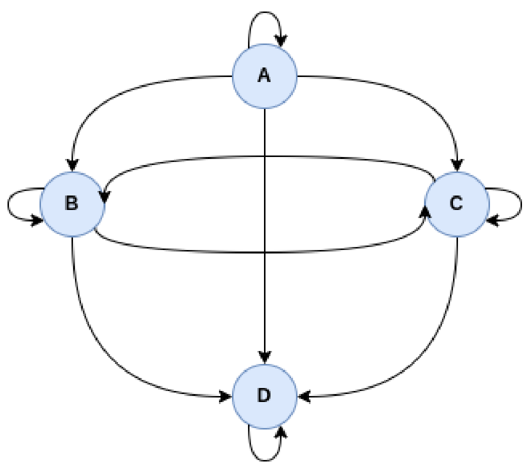
Example 2.4. The relation $R = \{(a,a), (a,c), (c,a), (c,c), (b,b), (b,d), (d,b), (d,d)\}$ is an equivalence relation. Its graph representation is:



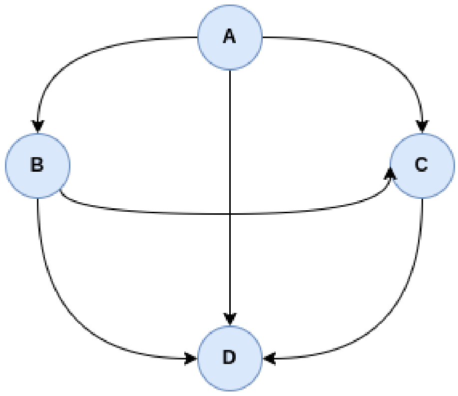
The relation $R = \{(a,a), (a,c), (a,d), (c,c), (c,d), (d,d), (b,b)\}$ is a preorder which is not a complete preorder. Its graph representation is:



The relation $R = \{(a,a), (a,c), (a,d), (c,c), (c,d), (d,d), (b,b), (a,b), (b,d), (b,c), (c,b)\}$ is a complete preorder. Its graph representation is:



The relation $R = \{(a,b), (a,c), (a,d), (b,c), (b,d), (c,d)\}$ is a total order. Its graph representation is:



Exercise 2.1. Let B be a binary relaiton on a set $X = \{a, b, c, d, e, f\}$ defined by

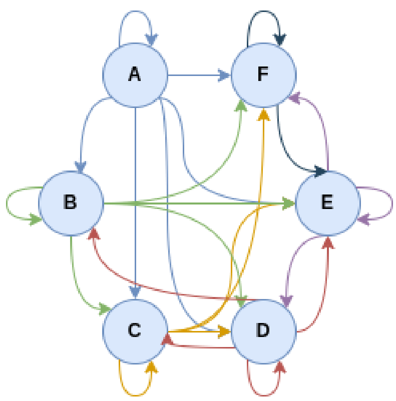
$$\begin{aligned} aBa, aBb, aBc, aBd, aBe, aBf \\ bBb, bBc, bBd, bBe, bBf \\ cBc, cBd, cBe, cBf \\ dBb, dBc, dBd, dBe \\ eBd, eBe, eBf \\ fBe, fBf \end{aligned}$$

Give a matrix and a graph representation of B

The matrix form of B is the following:

$$\begin{array}{c c} & \begin{array}{cccccc} a & b & c & d & e & f \end{array} \\ \begin{array}{c} a \\ b \\ c \\ d \\ e \\ f \end{array} & \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

And the graph representation:



Is B reflexive? Symmetric? Asymmetric? Transitive? Negative transitive? Semi-transitive?

- Reflexive: Yes, since $xBx, \forall x \in X$.
- Symmetric: No, aBb but *not* (bBa).
- Asymmetric: No, eBf but fBe.

- Transitive: No, dBe and eBf, but not dBf.
- Negative transitive: No, $not\,(fBd)$ and $not\,(dBf)$ but fBf.
- Semi-transitive: Yes. Let's do this little by litte.

  If we take $x = a$, then $aBt, \forall t$, so it will hold. If $y = a$, then $\nexists x$ such that xBa, so it will hold. Same happens if $z = a$. Finally, if $t = a$, then $aBz, \forall z \in X$, so it also holds.

  If we eliminate a (since we have seen all cases in which a is involved), we obtain the relation:

$$\begin{array}{c c} & \begin{array}{ccccc} b & c & d & e & f \end{array} \\ \begin{array}{c} b \\ c \\ d \\ e \\ f \end{array} & \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

  Using the same argument, b will work now. When we remove it, we get:

$$\begin{array}{c c} & \begin{array}{cccc} c & d & e & f \end{array} \\ \begin{array}{c} c \\ d \\ e \\ f \end{array} & \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

  And once again we can use c. When we remove it, we get:

$$\begin{array}{c c} & \begin{array}{ccc} d & e & f \end{array} \\ \begin{array}{c} d \\ e \\ f \end{array} & \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \end{array}$$

  Now we can repeat the same with e, and when we remove it we obtain a two-element relation, which is always semi-transitive.

Exercise 2.2. Let B and B' be two equivalence relations on a set X:

Prove that $B \cap B'$ is an equivalence relation, where

$$x\,(B \cap B')\,y \iff [xBy \wedge xB'y]\,, \forall x, y \in X.$$

We need to see that $B \cap B'$ is reflexive, symmetric and transitive:

- Reflexive: take $x \in X$, then $x\,(B \cap B')\,x \iff xBx \wedge xB'x$, which holds since B and B' are reflexive.
- Symmetric: take $x, y \in X$ such that $x\,(B \cap B')\,y$. This means that $xBy \wedge xB'y$. Since B and B' are symmetric, we obtain that $yBx \wedge yB'x$, and therefore $y\,(B \cap B')\,x$.
- Transitive: take $x, y, z \in X$ such that $x\,(B \cap B')\,y \wedge y\,(B \cap B')\,z$. This means that $xBy, xB'y, yBz$ and $yB'z$. Now, since B and B' are transitive, we obtain xBz and $xB'z$, so $x\,(B \cap B')\,z$.

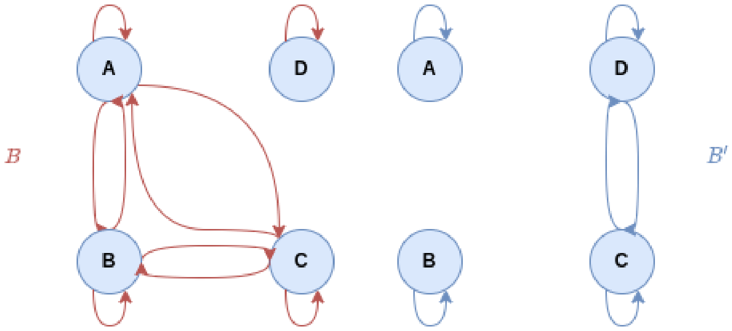
Is $B \cup B'$ an equivalence relation, where

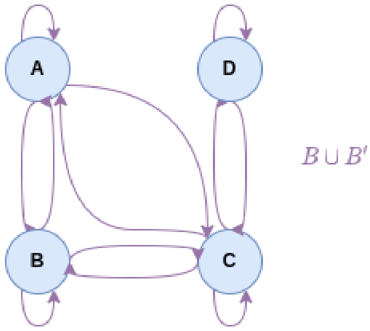
$$x\,(B \cup B')\,y \iff [xBy \vee xB'y]\,, \forall x, y \in X?$$

Reflexivity and symmetry are preserved, but what about transitivity? It is not... Take $X = \{a, b, c, d\}$ and the relations

$$B = \{(a,a), (a,b), (a,c), (b,b), (b,a), (b,c), (c,c), (c,b), (c,a), (d,d)\},$$
$$B' = \{(a,a), (b,b), (c,c), (c,d), (d,d), (d,c)\}.$$

Then, they are both equivalence relations, but their union is not. This is shown below:





As we can see, in the union, we find (a, c) and (c, d), but not (a, d), so it is not transitive.

Could we have the same conclusions if B and B' are two complete preorders on a set X?

We are been asked if $B \cap B'$ and $B \cup B'$ are also complete preorders. The answers to both questions is no.

In the case of the intersection, the transitivity is preserved, following the same argument we did for equivalence relations, but completeness is not preserved. To see this, take $X = \{a, b, c, d\}$ and the relations

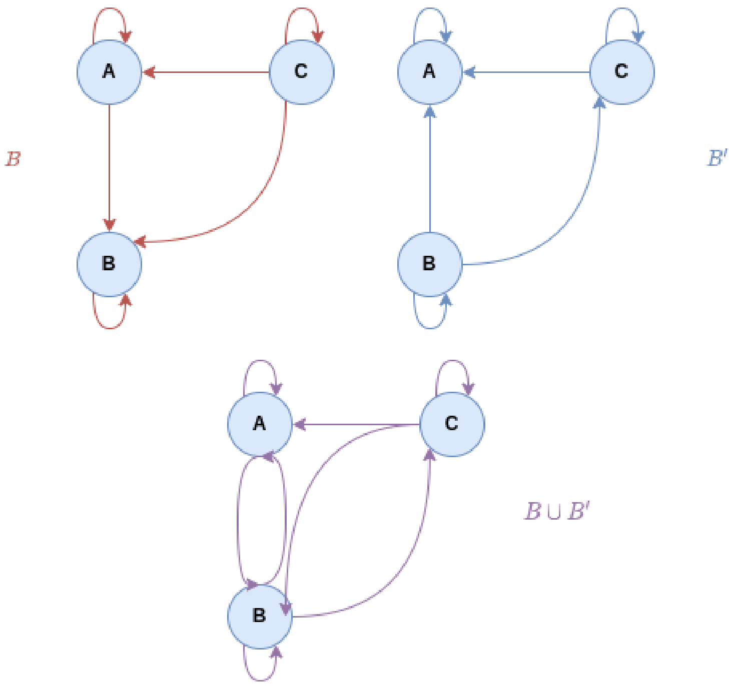
$$B : (a, b) \succ (c, d)$$
$$B' : (c, d) \succ (a, b) \,.$$

Then, the intersection is

$$B \cap B' : a \equiv b, c \equiv d,$$

which is not complete.

As for the union, the opposite happens. Completeness is preserved, because the relations involved are complete. However, transitivity is not preserved. As a counter example, take the following graph visualization:



As we can see, in the union, we find (a, b) and (b, c) but not (a, c).

Definition 2.7. The **reflexive closure** of a binary relation, R, is

$$r(R) = R \cup Eq,$$

where Eq is the reflexive, relation, i.e., $Eq = \{(x, x) : x \in X\}$.

Therefore, the reflexive closure of a relation is the relation resulting when adding all needed edges for the initial relation to be reflexive. As a example, see Figure 1.

A similar concept is that of symmetric closure:

Definition 2.8. The **symmetric closure** of a relation R is

$$s(R) = R \cup R^c,$$

where R^c is the converse relation, i.e.,

$$xR^c y \iff yRx.$$

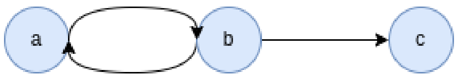
In this case, we are adding all needed edges for the relation to be symmetric. Again, see Figure 1 for an example.

Finally, there is the transitive closure:

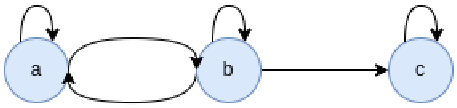
Definition 2.9. The **transitive closure** of a relation R is

$$t(R) = R \cup R^2 \cup R^3 \cdots = \bigcup_{i \geq 1} R^i.$$

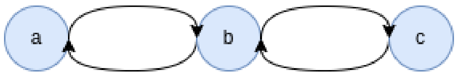
In this case, we add all possible transitive relations by relating all elements with a path between them. Note that if the set X has size N, then we only need to use $N - 1$ step at maximum to reach any reachable node. An example is also shown in Figure 1.



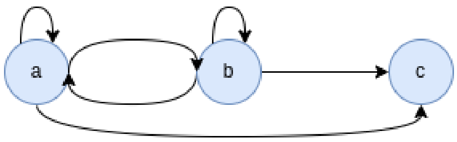
(a) Original relation, R.



(b) Reflexive closure, $r(R)$.



(c) Symmetric closure, $s(R)$.



(d) Transitive closure, $s(R)$.

Figure 1: Closures visualization.

A natural question at this point is if we can get any of our properties by applying some kind of closure, and the answer in the general case is no. However, there are ways to obtain reasonable solutions.

For instance, consider the problem of extending a partial pre-order to a complete pre-order. In such case, we need to define how unrelated elements should be related. One way to solve this is by applying what is called a **topological sorting**, which creates a complete pre-order that respects all the partial pre-order preferences, when there are no cycles. More formally:

Definition 2.10. A **Directed Acyclic Graph (DAG)** is a directed graph with no cycles, i.e., there is no $x \in X$ such that there is a path from x to itself.

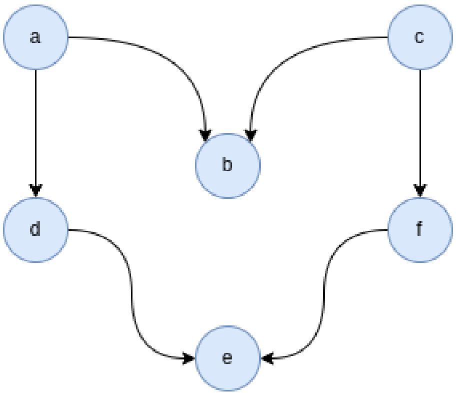
Definition 2.11. A **topological sorting** for a DAG is a linear ordering of vertices such that, for every directed edge, (u, v), vertex u comes before v in the ordering.

There are several ways to perform topological sorting, obtaining different results, and all of them respecting the initial preferences. As a general algorithm, see Algorithm 1. Note that we are first taking those nodes with no outgoing edges, but we could do it also by taking the nodes with no ingoing edges, in which case the list would not need to be reversed at the end.

```
input:
  set X
  relation R (must be DAG)

initialize:
  list L

execute:
  while R is not empty:
    free = {x in X : out_degree(x,R)=0 } # Take all nodes without outgoing edges
    R = R - edges(free) # Remove edges going into nodes in free from R
    list.append(free) # Append each node to the list

  list.append(rest of nodes)

  return list.reverse # The order is reversed
```

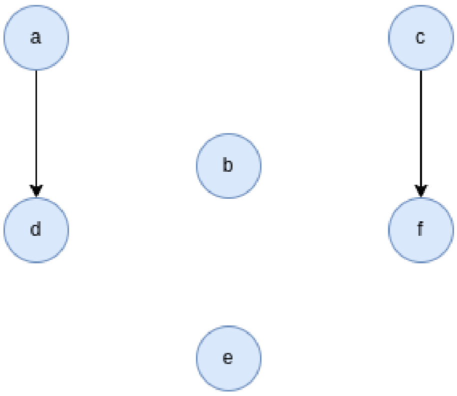
Algorithm 1: Topological sorting.

Example 2.5. Create a topological sorting of the following relation:

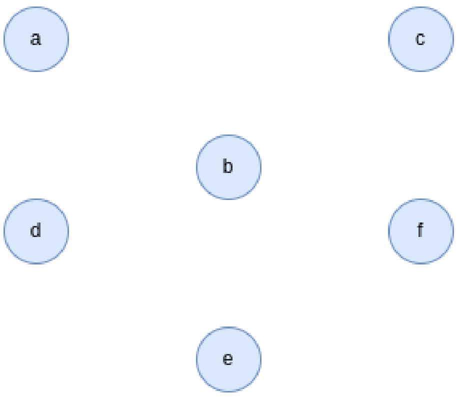


Let's follow our pseudocode:

- Input: $X = \{a, b, c, d, e, f\}$, $R = \{aRb, aRd, cRb, cRf, dRe, fRe\}$
- $L = []$
- $free = \{e, b\}$
- $R = R - edges\,(\{e, b\}) = \{aRd, cRf\}$



- $L = [e, b]$
- $free = \{d, f\}$
- $R = \{\}$



- $L = [e, b, d, f]$
- Append the rest of nodes: $L = [e, b, d, f, a, c]$.

Therefore, a linear ordering respecting the preferences in the original partial preorder is $c \succsim a \succsim f \succsim d \succsim b \succsim e$.

2.2 Numerical representation

The idea of the numerical representation is to try to construct a binary relation $\succsim$ on a set X such that there exists a numerical function $f : X \to \mathbb{R}$ satisfying the property

$$x \succsim y \iff f(x) \geq f(y).$$

Therefore, we are trying to convey the information given by the relation via a mathematical function. When doing this, we usually assume $\succeq$ to be a preorder.

We also usually use the following notation:

- $x \succsim y$ means *x is at least as good as y.*
- $\succ$ is the asymmetric part of $\succsim$.
- $\sim$ is the symmetric part of $\succsim$.

Cantor proved a theorem that answered this problem:

Theorem 2.1. *Cantor, 1895*
Let X be a countable set and $\succsim$ a binary relation on X. Then:

$$\begin{gathered}[\exists f : X \to \mathbb{R} | \forall x, y \in X, x \succsim y \iff f(x) \geq f(y)] \\ \Updownarrow \\ \succsim \textit{ is a complete preorder on } X.\end{gathered}$$

Proof. $[\implies]$

- Completeness: pick $x, y \in X$, then we compute $f(x)$ and $f(y)$. Since $\mathbb{R}$ is a linearly ordered set, it is either $f(x) \geq f(y)$ or $f(y) \geq f(x)$. This, by hypothesis means that $x \succsim y$ or $y \succsim x$.
- Transitivity: pick $x, y, z \in X$ with $x \succsim y$ and $y \succsim z$. Then, $f(x) \geq f(y)$ and $f(y) \geq f(z)$. The transitivity in the real line gives us $f(x) \geq f(z)$, which by hypothesis means $x \succsim z$.

Therefore, $\succsim$ is a complete preorder on X.

$[\impliedby]$ We assume that $\succsim$ is a complete preorder on X. Since X is countable, we can enumerate its elements, i.e., we can express

$$X = \{x_i : i \in K \subset \mathbb{N}\} = \{x_1, x_2, ...\}.$$

Let $N(y) = \{i \in K : y \succsim x_i\}$, i.e., the set of indices of elements put after y.

Now, we define

$$\begin{array}{rccc} f: & X & \to & \mathbb{R} \\ & y & \mapsto & f(y) = \sum_{i \in N(y)} \frac{1}{2^i} \end{array}.$$

This series is convergent, since $\sum_{n \in \mathbb{N}} \frac{1}{2^n} = 2$ and we are taking a subset of it. Let's see that $x \succsim y \iff f(x) \geq f(y)$:

- If $x \succsim y$, then, the transitivity of the relation ensures that

$$N(x) \supseteq N(y),$$

therefore, it follows that

$$f(x) \geq f(y),$$

simply because we are summing over more elements.

- Conversely, assume that $f(x) \geq f(y)$ but not $x \succsim y$. Since $\succsim$ is complete, then it must be $y \succsim x$, which combined with *not* $(x \succsim y)$ gives us that it must be $N(y) \supsetneq N(x)$. This means that there are elements in $N(y)$ which are not in $N(x)$, so $f(y) > f(x)\#$ which is a contradiction. Therefore, it must be $x \succsim y$.

□

Note that this theorem ensures that such a function can only be found if the relation is a complete preorder, and therefore all partial preorders that are not complete cannot be assigned such a function. Note also that the defined function is not unique (for instance, add 1 to this function, and it still works).

In addition, the following proposition proves that even if there are infinite possible such functions, they are all related:

Proposition 2.2. *Let $\succsim$ be a complete preorder on X, representable by a function $f: X \to \mathbb{R}$. Then, the following two properties are equivalent:*

1. *$v: X \to \mathbb{R}$ is a function representing $\succsim$.*
2. *There exists a strictly increasing function $\varphi: f(X) \to \mathbb{R}$ such that $v = \varphi \circ f$.*

Proof. [2 $\implies$ 1] $x \succsim y \overset{f \text{ represents } \succsim}{\iff} f(x) \geq f(y) \overset{\varphi \text{ strictly increasing}}{\iff} \varphi(f(x)) \geq \varphi(f(y)) \overset{\text{definition of } g}{\iff} g(x) \geq g(y)$.

[1 $\implies$ 2] Define $\varphi: f(X) \to \mathbb{R}$ as

$$\varphi(u) = v(x) : f(x) = u.$$

We have to see that φ is well-defined, strictly increasing and that, indeed, $v = \varphi \circ f$.

- Well-defined: For each $u \in f(X)$, there exist $x \in X$ such that $f(x) = u$, so $\{v(x) : f(x) = u\}$ is non-empty and therefore $\varphi(u)$ is defined.
- Strictly increasing: Take $u_1, u_2 \in f(X)$ with $u_1 > u_2$. Then, there exist $x_1, x_2 \in X$ with $f(x_1) = u_1$ and $f(x_2) = u_2$. This means that $x_1 \succsim x_2$ and not $x_2 \succsim x_1$, since $f(x_1) > f(x_2)$. But then, it is $v(x_1) > v(x_2)$, and so $\varphi(u_1) > \varphi(u_2)$.
- The equality: Take $x \in X$, then

$$\varphi(f(x)) = v(x) \,|f(x) = f(x) = v(x).$$

□

Remark 2.1. This representation functions are **ordinal scales**, since the importance does not rely on the absolute values they provide, but in comparating the values for different elements.

3 Voting Rules as Group Decision Making Models

The objectives of this section are to study a set of decision problems in which a group has to take a decision among several alternatives and analyzing a number of properties of electoral systems. For this, we will present a few elements of the classical **social choice theory**, for which we will see that several parameters need to be taken into account. These are the nature of the decision, the size of the group and its nature.

> **Problem**
> Study election problems in which a society has to take a decision among several candidates.

When electing a candidate, we assume that the choice of the candidate will affect the member of society, and should therefore take the opinion of all members into account.

The intuition then is that a democratic approach is preferred, through an election system in which a candidate should be elected through some kind of majority rule.

Of course, this approach has many political considerations, like if the democracy is direct or indirect, how parties play a role, who can vote, etc.

But there are also technical considerations involved in this problem. For instance, the definition of majority is obvious when there are two candidates: the candidate receiving more than half the votes is preferred to the other one. However, when there are more than two candidates, there are many ways to extend this simple idea, which are not equivalent and can lead to different and even unwanted results.

In addition, elections can by of different types:

- Type of ballots admitted: only one candidate; or a ranking of all candidates; or all acceptable candidates; or grading candidates; etc.
- Method for organizing the election and tallying ballots.

The consequences of all this is that there are many possible types of elections, many of which have been proposed and even used in practice.

> **Hypotheses**
> 1. All voters are able to rank the set of all candidates, admitting for ties.
> 2. Voters are sincere: this means that they don't change their preferences strategically.

3.1 Methods of Voting

3.1.1 Plurality Voting

> **Plurality: rules**
> 1. One round of voting.
> 2. Ballots with one name.
> 3. The most voted candidate wins.

Remark 3.1. Ties are neglected because they are unlikely for large populations. They can be solved in several ways, for example if one voter has special power (the king or queen); or one candidate receives special treatment (the older one is elected in case of tie); or even choosing randomly.

Example 3.1. Consider 3 candidates $\{a, b, c\}$ and 21 voters, with the following distribution:

$$10\ voters : a \succ b \succ c$$
$$6\ voters : b \succ c \succ a$$
$$5\ voters : c \succ b \succ a$$

In this case, candidate a is elected, according to plurality voting, since it's the most voted one.

Note, however, how a majority of voters (11/21) prefer any of the losing candidates to the elected one.

Remark 3.2. Problems are expected as soon as there are more than 2 candidates. A system based on an idea of majority, like this, may violate the will of a majority of voters. Also, the sincere hypothesis becomes hard to believe. For example, c's voters could change their vote for b, making it win the election and improving the outcome according to their preference.

3.1.2 Plurality with Runoff

Plurality with Runoff: rules

1. Ballots with one name.
2. Two rounds of voting:
   (a) First round:
      i. The candidate with most votes if elected if he receives more than 50% of votes.
      ii. Otherwise, go to the second round.
   (b) Second round:
      i. Keep the two most voted candidates.
      ii. Apply plurality votes to these two.

Example 3.2. Consider 3 candidates $\{a, b, c\}$ and 21 voters, with the following distribution:

$$10\ voters : a \succ b \succ c$$
$$6\ voters : b \succ c \succ a$$
$$5\ voters : c \succ b \succ a$$

First round: the most voted candidate is a with 10 votes, less than 50%. Therefore, we go to the second round.

Second round: we only consider the two most voted candidates, a and b:

$$10\ voters : a \succ b$$
$$6\ voters : b \succ a$$
$$5\ voters : b \succ a$$

Therefore, in this case b wins with 11 votes.

Example 3.3. Consider 4 candidates $\{a, b, c, d\}$ and 21 voters, with the following distribution:

$$10\ voters : b \succ a \succ c \succ d$$
$$6\ voters : c \succ a \succ d \succ b$$
$$5\ voters : a \succ d \succ b \succ c$$

In this case, with plurality with runoff. the first round is won by b, with 10 votes, which is not enough, and then the second round is between b and c. b wins with 15 votes in the second round.

However, note that there is a majority of voters (11/21) preferring a and d to b.

Some takeaways:

- The French system does only a little better than the UK one.
- Preferences used in the previous example are not extremely rare.

Manipulation Manipulation occurs when some voters are not sincere, to improve the result of the election according to their preferences. Let's see this through an example.

Example 3.4. In the previous example, b was elected. Howeber, suppose that the 6 voters with preferences $c \succ a \succ d \succ b$ vote as if their preferences were $a \succ c \succ d \succ b$. Then, a would be elected in the first round with 11 votes. Therefore, it is profitable to the six manipulating voters, for which $a \succ b$.

> **Definition 3.1.** A voting rule is **manipulable** if it may happen that some voters may have an interest to vote in a non-sincere way.

We have just seen with the previous example that plurality with runoff is manipulable.

Monotonicity Monotonicity refers to a property of a voting method. It means that if a candidate increments its valoration for some population, the outcome of the election should not worsen for this candidate.

Example 3.5. Consider 3 candidates $\{a, b, c\}$ and 17 voters, with the following distribution:

$$\begin{aligned} 6\ voters :& a \succ b \succ c \\ 5\ voters :& c \succ a \succ b \\ 4\ voters :& b \succ c \succ a \\ 2\ voters :& b \succ a \succ c \end{aligned}$$

In this case, a would be elected in the second round. However, imagine that a now gets more money and can increase their campaign power. Through this campaign, they manage to change the preferences of the two last voters, effectively changing the distribution to:

$$\begin{aligned} 8\ voters :& a \succ b \succ c \\ 5\ voters :& c \succ a \succ b \\ 4\ voters :& b \succ c \succ a \end{aligned}$$

Now, c is elected! The good campaign done by a, made him lose.

With this example, we observe that when the voting method is non-monotonic, the possibilities of manipulation increase.

3.1.3 Condorcet Voting Rule

> **Condorcet: rules**
>
> 1. Compare all candidates by pair.
> 2. Declare that a is socially preferred to b if strictly more voters prefer a to b.
> 3. **Condorcet principle**: if one candidate is preferred to all other candidates, this one should be elected. This is the **Condorcet winner**.

Remark 3.3. The condorcet winner, if there is one, must be unique.

Plurality rule and plurality with runoff violate Condorcet's principle.

Condorcet's principle does not solve the 'dictature of the majority' difficulty. This is, a majority of the population only take their objectives into account. Since they are more, they can decide the elections.

A Condorcet winner is not necessarily ranked high by voters.

Example 3.6. Consider 3 candidates $\{a, b, c\}$ and 21 voters, with the following distribution:

$$\begin{aligned} 10\ voters :& a \succ b \succ c \\ 6\ voters :& b \succ c \succ a \\ 5\ voters :& c \succ b \succ a \end{aligned}$$

The results according to Condorcet are:

- a is not preferred to b nor c
- b is preferred to a and c
- c is preferred to a but not to b

Therefore, b is the Condorcet winner.

Example 3.7. Consider 4 candidates $\{a, b, c, d\}$ and 21 voters, with the following distribution:

$$\begin{aligned} 10\ voters :& b \succ a \succ c \succ d \\ 6\ voters :& c \succ a \succ d \succ b \\ 5\ voters :& a \succ d \succ b \succ c \end{aligned}$$

In this case, a is the Condorcet winner.

Condorcet's Paradox The social strict preference relation may have circuits:

Example 3.8. Consider 3 candidates $\{a, b, c\}$ and 3 voters, with the following distribution:

$$\begin{aligned} 1\ voter :& a \succ b \succ c \\ 1\ voter :& b \succ c \succ a \\ 1\ voter :& c \succ a \succ b \end{aligned}$$

In this case, a is preferred to b, b is preferred to c and c is preferred to a.

Therefore, electing the Condorcet winner is an attractive idea, but not always possible or effective.

4 Preferences Aggregation: the Multri-Attribute Utility Theory (MAUT) Approach

4.1 Multi-Criteria Decision Analysis (MCDA)

Think about a situation in which a Decision Maker (DM) is facing a decision problem, i.e., the DM has to deal with different alternatives and compare them, to take a decision. The alternatives are described by several attributes.

A **criterion** is an attribute associated to a preference relation, also called a monotonic attribute.

Criteria cannot be reduced to only one criterion, since they can potentially be in conflict.

More formally, let $X = \{x_1, ..., x_m\}$ be the set of alternatives, that are evaluated on a finite set of criteria, $N = \{1, ..., n\}$. X can be also written as $X = X_1 \times X_2 \times ... \times X_n$, where $X_i \in \mathbb{R}^m$ is the vector of all alternatives evaluated with criterion i.

There exist preferences on the values for each criterion i, by means of an utility funtion, a qualitative preference relation $\succsim_i$, etc.; and a representation of the importance of each criterion or set of criteria, by means of weights, importance relations, etc.

Now, using the input information, we want to elaborate a decision rule allowing to compare two different alternatives. That is, given two alternatives, $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$, we want to be able to say whether $x \succsim y$ or $y \succsim x$.

Example 4.1. Consider the following students grades:

	Math	Stats	Lang
a	16	13	7
b	16	11	9
c	6	13	7
d	6	11	9

How we can rank these students?

There are many ways to do this, for example, we can assign weights to each course and use the weighted sum to make the ranking, for example, with $w_M = 1, w_S = 0, w_L = 1$, the ranking would be: $b \succ a \succ d \succ c$.

Another way is to use a rule system, such as: Compare Math -> If tied, compare Stats -> If tied, compare Lang. In this case, the ranking would be $a \succ b \succ c \succ d$.

MCDA presents several difficulties:

- It is not easy.
- There is no best method for all situations.
- All methods have some structural bias.

There are three main problems treated with MCDA:

- **Problem of Choice**: choose the best out of a set of alternatives.
- **Problem of Ranking**: rank a set of alternatives from best to worst.
- **Problem of Sorting**: sort the alternatives into a set of pre-defined categories, which are usually ordered.

4.2 Some Simple Models

4.2.1 Pareto Dominance

Definition 4.1. Given two alternatives, x and y, evaluated on a set of criteria, $N = \{1, ..., n\}$, we say that x (pareto) dominates y, $x \succsim y$, if it is considered better on all the criteria:

$$x \succsim y \iff [\forall i \in N, x_i \succsim_i y_i] .$$

We say that x (pareto) strictly dominates y, $x \succ y$, if it is considered strictly better on all the criteria:

$$x \succ y \iff [\forall i \in N, x_i \succ y_i] .$$

The **pareto front** is the set of all non-dominated alternatives, i.e., those that are not dominated by any other alternative.

Remark 4.1. The optimal solution is necessarily in the pareto front.

This model is not very interesting, because most alternatives will not be pareto comparable, i.e., one will be better in some criteria, and the other will be better on different criteria. This means that the pareto front would usually be very similar to the whole set of alternatives.

4.2.2 Weighted Sum

Definition 4.2. Given two alternatives, x and y, evaluated on a set of real-valued criteria, $N = \{1, ..., n\}$, and a set of weights $W = \{w_1, ..., w_n\}$, we define the following preference relationship:

$$w \succsim y \iff \sum_{i=1}^{n} w_i x_i \geq \sum_{i=1}^{n} w_i y_i.$$

4.2.3 Majority Rule

Definition 4.3. Given two alternatives, x and y, evaluated on a set of criteria, $N = \{1, ..., n\}$, we say that x is preferred to y if it is considered better in a majority of the criteria:

$$x \succsim y \iff |\{i \in N : x_i \succsim y_i\}| \geq |\{i \in N : y_i \succsim x_i\}|.$$

Example 4.2. Consider the following example:

	Speed	Robustness	Price
A	20	Very Good	600
B	15	Good	500
C	25	Bad	550

We find that $A \succsim B, B \succsim C$ and $C \succsim A$, so we cannot make a decision. This is called the **Condorcet Paradox**.

All the previous approaches can be grouped into two main kinds:

1. **Multi Attribute Utility Theory**: a quantitative approach based on 'aggregate, then compare':

$$x \succsim y \iff U(x) \geq U(y),$$

   where U is an utility function.

2. **Outranking**: a qualitative approach, based on 'compare, then aggregate':

$$x \succsim y \iff |\{i \in N : x_i \succsim_i y_i\}| \triangleright |\{i \in N : y_i \succsim_i x_i\}|.$$

We are going to focus on MAUT approach.

4.3 Multi Attribute Utility Theory (MAUT)

Let X be a set of alternatives evaluated on a finite set of criteria, $N = \{1, ..., n\}$, as before, $X = X_1 \times X_2 \times ... \times X_n$.

Let $\succsim_X$ be a complete preorder on X, indicating the preferences of the Decision Maker (DM). We assume that $\succsim_X$ is representable by an overal utility function. That is,

$$\exists F, U : \forall x, y \in X, x \succsim y \iff F(U(x)) \geq F(U(y)),$$

where x, y are alternatives in X, $U(x) = (u_1(x_1), ..., u_n(x_n))$ with $u_i : X_i \to \mathbb{R}$ is a marginal utility function and $F : \mathbb{R}^n \to \mathbb{R}$ is an aggregation function.

F is usually characterized by a parameter vector, θ. This is, for example, a weight vector in the case of the weighted sum.

There are several considerations to take into account:

- How to choose the aggregation function F?
- How to construct the marginal utility functions $u_i : X_i \to \mathbb{R}$?
- The marginal utility functions u_i should have a signification for the DM:
  - Ordinal scales: absolute differences between values have no importance. They can represents order and preorders.
  - Cardinal scales: differences between values may be meaningful.

4.3.1 The Additive Model

In the additive model, the overall utility function can be represented as

$$F(U(x)) = \sum_{i=1}^{n} u_i(x_i),$$

or, equivalently,

$$F(U(x)) = \sum_{i=1}^{n} w_i u_i(x_i).$$

This is a simple method that involves a compensation between criteria, i.e., a bad performance on a criterion i could be compensated by a good performance on another criterion.

In the weighted sum, the weights represents the substitution rate between criteria. For example, consider $n = 2$ and $w_1 = b \cdot w_2$. In this case, the alternatives $(0, b)$ and $(1, 0)$ are indifferent.

If we have $(a, b) \sim (a - \delta, b + \gamma)$, then the gain of γ compensates the loss of δ. Indeed, we have

$$\begin{aligned} w_1 \cdot a + w_2 \cdot b &= w_1 \cdot (a - \delta) + w_2 \cdot (b + \gamma) \\ \iff w_1 \delta &= w_2 \gamma \\ \iff \frac{w_1}{w_2} &= \frac{\gamma}{\delta}. \end{aligned}$$

Implicitly, this implies that all the criteria could be expressed indirectly in the same unit.

Note that the additive model requires to normalize the criteria. In general, we set, $\forall i \in N$, $u_i : X_i \to [0, 1]$.

For example, we can use:

$$u_i(x_i, X) = \frac{x_i}{\max_{y \in X} y_i},$$

or

$$u_i(x_i, X) = \frac{x_i - \min_{y \in X} y_i}{\max_{y \in X} y_i - \min_{y \in X} y_i}.$$

The Mutual Preferencial Independence Axiom
The additive model requires to satisfy the Mutual Preferencial Independence Axiom, i.e., the criteria are independent in the sense of preferences. This is formalized as

$$\forall i \in N, \forall z_i, t_i \in X_i, \forall x, y \in X, (z_i, x_{N-1}) \succsim (z_i, y_{N-1}) \iff (t_i, x_{N-1}) \succsim (t_i, y_{N-1}).$$

This means that an attribute is preferentially independent from all the other attributes if changes in the rank ordering of preferences of other attributes do not change the preference order of the attribute.

4.3.2 MAUT in Practice

In practice, we ask the DM some preferencial information $\succsim_{X'}$ on a **reference set** (or learning set), $X' \subset X$. Then, the parameter vector that represents the overall utility function is constructed so that $\succsim_X$ is an extension of $\succsim_{X'}$, so that the modelobtained in X' is automatically extended to X.

4.3.3 The UTilités Additives (UTA) Approach

The UTA method aims at inferring one or more additive value functions from a given ranking on a reference set A_R.

The method uses linear programming techniques to assess these functions so that the reankings obtained through these functions on A_R are consistent with the given one.

The input data of UTA is a set of criteria N, on which a set of alternatives, X, is evaluated; a preorder $\succsim_{X'}$ on $X' \subset X$.

For each element $x \in X$, it is assumed that

$$U(x) = \sum_{i=1}^{n} u_i(x_i),$$

where $u_i : X_i \to \mathbb{R}$ are marginal utility functions.

Then, for each element $x \in X'$, we set

$$V(x) = U(x) + \sigma(x),$$

where $\sigma(x)$ is a non-negative real valued function estimating the error of the estimation of the value $U(x)$.

The value $\sigma(x)$ will be minimized by the linear program. That is:

$$\begin{array}{ll}
\min \sum_{x \in X'} \sigma(x) & \\
s.t. & \\
\quad V(x) \geq V(y) + \delta & if\ x \succ y \\
\quad V(x) \sim V(y) & if\ x \sim y \\
\quad u_i\left(x_i^{j+1}\right) - u_i\left(x_i^{j}\right) \geq 0 & if\ x_i^{j+1} \succ_i x_i^{j} \\
\quad \sigma(x) \geq 0 & \forall x \in X'
\end{array}$$

If the optimal solution is equal to 0, then $\succsim_{X'}$ is representable by an additive model.

Note that there are many variations of this method. For example, the $\boldsymbol{UTA^{GMS}}$ **Approach**, which generalizes the UTA approach taking into account all additive value functions compatible with indirect preference information, while UTA is using only one such function. The marginal value functions are general monotonic non-decreasing functions, and not just piecewise linear. This method produces two rankings in the set of alternatives X, such that, for any pair of alternatives $x, y \in X$:

- In the **necessary order**, $x \succsim_X y \iff U(x) \geq U(y)$, for all value functions U compatible with $\succsim_{X'}$.
- In the **possible order**, $x \succsim_X y \iff U(x) \geq U(y)$, for at least one value function U compatible with $\succsim_{X'}$.

4.4 Final Thoughts

The Multi-Criteria Decision Aid (MCDA) phenomena has three branches:

- Prescriptive approach: help the DM by proposing a solution using a model.
- Descriptive approach: describe the DM preferences using a model.
- Elicitation: obtain the parameters of a decision model that explains the past decisions, in order to help with future decisions. This can be done in two main ways:
  - Explicit elicitation: explain the model to the DM and let the DM choose the parameters of the model.
  - Implicit eliciation: present some alternatives to the DM and ask him to compare them. From these, deduct the parameters of the model by solving an optimization problem.

5 The Outranking Approach

Now, we are going to study the outranking approach, based on the premise 'compare, then aggregate':

$$x \succsim y \iff \left|\{i \in N : x_i \succsim_i y_i\}\right| \triangleright \left|\{i \in N : y_i \succsim_i x_i\}\right|.$$

The principle of this approach is to compare alternatives criterion by criterion, and then aggregated the comparisons:

$$x \succsim y \iff U\left(c\left(x_1, x_2\right), ..., c\left(x_n, y_n\right)\right) \geq 0.$$

The objectives of the approach is to build a relation on the alternatives, called the **outranking relation**, and exploit it to solve on the MCDA problems. One of the particularities of outranking methods is that the relation built on the set X allows three types of comparisons of alternatives, namely: *preference, indifference* and *incomparability*.

This approach allows to represents doubts or hesitations of the DM, which may result from phenomena like uncertainty, conflicts or contradictions.

5.1 Elaboration of the Outranking Relation

Let X be a set of alternatives evaluated on n real-valued criteria, $N = \{1, ..., n\}$. Let $W = (w_1, ..., w_n)$ be a non-negative weight vector assigned to the different criteria. Without loss of generality, we can suppose that $\sum_{i=1}^n w_i = 1$. Let $p_i \geq 0$ for $i = 1, ..., n$ be a non-negative preference threshold assigned to each criterion i. If the value $0 < x_i - y_i < p_i$, we assume that this difference is not significant. Hence, on this criterion, the two alternatives should be considered indifferent.

Definition 5.1. The **partial concordance index** is a map $c_i : X \times X \to [0,1]$ defined on each criterion $i \in N$ by

$$c_i\left(x, y\right) = \begin{cases} 1 & if\ x_i - y_i > p_i \\ 0 & if\ x_i - y_i \leq p_i \end{cases}.$$

The **concordance index** is an aggregation of partial concordance indices c_i:

$$c\left(x, y\right) = \sum_{i=1}^{n} w_i \cdot c_i\left(x, y\right).$$

Definition 5.2. The **outranking relation**, S_λ, given the **cutting level** $\lambda \in [0,1]$, on X, is a binary relation defined by

$$xS_\lambda y \iff c\left(x, y\right) \geq \lambda.$$

Remark 5.1. The concordance indicates whether there are enough reasons to say that x is at least as good as y. An alternative $x \in X$ outranks an alternative $y \in X$ if it can be considered ***'at least as good'*** as the latter, given the values of x and y for the n criteria. However, if there are some criteria in which x is worse than y, then x may outrank y or not, depending on the relative importance of those criteria and the differences in the evaluations, since p_i may allow for ignoring small differences.

From S_λ, we can derive three binary relations:

- **Strictly better than** relation:

$$xP_\lambda y \iff [xS_\lambda y \wedge \mathit{not}\,(yS_\lambda x)]\,.$$

- **Indifferent to** relation:

$$xI_\lambda y \iff [xS_\lambda y \wedge yS_\lambda x]\,.$$

- **Incomparable** relation:

$$xU_\lambda y \iff [\mathit{not}\,(xS_\lambda y) \wedge \mathit{not}\,(yS_\lambda x)]\,.$$

Remark 5.2. Notice that an outranking relation is not necessarily transitive, and that it can be defined using qualitative data as well.

Moreover, we could also use a **discordance threshold**, which indicates whether there are not enough important reasons to say that x is worse than y.

5.2 ELECTRE TRI Method

Let us consider r ordered categories, $C_1, ..., C_r$, with C_1 being the worst one, and C_r the best one. Each category, C_k, is modeled by using limiting profiles, i.e., $C_k = [\pi_k, \pi_{k+1})$, where $\pi_{k+1} \succ \pi_k$ and it holds:

$$\forall x \in X, xP_\lambda \pi_1, \pi_{r+1}P_\lambda x.$$

ELECTRE TRI is MCDA method that uses limiting profiles to classify the elements in X into the r classes.

Definition 5.3. Pessimistic version of ELECTRE TRI
For each $x \in X$:

1. Decrease k, from $r+1$, until the first k such that $xS_\lambda \pi_k$:

$$k = \arg\max_k \{k : xS_\lambda \pi_k\}\,.$$

2. Assign alternative x to C_k.

The pessimistic version of ELECTRE TRI assigns an alternative x to the unique class C_k such that x is at least as good as to its lower limiting profile, and not at least as good as its upper limiting profile.

Definition 5.4. Optimistic version of ELECTRE TRI
For each $x \in X$:

1. Increase k, from 1, until the first k such that $\pi_k P_\lambda a$:

$$k = \arg\min_k \{k : \pi_k P_\lambda x\}\,.$$

2. Assign alternative x to C_{k-1}.

The optimistic version of ELECTRE TRI assigns an alternative x to the unique class C_k such that its upper limiting profile is better than x and its lower limiting profile is not better than x.

Theorem 5.1. *If $x \in X$ is assigned to the class C_k by the pessimistic version, and to the category C_l by the optimistic version, then $k \leq l$.*

5.3 Majority Rule Sorting (MR-Sort)

MR-Sort is a simplified version of the ELECTRE TRI method. The general principle of MR-Sort is to assign alternatives by comparing their performances to those to profiles delimiting the categories. An alternative is assigned to a category above a profile if, and only if, it's at least as good as the profile on a weighted majority of criteria.

That it, $x \in X$ is assigned to category C_k if

$$\sum_{i:a_i \geq \pi_{k,i}} w_i \geq \lambda$$

and

$$\sum_{i:a_i \geq \pi_{k+1,i}} w_i < \lambda.$$

The MR-Sort method involves $r \times n + 1$ parameters.

The problem of learning the parameters of MR-Sort model on the basis of assignment examples can be formulated as a MIP problem, but only instances of modest size can be solved in reasonable computing times.

For instance, a problem involving 1000 alternatives, 10 criteria and 5 categories requires 21000 binary variables.

Learning only the weights and the majority threshold of an MR-Sort model on the basis of assignment examples can be done using an ordinary linear program, without binary nor integer variables. On the contrary, learning profile evaluations is not possible by linear programming without binary variables.

Example 5.1. Given the following alternatives, criteria and weights:

	1	2	3	4
a	5	7	2	5
b	8	4	6	2
c	4	8	7	5
d	6	4	5	7
e	2	6	2	8
f	3	5	6	4
Weights	0.4	0.3	0.1	0.2

Assign them to the categories $C_1 \equiv Bad$, $C_2 \equiv Medium$ and $C_3 \equiv Good$, with limiting profiles:

	1	2	3	4
π_4	10	10	10	10
π_3	6	6	5	5
π_2	5	5	3	4
π_1	1	1	1	1

With $\lambda = 0.7$

- a:

$$a - \pi_3 : 0 + 0.3 + 0 + 0.2 = 0.5 < 0.7$$

$$a - \pi_2 : 0.4 + 0.3 + 0 + 0.2 = 0.9 \geq 0.7$$

  So $a \in C_2$.

- b:

$$b - \pi_3 : 0.4 + 0 + 0.1 + 0 = 0.5 < 0.7$$

$$a - \pi_2 : 0.4 + 0 + 0.1 + 0 = 0.5 < 0.7$$

  So $b \in C_1$.

- c:
$$c - \pi_3 : 0 + 0.3 + 0.1 + 0.2 = 0.6 < 0.7$$
$$c - \pi_2 : 0 + 0.3 + 0.1 + 0.2 = 0.6 < 0.7$$
So $c \in C_1$.
- d:
$$d - \pi_3 : 0.4 + 0 + 0.1 + 0.2 = 0.7 \geq 0.7$$
So $d \in C_3$.
- e:
$$e - \pi_3 : 0 + 0.3 + 0 + 0.2 = 0.5 < 0.7$$
$$e - \pi_2 : 0 + 0.3 + 0 + 0.2 = 0.5 < 0.7$$
So $e \in C_1$.
- f:
$$f - \pi_3 : 0 + 0 + 0.1 + 0 = 0.1 < 0.7$$
$$f - \pi_3 : 0 + 0.3 + 0.1 + 0.2 = 0.6 < 0.7$$
So $f \in C_1$.

Therefor, $C_1 = \{b, c, e, f\}, C_2 = \{a\}$ and $C_3 = \{b\}$.

6 Sample Exam

Exercise 6.1. Voting Rules

Let us consider the following preferences of $n = 17$ voters, given on a set $X = \{a, b, c, d\}$ of $m = 4$ candidates:

$$5\ voters : c \succ b \succ a \succ d$$
$$4\ voters : b \succ a \succ c \succ d$$
$$4\ voters : d \succ b \succ c \succ a$$
$$2\ voters : a \succ c \succ b \succ d$$
$$2\ voters : d \succ b \succ a \succ c$$

Which candidate is elected by using the following voting rules or principles:

1. **Plurality voting.**

   The elected candidate is d, with 6 votes.

2. **Plurality runoff voting (plurality with two rounds).**

   Since the most voted candidate is d with 6 votes, a second round is needed. c and d go to the second round, it remains:

   $$5\ voters : c \succ d$$
   $$4\ voters : c \succ d$$
   $$4\ voters : d \succ c$$
   $$2\ voters : c \succ d$$
   $$2\ voters : d \succ c$$

   So the winner is c with 11 votes.

3. **Condorcet principle.**

   We obtain the following Condorcet matrix:

	a	b	c	d
a		0	0	1
b	1		1	1
c	1	0		1
d	0	0	0	

   So, the Condorcet winner is b.

4. **Borda principle.**

   We assign the ranks and compute the Borda scores:

   $$B(a) = 5 \cdot 3 + 4 \cdot 2 + 4 \cdot 4 + 2 \cdot 1 + 2 \cdot 3 = 47$$
   $$B(b) = 5 \cdot 2 + 4 \cdot 1 + 4 \cdot 2 + 2 \cdot 3 + 2 \cdot 2 = 32$$
   $$B(c) = 5 \cdot 1 + 4 \cdot 3 + 4 \cdot 3 + 2 \cdot 2 + 2 \cdot 4 = 41$$
   $$B(d) = 5 \cdot 4 + 4 \cdot 4 + 4 \cdot 1 + 2 \cdot 4 + 2 \cdot 1 = 50$$

   So, the Borda winner is b.

5. **Copeland principle. The Copeland principle associates to each candidate, x, a score calculated as follows:**

   $$S_{Cop}(x) = \sum_{y \in X, y \neq x} Cop(x, y),$$

where X is the set of candidates and $Cop\,(x,y) = \begin{cases} 1 & \textbf{if a majority of voters prefer } x \textbf{ to } y \\ -1 & \textbf{if a majority of voters prefer } y \textbf{ to } x \\ 0 & \textbf{if the both two previous situations arise simultan} \end{cases}$

The candidate with the highest Copeland score is elected.

$$S_{Cop}\,(a) = -1 - 1 + 1 = -1$$
$$S_{Cop}\,(b) = 1 + 1 + 1 = 3$$
$$S_{Cop}\,(c) = 1 - 1 + 1 = 1$$
$$S_{Cop}\,(d) = -1 - 1 - 1 = -3$$

So, the Copelan winner is b.

6. **Kramer-Simpson principle. The Kramer-Simpson principle associates to each candidate, x, a score:**

$$KS\,(x) = \min_{y \in X, y \neq x} n\,(x,y)\,,$$

where X is the set of candidates and $n\,(x,y)$ is the number of voters who prefer x to y. The candidate with the highest Kramer-Simpson score is elected.

In the following table, we represent the number of voters that prefer the candidate in the row to the candidate in the column:

	a	b	c	d
a		2	8	11
b	15		10	11
c	9	7		11
d	6	6	6	

$$KS\,(a) = 2$$
$$KS\,(b) = 10$$
$$KS\,(c) = 7$$
$$KS\,(d) = 6$$

So, the Kramer-Simpson winner is b.

In general, does the Copeland principle elect the Condorcet winner, if this latter exists? Justify your answer.

Assume the Condorcet winner exists, say W. Then, for all other candidate, c, we know that W is preferred to c by a majority of voters. This means that, for all other candidate c, it is

$$Cop\,(W,c) = 1,$$

so

$$S_{Cop}\,(W) = n - 1,$$

where n is the total number of candidates. Note this is the maximum possible Copeland value, so we only have to check if no other candidate can reach this value.

Indeed, since $Cop\,(W,c) = 1$, for all c, then $Cop\,(c,W) = -1$, and so no other candidate c can reach the maximum score. Therefore, W is also the Copeland winner.

In general, does the Kramer-Simpson principle elect the Condorcet winner, if this latter exists? Justify your answer.

Assume the Condorcet winner exists, say W. Then, for all other candidate, c, we know that W is preferred to c by a majority of voters. This means that, for all other candidate c, it is

$$n\,(W,c) > \frac{n}{2}.$$

Therefore,

$$KS(W) = \min_{c \neq W} n(W, c) > \frac{n}{2}.$$

On the other hand, it must be

$$n(c, W) < \frac{n}{2},$$

so

$$KS(c) < \frac{n}{2} < KS(W).$$

Therefore, W is also the Kramer-Simpson winner.

Let us consider the followwing preferences of 4 voters given on a set $X = \{a, b, c, d, e\}$ of 5 candidates:

$$\begin{aligned}
1 \; voter &: a \succ b \succ c \succ d \succ e \\
1 \; voter &: b \succ c \succ d \succ a \succ e \\
1 \; voter &: c \succ d \succ a \succ b \succ e \\
1 \; voter &: d \succ a \succ b \succ c \succ e
\end{aligned}$$

The Borda principle is chosen to elect the winner of this election.

1. **Who is elected?**

$$\begin{aligned}
B(a) &= 1 \cdot 1 + 1 \cdot 4 + 1 \cdot 3 + 1 \cdot 2 = 10 \\
B(b) &= 1 \cdot 2 + 1 \cdot 1 + 1 \cdot 4 + 1 \cdot 3 = 10 \\
B(c) &= 1 \cdot 3 + 1 \cdot 2 + 1 \cdot 1 + 1 \cdot 4 = 10 \\
B(d) &= 1 \cdot 4 + 1 \cdot 3 + 1 \cdot 2 + 1 \cdot 1 = 10 \\
B(e) &= 1 \cdot 5 + 1 \cdot 5 + 1 \cdot 5 + 1 \cdot 5 = 20
\end{aligned}$$

There is a quadruple tie.

2. **By adding 3 new voters to the previous 4 ones (and we have now a total of 7 voters), is it possible to provide the preferences of these new voters such that the candidate e is elected (the winner is not necessary unique)? Justify your answer.**

The minimum score that e can reach is

$$B'(e) = B(e) + 3 \cdot 1 = 23,$$

if it's the first choice for the 3 new voters.

From 3 voters, there are a total of $3 \cdot (1 + 2 + 3 + 4 + 5) = 45$ votes to allocate. We have allocated 3 of them to e, so there are 42 points left to allocate. In order for e to be the winner, we need

$$B'(e) \leq B'(x),$$

for $x \in \{a, b, c, d\}$. This means that

$$23 = B'(e) \leq B'(x) = 10 + k_x,$$

which implies that $k_x \geq 13$, for all x. So

$$\sum_x k_x \geq 52.$$

However, we only have 42 points left. Therefore, e cannot be the Borda winner by adding 3 voters.

3. **By adding 4 new voters to the previous one (and we have now a total of 8 voters), is it possible to provide the preferences of these new voters such that the candidate e is elected (the winner is not necessary unique)? Justify your answer.**

   In this case, the minimum score that e can reach is

$$B'(e) = B(e) + 4 \cdot 1 = 24.$$

   From 4 voters, there are a total of $4 \cdot (1+2+3+4+5) = 60$ points to allocate. We have allocated 4 of them to e, so there are 56 points left to allocate. In order for e to be the winner, we need

$$B'(e) \leq B'(x),$$

   for $x \in \{a, b, c, d\}$. This means that

$$24 = B'(e) \leq B'(x) = 10 + k_x,$$

   which implies that $k_x \geq 14$, for all x. So

$$\sum_x k_x \geq 56,$$

   which is a quantity that we can afford. The following 4 voters would make e win (with ties):

$$\begin{aligned} 1 \ voter :& e \succ a \succ b \succ c \succ d \\ 1 \ voter :& e \succ b \succ a \succ d \succ c \\ 1 \ voter :& e \succ c \succ d \succ a \succ b \\ 1 \ voter :& e \succ d \succ c \succ b \succ a \end{aligned}$$

Is it possible to add new voters to the previous 4 ones (by keeping the same 4 types of preferences) such that a is the unique Condorcet winner? Justify your answer.

Yes, we just need to add this voter to the first profile:

$$\begin{aligned} 5 \ voter :& a \succ b \succ c \succ d \succ e \\ 1 \ voter :& b \succ c \succ d \succ a \succ e \\ 1 \ voter :& c \succ d \succ a \succ b \succ e \\ 1 \ voter :& d \succ a \succ b \succ c \succ e \end{aligned}$$

In this case, a would be the unique Borda winner.

Exercise 6.2. Ranking or sorting?

We consider the following six students a, b, c, d, e, f evaluated on three subjects $N = \{1, 2, 3\}$. The scores on each criterion are given in the interval $[0, 100]$. The strict preference, given by the Decision Maker (DM), is denoted by $\succ$, while his indifference preference is denoted by $\sim$. The performance matrix of the students evaluations is the following:

	1: Mathematics	2: Statistics	3: Language
a	85	90	75
b	80	70	70
c	80	65	70
d	85	90	60
e	50	65	75
f	50	70	60

The two parts below are independent and can be solved separately.

Part 1: Ranking

1. **We assume that the DM provides the following preference information: $a \succ b$, $c \succ d$ and $e \succ f$. Are these preferences representable by an additive model? Justify your answer.**

   Yes, we can take $w_1 = w_2 = 0, w_3 = 1$ and the identity as marginal utility functions. This way $F(a) > F(b), F(c) > F(d)$ and $F(e) > F(f)$.

2. **We assume that the DM provides the following preference information: $a \succ b$, $c \succ d$ and $e \sim f$. Are these preferences representable by an additive model? Justify your answer.**

   It is

   $$F(a) > F(b) \iff w_1 \cdot u_1(85) + w_2 \cdot u_2(90) + w_3 \cdot u_3(75) > w_1 \cdot u_1(80) + w_2 \cdot u_2(70) + w_3 \cdot u_3(75)$$

   $$\iff w_1 (u_1(85) - u_1(80)) + w_2 (u_2(90) - u_2(70)) + w_3 (u_3(75) - u_3(70)) > 0.$$

   Also,

   $$F(c) > F(d) \iff w_1 (u_1(80) - u_1(85)) + w_2 (u_2(65) - u_2(90)) + w_3 (u_3(70) - u_3(60)) > 0.$$

   And,

   $$F(e) = F(f) \iff w_1 (u_1(50) - u_1(50)) + w_2 (u_2(65) - u_2(70)) + w_3 (u_3(75) - u_3(60)) = 0$$

   $$\iff w_2 (u_2(65) - u_2(70)) + w_3 (u_3(75) - u_3(60)) = 0.$$

   Now, if we sum $(F(a) - F(b)) + (F(c) - F(d))$, the quantity must remain positive, i.e.,

   $w_1 (\cancel{u_1(85)} - \cancel{u_1(80)} + \cancel{u_1(80)} - \cancel{u_1(85)}) + w_2 (\cancel{u_2(90)} - u_2(70) + u_2(65) - \cancel{u_2(90)}) + w_3 (u_3(75) - \cancel{u_3(70)} + \cancel{u_3(70)} - u$

   that is

   $$w_2 (u_2(65) - u_2(70)) + w_3 (u_3(75) - u_3(60)) > 0.$$

   So we have the same quantity equalling 0, and being greater than 0. This is impossible, and therefore these preferences cannot be represented as an additive model.

3. **Now, we assume that the marginal utility function u_i associated to the criterion i, is exactly the obtained marks, i.e., $u_i(x_i) = x_i$, where x_i is a value on the criterion i.**

   (a) **Determine the ranking $\succsim_1$ of the 6 students by using a weighted sum, where the weights (ECTS) associated to the 3 subjects is the vector $(w_1 = 6, w_2 = 3, w_3 = 2)$, w_i being the weight associated to the criterion i.**

   Let's compute the scores:

	1: Mathematics	2: Statistics	3: Language	Score
a	85	90	75	930
b	80	70	70	830
c	80	65	70	815
d	85	90	60	900
e	50	65	75	645
f	50	70	60	630

   So the ranking is $a \succ d \succ b \succ c \succ e \succ f$.

   (b) **Is it possible to represent the following preferences by a weighted sum model?**

   - $d \succ c$ **and** $f \succ e$**.**
   - **Language is strictly more important than Mathematics.**
   - **There is no weight (of a criterion) equal to 0.**

   Justify your answer.

   The three preferences can be described mathematically as:

- $F(d) > F(c)$ and $F(f) > F(e)$.
- $w_3 > w_1$.
- $w_1, w_2, w_3 > 0$.

Let's analyze the implications. First

$$F(d) > F(c) \iff 85w_1 + 90w_2 + 60w_3 > 80w_1 + 65w_2 + 70w_3$$

$$\iff 5w_1 + 25w_2 - 10w_3 > 0 \iff w_1 + 5w_2 - 2w_3 > 0$$

$$F(f) > F(e) \iff 50w_1 + 70w_2 + 60w_3 > 50w_1 + 65w_2 + 75w_3$$

$$\iff 5w_2 - 15w_3 > 0 \iff w_2 > 3w_3.$$

The first restriction is equivalent to

$$5w_2 > 2w_3 - w_1,$$

using the second restriction, we obtain that

$$5w_2 > 15w_3,$$

so making

$$15w_3 > 2w_3 - w_1 \iff 13w_3 > -w_1$$

works. Since this is always true $(w_1, w_3 \geq 0)$, it is enough to ensure the second restriction. Therefore, choosing (w_1, w_2, w_3) such that $w_3 > w_1$ and $w_2 > 3w_3$ should work. For example, $(w_1 = 1, w_2 = 7, w_3 = 2)$. In this case, the scores are:

	1: Mathematics	2: Statistics	3: Language	Score
a	85	90	75	865
b	80	70	70	710
c	80	65	70	675
d	85	90	60	835
e	50	65	75	655
f	50	70	60	660

So, the ranking is $a \succ d \succ b \succ c \succ f \succ e$ and all conditions are met.

Part 2: Sorting

We aim at developing a multi-criteria method assigning the 6 students to some ordered categories. The envisaged method is baed on the elaboration of an outranking relation, as it is done, for instance, in MR-Sort method. However, unlike MR-Sort where each alternative is compared to reference profiles representing the boundaries of the categories, the outranking relation here is defined on the given set of alternatives.

- **In the sequel, we denote by $A = \{a, b, c, ...\}$ the set of alternatives to assign and N the set of criteria.**
- **The outranking relation $\succsim$ means "at least as good as", with $\succ$ its asymmetric part and $\sim$ its symmetric part. The binary relation $\succsim$ is defined by**

$$a \succsim b \iff \sum_{i | g_i(a) \geq g_i(b)} w_i \geq \lambda,$$

where $w_i \geq 0$ represents the weight associated to the criterion i $\left(\sum_{i\in N} w_i = 1\right)$, $g_i(a)$ represents the value of the alternative a on the criterion i and $\lambda \in [0.5, 1]$ is the majority threshold.

- **Two alternatives a and b are said "incomparable" if $[not\,(a \succsim b) \wedge not\,(b \succsim a)]$.**
- **The p ordered categories we consider are denoted $C_1, C_2, ...C_p$ (C_1 and C_p being respectively the worst and the best category).**

- $C(a)$ **represents the category where the alternative** a **is assigned.**
- $C(a) \geq C(b)$ **means that** a **is assigned to a category greater than the category where** b **is assigned.**
- $C(a) > C(b)$ **means that** a **is assigned to a category strictly greater than the category where** b **is assigned.**

Let us consider the following assignment principles:

$$\forall a, b \in A, C(a) > C(b) \implies a \succsim b,$$

$$\forall a, b \in A, a \succsim b \implies C(a) \geq C(b).$$

To assign the 6 students, we consider the following parameters and preferences:

- **We have 4 categories** C_1, C_2, C_3 **and** C_4**.**
- **The majority threshold is** $\lambda = 0.7$**.**
- **The weight vector is** $(w_1 = 0.3, w_2 = 0.3, w_3 = 0.4)$**.**
- **The student** a **belongs to the category** C_4**.**
- **The student** f **belongs to the category** C_1**.**

Determine the assignments of the other students by using these preferences and the adopted assignment principle.

In the following table, I compute the index $\sum_{i|g_i(a)\geq g_i(b)} w_i$:

	a	b	c	d	e	f
a		1	1	1	1	1
b	0		1	0.4	0.6	1
c	0	0.7		0.4	0.6	0.7
d	0.6	0.6	0.6		0.6	1
e	0.4	0.4	0.7	0.4		0.7
f	0	0.3	0.3	0.4	0.6	

Therefore, we have

$$a \succsim b, c, d, e, f$$

$$b \succsim c, f$$

$$c \succsim b, f$$

$$d \succsim f$$

$$e \succsim c, f$$

Since $b \succsim c$ and $c \succsim b$, we have that $C(b) \geq C(c)$ and $C(c) \geq C(b)$, so $C(b) = C(c)$.

Also, we have that $e \succsim c$, so $C(e) \geq C(c)$.

The following classification meets all these criteria:

	Student
C_4	a
C_3	e, d
C_2	b, c
C_1	f

But the solution is not unique, for example this also works:

	Student		Student
C_4	a	C_4	a, e
C_3	e	C_3	d
C_2	b, c, d	C_2	b, c
C_1	f	C_1	f

And many more. If we enable categories to be empty, there are even more possibilities, like

	Student		Student
C_4	a	C_4	a
C_3		C_3	b, c, e, d
C_2	b, c, d, e	C_2	
C_1	f	C_1	f

and so on.

6.1 Another Exam

Exercise 6.3. Ranking or sorting?

We consider 8 students of a Master program, evaluated on 6 subjects (criteria supposed to be maximized) as follows:

Alternatives	CR_1	CR_2	CR_3	CR_4	CR_5	CR_6
a_1	10	10	10	10	10	10
a_2	9	11	10	11	11	9
a_3	11	9	10	11	11	9
a_4	10	10	9	9	10	9
a_5	9	11	11	10	10	9
a_6	11	10	10	11	10	10
a_7	10	9	10	11	10	10
a_8	11	10	11	12	9	9

The two parts below are independent and can be solved separately.

Part 1: Ranking

1. **We assume that each score is given on a scale $[0, 20]$, corresponding to a value of the marginal utility function associated to a subject.**

   (a) **Determine the ranking $\succsim_1$ of the 8 students by using a weighted sum, where weights (ECTS) associated to the 6 subjects is the vector $W_1 = (6; 3; 2; 6; 2; 6)$.**

   Let's compute the complete scores.

   $W_1(a_1) = 250, W_1(a_2) = 249, W_1(a_3) = 255, W_1(a_4) = 236,$

   $$W_1(a_5) = 243, W_1(a_6) = 262, W_1(a_7) = 253, W_1(a_8) = 262$$

   So the ranking is

   $$[a_6, a_8] \succsim a_3 \succsim a_7 \succsim a_1 \succsim a_2 \succsim a_5 \succsim a_4$$

(b) **Determine the ranking $\succsim_2$ of the 8 students by using a weighted sum where the weights (ECTS) associated to the 6 subjects is the vector $W_2 = (3;4;2;6;2;4)$. Are the rankings $\succsim_1$ and $\succsim_2$ different?**
$W_2(a_1) = 210, W_2(a_2) = 215, W_2(a_3) = 213, W_2(a_4) = 198,$

$$W_2(a_5) = 209, W_2(a_6) = 219, W_2(a_7) = 212, W_2(a_8) = 221$$

So the ranking is

$$a_8 \succsim a_6 \succsim a_2 \succsim a_3 \succsim a_7 \succsim a_1 \succsim a_5 \succsim a_4.$$

And both rankings are different.

(c) **In fact, the director of the Master has some preferences given as follows:**

i. **The students a_3 and a_2 are judged equivalent**
ii. **The student a_2 is strictly preferred to the student a_1**

Does a weight vector W_3 exist such that these preferences are representable by a weighted sum? Justify your answer and if yes, give the ranking $\succsim_3$ obtained by applying W_3 to the student's dataset.
We have $W_3(a_3) = W_3(a_2)$ and $W_3(a_2) > W_3(a_1)$. This implies that:

$$\begin{aligned} & w_1 \cdot 11 + w_2 \cdot 9 + \cancel{w_3 \cdot 10 + w_4 \cdot 11} + w_5 \cdot 9 + w_6 \cdot 11 \\ = & \, w_1 \cdot 9 + w_2 \cdot 11 + \cancel{w_3 \cdot 10 + w_4 \cdot 11} + w_5 \cdot 11 + w_6 \cdot 9 \end{aligned}$$

We obtain

$$2w_1 - 2w_2 - 2w_5 + 2w_6 = 0,$$

or

$$w_1 - w_2 - w_5 + w_6 = 0.$$

On the other hand

$$\begin{aligned} & w_1 \cdot 9 + w_2 \cdot 11 + \cancel{w_3 \cdot 10} + w_4 \cdot 11 + w_5 \cdot 11 + w_6 \cdot 9 \\ > & \, w_1 \cdot 10 + w_2 \cdot 10 + \cancel{w_3 \cdot 10} + w_4 \cdot 10 + w_5 \cdot 10 + w_6 \cdot 10 \end{aligned}$$

So

$$-w_1 + w_2 + w_4 + w_5 - w_6 > 0.$$

Now, using the previous equation, we get $w_4 > 0$. It seems like choosing $w_1 = w_2 = w_5 = w_6$ and w_4 strictly positive, we get such ranking. Lets use $W_3 = (1;1;1;1;1;1)$. The scores are
$W_2(a_1) = 60, W_2(a_2) = 61, W_2(a_3) = 61, W_2(a_4) = 57,$

$$W_2(a_5) = 60, W_2(a_6) = 62, W_2(a_7) = 60, W_2(a_8) = 62$$

So the ranking is

$$[a_6, a_8] \succsim [a_2, a_3] \succsim [a_1, a_5, a_7] \succsim a_4.$$

(d) **Prove that the following preferences of the director of the Master are not representable by a weighted sum:**

i. **student a_1 is strictly preferred to the student a_3**
ii. **student a_7 is strictly preferred to the student a_1**
iii. **criterion 6 is more important than criterion 5**

In this case, we have $W_7(a_7) > W_4(a_1) > W_4(a_3)$. From the first inequality we get

$$\begin{aligned} & \cancel{w_1 \cdot 10} + w_2 \cdot 9 + \cancel{w_3 \cdot 10} + w_4 \cdot 11 + \cancel{w_5 \cdot 10} + \cancel{w_6 \cdot 10} \\ > & \, \cancel{w_1 \cdot 10} + w_2 \cdot 10 + \cancel{w_3 \cdot 10} + w_4 \cdot 10 + \cancel{w_5 \cdot 10} + \cancel{w_6 \cdot 10} \end{aligned}$$

which is equivalent to

$$-w_2 + w_4 > 0.$$

On the other hand,

$$\begin{aligned} & w_1 \cdot 10 + w_2 \cdot 10 + \cancel{w_3 \cdot 10} + w_4 \cdot 10 + w_5 \cdot 10 + w_6 \cdot 10 \\ & > w_1 \cdot 11 + w_2 \cdot 9 + \cancel{w_3 \cdot 10} + w_4 \cdot 11 + w_5 \cdot 9 + w_6 \cdot 11 \end{aligned}$$

which equates to

$$-w_1 + w_2 - w_4 + w_5 - w_6 > 0.$$

Since $w_2 - w_4 < 0$, it is sure that

$$-w_1 + w_2 - w_4 + w_5 - w_6 < -w_1 + w_5 - w_6.$$

Since $w_1 \geq 0$, it is sure that

$$-w_1 + w_2 - w_4 + w_5 - w_6 < -w_1 + w_5 - w_6 \leq w_5 - w_6.$$

And since $w_6 > w_5$, we obtain that

$$0 < -w_1 + w_2 - w_4 + w_5 - w_6 < 0$$

which is a contradiction. Therefore, this preferences are not representable by a weighted sum.

Part 2: Sorting

We aim at developing a multi-criteria method assigning the 8 students to some ordered categories. The envisaged method is based on the elaboration of an outranking relation, as it is done, for instance, in MR-Sort method. However, unlike MR-Sort where each alternative is compared to reference profiles representing the boundaries of the categories, the outranking relation here is defined on the given set of the alternatives.

In the sequel, we denote by $A = \{a, b, c, ...\}$ the set of alternatives to assign and N the set of n criteria.

The outranking relation $\succsim$ means 'at least as good as', with $\succ$ its asymmetric part and $\sim$ is symmetric part.

Two alternatives a and b are "incomparable" if $[not\,(a \succsim b) \wedge not\,(b \succsim a)]$.

The p categories we consider are denoted $C_1, C_2, ..., C_p$ (C_1 and C_p being respectively the worst and the best category).

- **$C(a)$ represents the category where the alternative a is assigned.**
- **$C(a) \geq C(b)$ means that a is assigned to a category greater than the category where b is assigned.**
- **$C(a) > C(b)$ means that a is assigned to a category strictly greater than the category where b is assigned.**

1. **Let us consider the following assignment principle:**

$$\forall a, b \in A, C(a) > C(b) \implies a \succsim b.$$

Prove that, if this principle is used, then two incomparable alternatives are necessarily assigned to the same category.

Suppose a and b are incomparable. By reductio ad absurdum, suppose they are not assigned to the same category. Without loss of generality, we may assume $C(a) > C(b)$. However, this implies that $a \succsim b$. # This is a contradiction, because a and b are incomparable. The same goes if $C(b) > C(a)$. Therefore, it must be $C(a) = C(b)$.

2. **Let us consider the following assignment principle:**

$$\forall a, b \in A, a \succsim b \implies C(a) \geq C(b).$$

Prove that this principle implies $a \succ b \implies C(a) > C(b)$ **and** $a \sim b \implies C(a) = C(b)$**.**

On one hand:

$a \sim b \iff a \succsim b \wedge b \succsim a \implies C(a) \geq C(b) \wedge C(b) \geq C(a) \implies C(a) = C(b)$.

The other implication cannot be proved using the given assumptions (I think xD).

3. **In this question, we consider the previous assignment principle. We suppose that the outranking relation is defined by**

$$a \succsim b \iff \sum_{i | g_i(a) \geq g_i(b)} w_i \geq \lambda,$$

where $w_i \geq 0$ **represents the weights associated to the criterion** i **and** $\sum_{i\in N} w_i = 1$**.** $g_i(a)$ **represents the value of the alternative** a **on the criterion** i **and** $\lambda \in [0.5, 1]$ **is the majority threshold.**

To assign the 8 students of the Master, we consider the following parameters and preferences:

- **We have 4 categories** C_1, C_2, C_3 **and** C_4
- **The majority threshold is** $\lambda = \frac{2}{3}$
- **The weight vector is** $W_5 = (0.1; 0.2; 0.05; 0.4; 0.05; 0.2)$
- **The students** a_2 **and** a_2 **belong to the category** C_3
- **The student** a_1 **belongs to the category** C_2

Determine the assignments of the other students by using these preferences and the adopted assignment principle.

Let's compute the comparison value for each pair of alternatives:

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
a_1		0.35	0.3	1	0.75	0.5	0.6	0.45
a_2	0.7		0.7	0.9	0.95	0.7	0.7	0.45
a_3	0.75	0.75		0.75	0.7	0.75	0.95	0.35
a_4	0.35	0.3	0.25		0.35	0.25	0.35	0.45
a_5	0.7	0.55	0.3	0.9		0.3	0.3	0.5
a_6	1	0.75	0.8	1	0.75		1	0.55
a_7	0.8	0.75	0.7	0.8	0.75	0.7		0.25
a_8	0.75	0.75	0.8	0.95	0.75	0.75	0.75	

Therefore

$$a_1 \succsim a_4, a_5$$
$$a_2 \succsim a_1, a_3, a_4, a_5, a_6, a_7$$
$$a_3 \succsim a_1, a_2, a_4, a_5, a_6, a_7$$
$$a_4 \succsim \emptyset$$
$$a_5 \succsim a_1, a_4$$
$$a_6 \succsim a_1, a_2, a_3, a_4, a_5, a_7$$
$$a_7 \succsim a_1, a_2, a_3, a_4, a_5, a_6$$
$$a_8 \succsim a_1, a_2, a_3, a_4, a_5, a_6, a_7$$

So $a_1 \sim a_5$, $a_2 \sim a_3 \sim a_5 \sim a_6 \sim a_7$, with

$$a_8 \succ [a_2, a_3, a_5, a_6, a_7] \succ [a_1, a_5]$$

So the classification must be

C_1	
C_2	a_1, a_5
C_3	a_2, a_3, a_5, a_6, a_7
C_4	a_8

References

[1] Brice Mayag. Decision modeling. Lecture Notes.