Decision Methods and Models

Master's Degree in Computer Science

Roberto Cordone

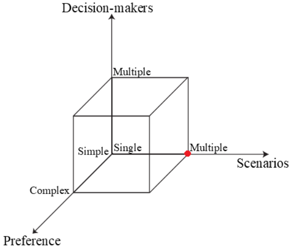
DI - Università degli Studi di Milano

Schedule:	Thursday 16.30 - 18.30 in Aula Magna (CS department)
	Friday 12.30 - 14.30 in classroom 301
Office hours:	on appointment
E-mail:	roberto.cordone@unimi.it
Web page:	https://homes.di.unimi.it/cordone/courses/2024-mmd/2024-mmd.html
Ariel site:	https://myariel.unimi.it/course/view.php?id=4467

Decision-making under uncertainty

We assume

- a preference relation Π that is a weak order
  with a known consistent value function $u(f)$ (replaced by a cost f)
- a uncertain environment: $|\Omega| > 1$
- a single decision-maker: $|D| = 1 \Rightarrow \Pi_d$ reduces to Π



The problem becomes $\min_{x \in X} f(x, \omega)$ with $\omega \in \Omega$ known after x is selected

If ω is known in advance, the problem falls into parametric optimisation:
solution $x^*(\omega)$ represents a strategy prepared for every possible scenario

Applications

The situation has applications in many different fields

- finance: decide how much to invest and in what
- marketing: decide prices and advertising campaigns
- research and development: decide which research projects to finance and how much
- ...

but also less obvious fields such as algorithmics

- X is the set of alternative algorithms for a problem
- Ω is the set of instances
- $f(x, \omega)$ is the optimality gap (or computational time) achieved by algorithm x on instance ω

Varieties of uncertainty

Descriptive models provide information on the scenarios

Different kinds of information generate different classes of problems

1. ignorance: the only information is the scenario set Ω
2. risk: we know the scenario set Ω and a probability function π_ω
3. ...

Notice that even ruling out unpredicted scenarios is hard ("black swans")

Given X and Ω, other descriptive models provide the impacts $f(x, \omega)$

Strong dominance

Suitable conditions allow to introduce a relation on alternatives
that can filter out some of them and simplify (or even solve) the problem

$$x \preceq x' \Leftrightarrow f(x,\omega) \preceq f(x',\omega) \quad \text{for all } \omega \in \Omega$$

If the strong domination is strict, solution x' can be safely removed

The definition is formally equivalent to Paretian preference,
with scenario $\omega \in \Omega$ replacing indicators $l \in P$: $f_l(x) \rightarrow f(x,\omega)$

- similar methods find nondominated solutions and Paretian solutions
- but, while the indicators are p, the scenarios can be infinitely many (*in that case, the extension is not trivial*)

Example

Consider a set of investments with the corresponding impacts (benefits)

ROI	Recession	Moderate growth	Strong growth
Stock fund	-25%	0%	35%
Bond fund	-10%	5%	15%
Treasuries	8%	8%	8%
“Toxic” fund	-5%	6%	8%

The toxic fund is dominated by treasuries

Probabilistic dominance

$$x \preceq x' \quad \Leftrightarrow \quad Pr(f(x) \leq f) \geq Pr(f(x') \leq f) \text{ for all } f \in F$$

State (die result)	1	2	3	4	5	6
Gamble A wins \$	1	1	2	2	2	2
Gamble B wins \$	1	1	1	2	2	2
Gamble C wins \$	3	3	3	1	1	1

first-order stochastic dominance (FSD)

$P[A \geq x] \geq P[B \geq x]$ for all x,

A strongly dominates B (yield as good for all possible outcomes of the die roll)

C does not strongly dominate B (B gives a better yield in scenarios 4 to 6)

C probabilistically dominates B because

- $Pr(B \geq 1) = Pr(C \geq 1) = 1$
- $Pr(B \geq 2) = Pr(C \geq 2) = 3/6$
- $Pr(B \geq 3) = 0 < Pr(C \geq 3) = 3/6$

A and C exhibit no probabilistic dominance

- $Pr(A \geq 2) = 4/6 > Pr(C \geq 2) = 3/6$
- $Pr(C \geq 3) = 3/6 > Pr(A = \geq 3) = 0$

Dominance and algorithm performance

A typical application is the selection of heuristic algorithms where

- an alternative x is a heuristic algorithm A
- a scenario ω is a problem instance I
- the impact $f(x, \omega)$ is the percent gap (relative difference) $\delta_A(I)$

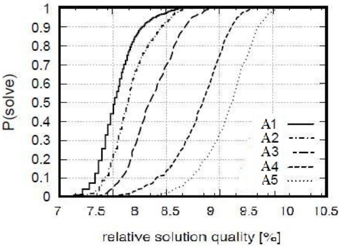
In this context

- strong dominance means a lower gap on all instances (very rare)

$$A \preceq A' \quad \Leftrightarrow \quad \delta_A(I) \leq \delta_{A'}(I) \text{ for all } I \in P$$

- probabilistic dominance means hitting below any given threshold $\bar{\delta}$ for more instances (higher *Solution Quality Distribution* diagram):

$$A \preceq A' \quad \Leftrightarrow \quad Pr(\delta_A(I) \leq \bar{\delta}) \geq Pr(\delta_{A'(I)} \leq \bar{\delta}) \text{ for all } f \in F$$



Different scenario modelling approaches

The scenario set Ω is in general a set of vectors in $\mathbb{R}^s$

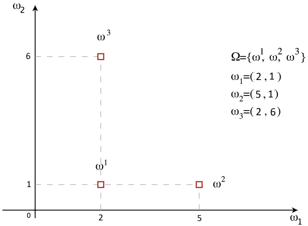
Two notable cases are

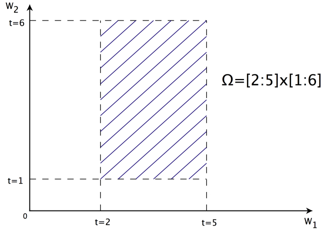
1. the scenario-based description (common in economy):
   Ω is a finite list of scenarios

$$\Omega = \left\{\omega^{(1)}, \ldots, \omega^{(|\Omega|)}\right\}$$

2. the interval-based description (common in engineering): Ω is the Cartesian product of a s real intervals on the exogenous variables

$$\Omega = \left[\omega_1^{\text{min}}, \omega_1^{\text{max}}\right] \times \ldots \times \left[\omega_s^{\text{min}}, \omega_s^{\text{max}}\right]$$





Example

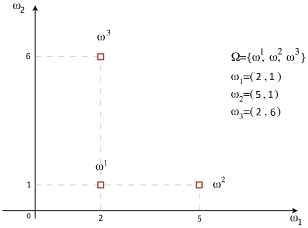
Suppose that we are computing the fastest path on a street network in which two arcs have uncertain travel times ω_1 and ω_2

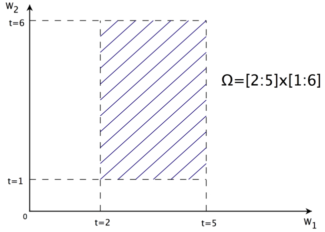
- in a scenario-based description, we might have

$$\omega = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \text{ or } \begin{bmatrix} 5 \\ 1 \end{bmatrix} \text{ or } \begin{bmatrix} 2 \\ 6 \end{bmatrix} \quad \Leftrightarrow \quad \Omega = \{(2,1),(5,1),(2,6)\}$$

- in an interval-based description, we might have

$$\begin{cases} 2 \leq \omega_1 \leq 5 \\ 1 \leq \omega_2 \leq 6 \end{cases} \quad \Leftrightarrow \quad \Omega = [2,5] \times [1,6]$$





How to choose?

The appropriate description depends on

- the precision of the representation
- the simplicity of the solution process (*is there an algorithm?*)

On the one hand, the scenario description has a finite number of cases

- are they all and only the possible cases? (precision)
- are they few or combinatorially many? (simplicity)

On the other hand, the interval description

- implies that the exogenous variable are independent (precision)
- the scenario are infinitely many, but the worst scenario might be
  - the same for all solutions
  - easy to find (simplicity)

For example, take the maximum for all travel times