Decision and belief functions

Sébastien Destercke

CNRS researcher, Laboratoire Heudiasyc, Compiègne



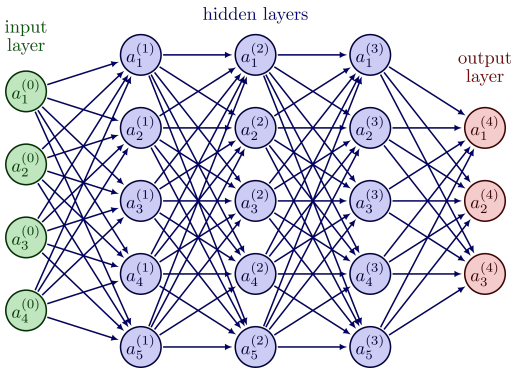
BFAS automn school courses

What I looked at

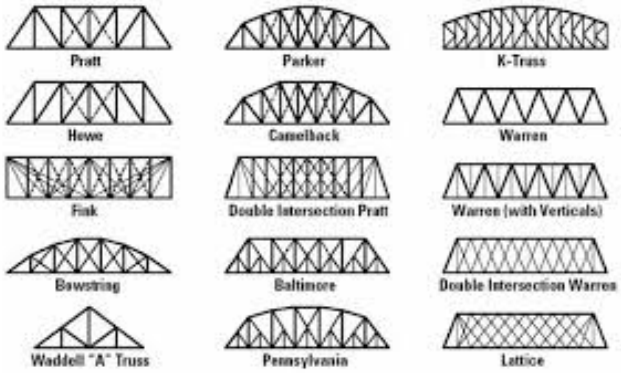


Decision is everywhere

Classification



System design



Planification



- Finance,
- Risk analysis,
- . . .

Questions you may have answer to after the talk

- How should I decide given my uncertainty?
- Do I want a unique recommendation, or just filter out bad decisions?
- What are the pros and cons of the different options?
- When do these problems become (even more) complex?

Outline

- Decision with extreme cases
  - Probabilities
  - Ignorance, complete order
  - Ignorance, partial orders
- Decision with belief functions
- Overview of other (advanced) topics

Decision setting

- Still a set $\mathscr{X}$ of states
- A set $\mathscr{A}$ of actions
- To each action $a : \mathscr{X} \to \mathbb{R}$ corresponds a mapping such that $a(x)$ is the reward/utility of performing a when x is true
- An uncertainty formalism modelling our knowledge about X

Decision problem (here): recommend one or multiple actions based on our knowledge about the states in $\mathscr{X}$

An example

What we see when walking at night

- States: the kind of animal
- Actions: what you choose to do



States $\mathscr{X}$

x_1 =(L)ioness x_2 =(P)uma x_3 =(C)at x_4=(O)celot






Actions $\mathscr{A}$

States $\mathscr{X}$

x_1 =(L)ioness x_2 =(P)uma x_3 =(C)at x_4=(O)celot






Actions $\mathscr{A}$

a_1 =(R)un



States $\mathscr{X}$

x_1 =(L)ioness x_2 =(P)uma x_3 =(C)at x_4=(O)celot






Actions $\mathscr{A}$

a_1 =(R)un a_2 =(S)hout




States $\mathscr{X}$

x_1 =(L)ioness x_2 =(P)uma x_3 =(C)at x_4=(O)celot






Actions $\mathscr{A}$

a_1 =(R)un a_2 =(S)hout a_3 =(D)o nothing





The matrix $\mathscr{U}$

	L	P	C	O
R	7	−3	3	5
S	−10	4	4	11
D	−5	−5	5	7

Which action to choose?

Outline

- Decision with extreme cases
  - Probabilities
  - Ignorance, complete order
  - Ignorance, partial orders
- Decision with belief functions
- Overview of other (advanced) topics

A quick reminder of expected utility

- Set of states $\mathscr{X}$
- Probability masses $p(x)$ for each state
- Act/decision a maps each state to a utility

$$a : \mathscr{X} \to \mathbb{R}$$

- Expected utility of a

$$\mathbb{E}(a) = \sum_x p(x)a(x)$$

An easy decision rule

One possible interpretation: $\mathbb{E}(a)$ is your price/value[1] of act a, then

$$a \succ b \text{ if and only if } \mathbb{E}(a) > \mathbb{E}(b)$$

that is, you would "spend" more to buy a than b.

[1] In a linear utility. Quique, que es muy elegante y intelligente, explained it yesterday.

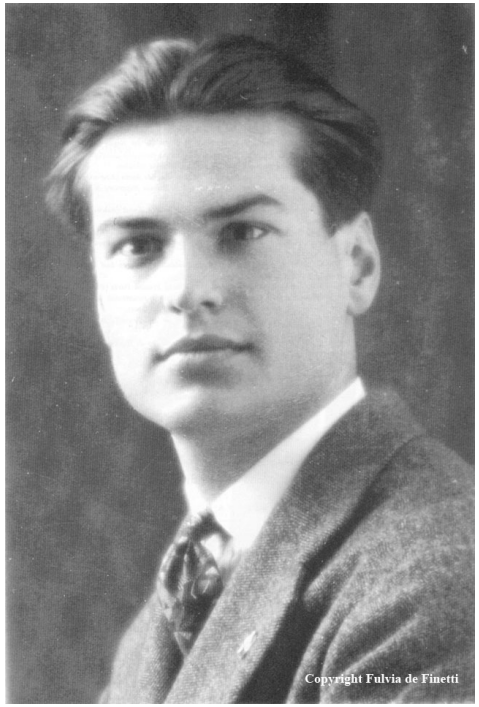
Three important guys



L. Savage [24]



J. Von Neumann [26]



B. De Finetti [16]

All justify probabilities (and expected utilities) as uncertainty models without frequencies.

An example

This looks like a big animal, so more a puma or lioness than a cat or ocelot ($P(L) = P(P) = 0.3$, $P(C) = P(O) = 0.2$)

	L	P	C	O	$\mathbb{E}(a)$
R	7	−3	3	5	2.8
S	−10	4	4	11	
D	−5	−5	5	7	
Max					

$$\mathbb{E}(R) = 0.3 \cdot 7 + 0.3 \cdot -3 + 0.2 \cdot 3 + 0.2 \cdot 5 = 2.8$$

An example

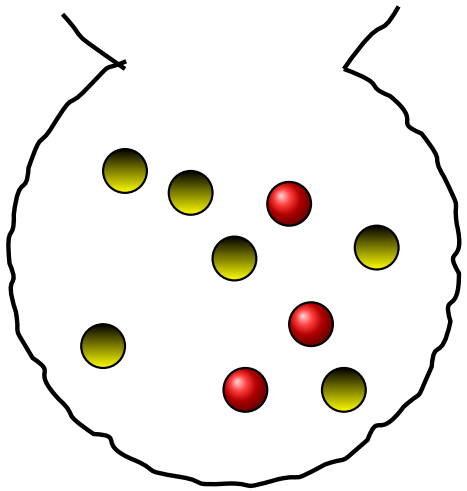
This looks like a big animal, so more a puma or lioness than a cat or ocelot ($P(L) = P(P) = 0.3$, $P(C) = P(O) = 0.2$)

	L	P	C	O	$\mathbb{E}(a)$
R	7	−3	3	5	2.8
S	−10	4	4	11	1.2
D	−5	−5	5	7	−0.6
Max					2.8

$$R \succ_{\mathbb{E}} S \succ_{\mathbb{E}} D$$

Urns and balls: case 1 [13]

9 balls, 3 are reds, 6 remaining are either yellow or black



What would you choose between A and B?

A

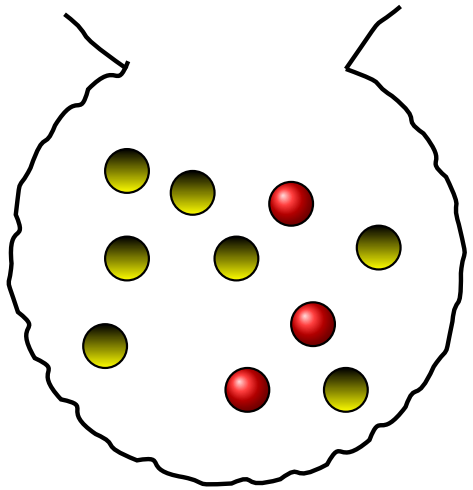
R(ed)	B(lack)	Y(ellow)
100 $	0 $	0$

B

R(ed)	B(lack)	Y(ellow)
0 $	100 $	0$

Urns and balls: case 2

9 balls, 3 are reds, 6 remaining are either yellow or black



What would you choose between C and D?

C

R(ed)	B(lack)	Y(ellow)
100 $	0 $	100$

D

R(ed)	B(lack)	Y(ellow)
0 $	100 $	100$

Ellsberg or probabilities?

9 balls, 3 are reds, 6 remaining are either yellow or black

A

R(ed)	B(lack)	Y(ellow)
100 $	0 $	0$

B

R(ed)	B(lack)	Y(ellow)
0 $	100 $	0$

C

R(ed)	B(lack)	Y(ellow)
100 $	0 $	100$

D

R(ed)	B(lack)	Y(ellow)
0 $	100 $	100$

- Most people: Prefer A to B, and D to C
- But $\mathbb{E}(A) = 100P(R) > \mathbb{E}(B) = 100P(B) \implies P(R) > P(B)$,
- and $\mathbb{E}(D) = 100(1 - P(R)) > \mathbb{E}(C) = 100(1 - P(B)) \implies P(B) > P(R)$
- **A clear contradiction!**

A first important conclusion

One cannot both think that

- decision theory should account Ellsberg experiment and;
- use only probability and expectation as a decision tool.

At least two solutions offered to you:

- Declare that most people are irrational and that probabilities are right → this was probably not the right school for you?
- Try to relax probability theory to account for the observations → path we will take with a (late) focus on belief functions

Outline

- Decision with extreme cases
  - Probabilities
  - Ignorance, complete order
  - Ignorance, partial orders
- Decision with belief functions
- Overview of other (advanced) topics

Maximin: pessimistic behaviour

- For each action a_i, compute $\underline{\mathbb{E}}(a_i) = \min_{x_j \in \mathscr{X}} a_i(x_j)$
- Say that $a_k \succ_{Mm} a_\ell$ if $\underline{\mathbb{E}}(a_k) > \underline{\mathbb{E}}(a_\ell)$

	L	P	C	O	$\underline{\mathbb{E}}(a)$
(R)	7	−3	3	5	−3
S	−10	4	4	11	−10
D	−5	−5	5	7	−5
Max					(-3)

- We get $R \succ_{Mm} D \succ_{Mm} S$, hence R is recommended
- Pessimistic attitude: best action in the worst case

Why (not) maximin?

Pros:

- Provide strong guarantees on selected option
- Is very-well studied in the literature

Cons:

- Can be very conservative
- Selected option may be suboptimal in all cases

	L	P	C	O	$\underline{\mathbb{E}}(a)$
R	7	−3	3	5	−3
S	−10	4	4	11	−10
D	−5	−5	5	7	−5
K	−2	−2	−2	−2	−2
Max					(-2)

K never the best solution!

Maximax: optimistic behaviour

- For each action a_i, compute $\overline{\mathbb{E}}(a_i) = \max_{x_j \in \mathscr{X}} a_i(x_j)$
- Say that $a_k \succ_{MM} a_\ell$ if $\overline{\mathbb{E}}(a_k) > \overline{\mathbb{E}}(a_\ell)$

	L	P	C	O	$\overline{\mathbb{E}}(a)$
R	7	−3	3	5	7
(S)	−10	4	4	11	11
D	−5	−5	5	7	7
Max					(11)

- We get $S \succ_{MM} D \approx_{MM} R$, hence S is recommended
- Optimistic attitude: best action in the best case

Why (not) maximax?

Pros:

- Often computationally easier (max/max problem)
- Solution always potentially optimal

Cons:

- May be rather bold (and in our case your last decision)
- Do not provide strong guarantees on reward

In-between: Hurwicz

- Pick a value $\alpha \in [0,1]$, called optimism index
- For a_i, compute

$$u_{H(\alpha)}(a_i) = \alpha \overline{\mathbb{E}}(a_i) + (1-\alpha)\underline{\mathbb{E}}(a_k)$$

- Say that $a_k \succ_\alpha a_\ell$ if $u_{H(\alpha)}(a_k) > u_{H(\alpha)}(a_\ell)$

	L	P	C	O	$\underline{\mathbb{E}}(a_i)$	$\overline{\mathbb{E}}(a_i)$	$u_{H(0.5)}(a_i)$
(R)	7	−3	3	5	−3	7	2
S	−10	4	4	11	−10	11	0.5
D	−5	−5	5	7	−5	7	1
Max							(2)

- We get $R \succ_{H(0.5)} D \succ_{H(0.5)} S$, hence R is recommended
- Try to balance between optimistic and pessimistic

Why (not) Hurwicz

Pros:

- Try to find a compromise between minimax/maximax
- Well-axiomatised

Cons:

- None of the guarantees of minimax/maximax
- All their defaults, but to a lesser extent
- Operational elicitation of α challenging

Savage Minimax regret

- For action a_i, compute $Re(a_i, x_j) = \max_{a_k \in \mathscr{A}} a_k(x_j) - a_i(x_j)$ the regret of picking a_i in x_j, instead of the best possible action
- For a_i, compute $Re^\star(a_i) = \max_j Re(a_i, x_j)$
- Say that $a_k \succ_{Re} a_\ell$ if $Re^\star(a_\ell) > Re^\star(a_k)$

	L	P	C	O	$Re^\star(a_i)$
R	7	−3	3	5	
$Re(R)$	0				
S	−10	4	4	11	
$Re(S)$					
D	−5	−5	5	7	
$Re(D)$					
Min					

Savage Minimax regret

- For action a_i, compute $Re(a_i, x_j) = \max_{a_k \in \mathscr{A}} a_k(x_j) - a_i(x_j)$ the regret of picking a_i in x_j, instead of the best possible action
- For a_i, compute $Re^\star(a_i) = \max_j Re(a_i, x_j)$
- Say that $a_k \succ_{Re} a_\ell$ if $Re^\star(a_\ell) > Re^\star(a_k)$

	L	P	C	O	$Re^\star(a_i)$
R	7	-3	3	5	
$Re(R)$	0	7			
S	-10	4	4	11	
$Re(S)$					
D	-5	-5	5	7	
$Re(D)$					
Min					

Savage Minimax regret

- For action a_i, compute $Re(a_i, x_j) = \max_{a_k \in \mathscr{A}} a_k(x_j) - a_i(x_j)$ the regret of picking a_i in x_j, instead of the best possible action
- For a_i, compute $Re^\star(a_i) = \max_j Re(a_i, x_j)$
- Say that $a_k \succ_{Re} a_\ell$ if $Re^\star(a_\ell) > Re^\star(a_k)$

	L	P	C	O	$Re^\star(a_i)$
R	7	−3	3	5	
$Re(R)$	0	7	2	6	7
S	−10	4	4	11	
$Re(S)$					
D	−5	−5	5	7	
$Re(D)$					
Min					

Savage Minimax regret

- For action a_i, compute $Re(a_i, x_j) = \max_{a_k \in \mathscr{A}} a_k(x_j) - a_i(x_j)$ the regret of picking a_i in x_j, instead of the best possible action
- For a_i, compute $Re^\star(a_i) = \max_j Re(a_i, x_j)$
- Say that $a_k \succ_{Re} a_\ell$ if $Re^\star(a_\ell) > Re^\star(a_k)$

	L	P	C	O	$Re^\star(a_i)$
R	7	−3	3	5	
$Re(R)$	0	7	2	6	7
S	−10	4	4	11	
$Re(S)$	17	0	1	0	17
D	−5	−5	5	7	
$Re(D)$	12	9	0	4	12
Min					

Savage Minimax regret

- For action a_i, compute $Re(a_i, x_j) = \max_{a_k \in \mathscr{A}} a_k(x_j) - a_i(x_j)$ the regret of picking a_i in x_j, instead of the best possible action
- For a_i, compute $Re^\star(a_i) = \max_j Re(a_i, x_j)$
- Say that $a_k \succ_{Re} a_\ell$ if $Re^\star(a_\ell) > Re^\star(a_k)$

	L	P	C	O	$Re^\star(a_i)$
R	7	−3	3	5	
$Re(R)$	0	7	2	6	7
S	−10	4	4	11	
$Re(S)$	17	0	1	0	17
D	−5	−5	5	7	
$Re(D)$	12	9	0	4	12
Min					7

- We get $R \succ_{Re} D \succ_{Re} S$, hence R is recommended
- Minimize regret, but sensitive to addition of non-optimal alternatives

Minimax regret vs maximin

Consider the following case:

	x_1	$\cdots$	x_{99}	x_{100}	$R^\star(a_i)$	$\underline{\mathbb{E}}(a_i)$
a_1	10	$\cdots$	10	1		
$R(a_1)$						
a_2	2	$\cdots$	2	2		
$R(a_2)$						
Min						

Minimax regret vs maximin

Consider the following case:

	x_1	$\cdots$	x_{99}	x_{100}	$R^\star(a_i)$	$\underline{\mathbb{E}}(a_i)$
a_1	10	$\cdots$	10	1		1
$R(a_1)$	0	$\cdots$	0	1	1	
a_2	2	$\cdots$	2	2		2
$R(a_2)$	8	$\cdots$	8	0	8	
Min					1	2

- Maximin: a_2
- Minimax regret: a_1

Minimax regret and irrelevant alternatives

Before: $R \succ_{Re} D \succ_{Re} S$

Now I get (A)nti-puma!



	L	P	C	O	$Re^{\star}(a_i)$
R	7	−3	3	5	
$Re(R)$					
S	−10	4	4	11	
$Re(S)$					
D	−5	−5	5	7	
$Re(D)$					
A	−15	15	−2	0	
$Re(A)$					
Min					

After A is possible:

Minimax regret and irrelevant alternatives

Before: $R \succ_{Re} D \succ_{Re} S$

Now I get (A)nti-puma!



	L	P	C	O	$Re^{\star}(a_i)$
R	7	−3	3	5	
$Re(R)$	0	18	2	6	18
S	−10	4	4	11	
$Re(S)$	17	11	1	0	17
D	−5	−5	5	7	
$Re(D)$	12	10	0	4	12
A	−15	15	−2	0	
$Re(A)$	22	0	5	11	22
Min					12

After A is possible: $D \succ_{Re} S \succ_{Re} R \succ_{Re} A$ (change on preferences not involving A)

Complete ordering: summary

- Minimax=pessimistic [29]
- Maximax=optimistic
- Hurwicz=in-between [1]
- Savage=Minimizing felt regret [24]

Whatever the chosen rule, we always get one optimal action. But we need to commit to a peculiar behaviour.

What if DM does not want to commit to peculiar behaviour?

What if DM wants to only know the actions that are potentially optimal, given our uncertainty?

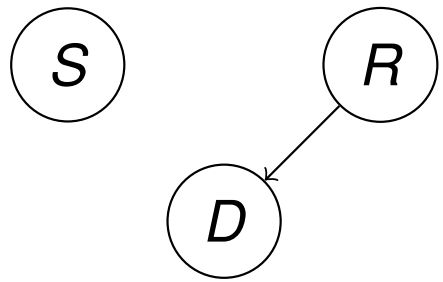
Outline

- Decision with extreme cases
  - Probabilities
  - Ignorance, complete order
  - Ignorance, partial orders
- Decision with belief functions
- Overview of other (advanced) topics

Lattice ordering

- Say that $a_k \succeq_L a_\ell$ if $\overline{\mathbb{E}}(a_k) \geq \overline{\mathbb{E}}(a_\ell)$ and $\underline{\mathbb{E}}(a_k) \geq \underline{\mathbb{E}}(a_\ell)$

	L	P	C	O	$\underline{\mathbb{E}}(a_i)$	$\overline{\mathbb{E}}(a_i)$
R	7	−3	3	5	−3	7
S	−10	4	4	11	−10	11
D	−5	−5	5	7	−5	7



- Only existing dominance is D by R, hence only D is considered non-optimal
- Can be seen as a robust Hurwicz (considering all α as possibilities)
- Note that with this criterion, we eliminate the best action in state C
- This rule looks good, why should I not use it?

Information-monotonic preferences

Principle

A (partial) decision rule is information monotonic if gaining more information can only refine our current preferences

Definition for imprecise information

A decision rule $\succeq$ is information monotonic if for any pair of actions a, b and for $E' \subseteq E$, we have

$$(E \implies a \succeq b) \implies (E' \implies a \succeq b),$$

meaning that the preference relation can only be refined by getting more precise

Lattice ordering and information monotonicity

	x_1	x_2	x_3	x_4	$\underline{\mathbb{E}}(a_i)$	$\overline{\mathbb{E}}(a_i)$
a	10	12	14	15	10	15
b	13	11	16	14	11	16

$$b \succ a$$

All states possible

Lattice ordering and information monotonicity

	~~x_1~~	x_2	x_3	x_4	$\underline{\mathbb{E}}(a_i)$	$\overline{\mathbb{E}}(a_i)$
a	~~10~~	12	14	15	12	15
b	~~13~~	11	16	14	11	16

$$b \succ\prec a$$

We learn (gain info) x_1 impossible

a and b becomes incomparable.

Lattice ordering and information monotonicity

	~~x_1~~	x_2	~~x_3~~	x_4	$\underline{\mathbb{E}}(a_i)$	$\overline{\mathbb{E}}(a_i)$
a	~~10~~	12	~~14~~	15	12	15
b	~~13~~	11	~~16~~	14	11	14

$$b \prec a$$

We learn (gain info) x_3 impossible

a is now preferred to b.

Refining Pareto

Principle

A (partial) decision rule refines Pareto if it ensures that its preference order refines Pareto dominance

Definition for imprecise information

A decision rule $\succeq$ refines Pareto if for any pair of actions a, b

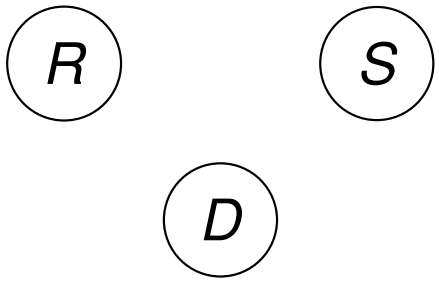
$$a \succeq_P b \implies a \succeq b,$$

where $\succeq_P$ is Pareto dominance

Interval dominance

- Say that $a_k \succ_{ID} a_\ell$ if $\underline{\mathbb{E}}(a_k) > \overline{\mathbb{E}}(a_\ell)$

	L	P	C	O	$\underline{\mathbb{E}}(a_i)$	$\overline{\mathbb{E}}(a_i)$
R	7	−3	3	5	−3	7
S	−10	4	4	11	−10	11
D	−5	−5	5	7	−5	7

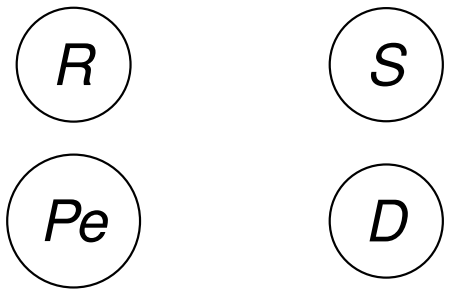


- no dominance at all
- overcautious criterion → may retain Pareto-dominated solutions

Interval dominance: drawback example

- We add a fourth possible action *Pe*=Pet the animal

	L	P	C	O	$\underline{\mathbb{E}}(a_i)$	$\overline{\mathbb{E}}(a_i)$
R	7	−3	3	5	−3	7
S	−10	4	4	11	−10	11
D	−5	−5	5	7	−5	7
Pe	−5	−5	2	−5	−5	2

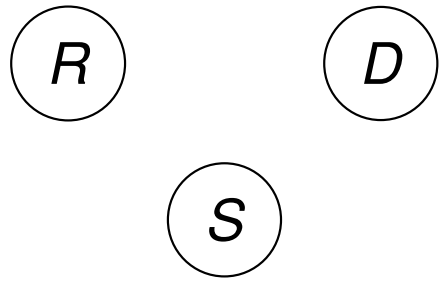


- no dominance, all intervals intersect
- even if *D* better (sometimes strictly) than *Pe* in **every** situation!

Difference dominance

- Say that $a_k \succeq_D a_\ell$ if $a_k(x_j) - a_\ell(x_j) \geq 0$ for all x_j ($\succ$ if > 0 for at least one x_j)

	L	P	C	O
R	7	−3	3	5
S	−10	4	4	11
D	−5	−5	5	7
$R-S$	17	−7	−1	−6

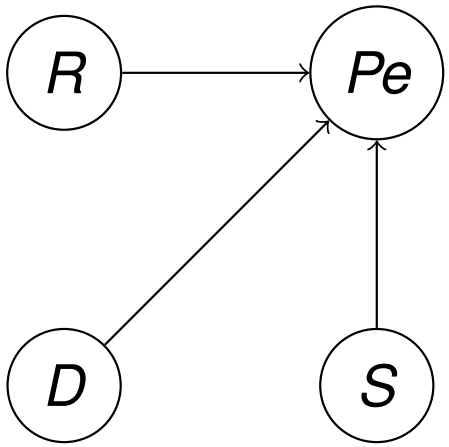


- no dominance at all, again
- do we have the same problem as with interval dominance?

Difference comparison

- We add a fourth possible action *Pe*=Pet the animal

	L	*P*	*C*	*O*
R	7	−3	3	5
S	−10	4	4	11
D	−5	−5	5	7
Pe	−5	−5	2	−5
D − *Pe*	0	0	3	12



- In the ignorance case, exactly Pareto ordering!

So far...

Options when true state of the world completely unknown:

- Complete ordering/one top recommendation
  - Maximin: pessimistic DM
  - Maximax: optimistic DM
  - Hurwicz: attempt to in-between
  - Regret minimax: pessimistic DM w.r.t. regret
- Partial ordering/multiple recommendations reflecitng lack of knowledge
  - Lattice ordering: robust hurwicz, may miss potentially optimal actions
  - Difference dominance: will keep every non-Pareto dominated solution
  - Interval dominance: very conservative, may keep Pareto dominated options

A summary of the different rules

Property	$\succ_{Mm}$	$\succ_{MM}$	$\succ_{R}$	$\succ_{H(\alpha)}$	$\succ_{L}$	$\succ_{D}$	$\succ_{ID}$
Complete	✓	✓	✓	✓	x	x	x
Reco possibly optimal	x*	✓	x*	x*	✓	✓	✓
Guaranteed value	✓	x	✓	x	N.A.	N.A.	N.A.
$a \succ b$ depends only on a, b	✓	✓	x	✓	✓	✓	✓
Refines Pareto	✓	✓	✓	✓	✓	✓	x
Information monotonic	x	x	x	x	x	✓	✓

We also have $\succ_L \subseteq \succ_D \subseteq \succ_{ID}$, going from the most precise to the least

* Not guaranteed to be.

Outline

- Decision with extreme cases
- Decision with belief functions
- Overview of other (advanced) topics

From ignorance to set-valued decision

- Consider set E and action a
- We can associate to E a worst case

$$\underline{\mathbb{E}}^{E}(a) = \inf_{x \in E} a(x)$$

- We can associate to E a best case

$$\overline{\mathbb{E}}^{E}(a) = \sup_{x \in E} a(x)$$

From sets to belief functions

- Consider a belief function $m : 2^{\mathcal{X}} \to \mathbb{R}$ and action a
- We can associate a worst-case evaluation

$$\underline{\mathbb{E}}^m(a) = \sum_{E \subseteq \mathcal{X}} m(E)\, \underline{\mathbb{E}}^E(a) = \sum_{E \subseteq \mathcal{X}} m(E) \inf_{x \in E} a(x)$$

- We can associate a best-case evaluation

$$\overline{\mathbb{E}}^m(a) = \sum_{E \subseteq \mathcal{X}} m(E)\, \overline{\mathbb{E}}^E(a) = \sum_{E \subseteq \mathcal{X}} m(E) \sup_{x \in E} a(x)$$

- Probabilistic interpretation: considering Bel/Pl from m, we define

$$\mathcal{P}^m = \{P : \forall A, Bel(A) \leq P(A) \leq Pl(A)\}$$

and we do have

$$\underline{\mathbb{E}}^m(a) = \inf_{p \in \mathcal{P}^m} \mathbb{E}^p(a), \quad \overline{\mathbb{E}}^m(a) = \sup_{p \in \mathcal{P}^m} \mathbb{E}^p(a)$$

Previous decision rules adaptation [6]

In general, replace $\overline{\mathbb{E}}$ by upper expectation $\overline{\mathbb{E}}^m$, $\underline{\mathbb{E}}$ by lower expectation $\underline{\mathbb{E}}^m$

Total order

- Maximax: $a \succeq_{MM} b$ if $\overline{\mathbb{E}}^m(a) \geq \overline{\mathbb{E}}^m(b)$
- Maximin: $a \succeq_{Mm} b$ if $\underline{\mathbb{E}}^m(a) \geq \underline{\mathbb{E}}^m(b)$
- Hurwicz: $a \succeq_{\alpha} b$ if $\alpha\overline{\mathbb{E}}^m(a) + (1-\alpha)\underline{\mathbb{E}}(a) \geq \alpha\overline{\mathbb{E}}^m(b) + (1-\alpha)\underline{\mathbb{E}}(b)$

Partial order

- Interval dominance: $a \succ_{ID} b$ if $\overline{\mathbb{E}}^m(b) \leq \underline{\mathbb{E}}^m(a)$
- Lattice: $a \succ_{L} b$ if $\overline{\mathbb{E}}^m(b) \leq \overline{\mathbb{E}}^m(a) \wedge \underline{\mathbb{E}}^m(b) \leq \underline{\mathbb{E}}^m(a)$
- Difference dominance: $a \succ_{D} b$ if $\underline{\mathbb{E}}^m(a-b) \geq 0$

→ they suffer from same pros/cons discussed in the case of ignorance.

Back to Ellsberg

9 balls, 3 are reds, 6 remaining are either yellow or black

A

R(ed)	B(lack)	Y(ellow)
100 $	0 $	0$

B

R(ed)	B(lack)	Y(ellow)
0 $	100 $	0$

C

R(ed)	B(lack)	Y(ellow)
100 $	0 $	100$

D

R(ed)	B(lack)	Y(ellow)
0 $	100 $	100$

- What are the possible probability values? In terms of bounds over each colour?
- Compute the lower/upper expectations for each act
- What kind of comparison explain the most frequent behaviour $A \geq B$ but $D \geq C$?

Back to Ellsberg

9 balls, 3 are reds, 6 remaining are either yellow or black

A

R(ed)	B(lack)	Y(ellow)
100 $	0 $	0$

B

R(ed)	B(lack)	Y(ellow)
0 $	100 $	0$

C

R(ed)	B(lack)	Y(ellow)
100 $	0 $	100$

D

R(ed)	B(lack)	Y(ellow)
0 $	100 $	100$

- $m(\{R\}) = 1/3$, $m(\{B, Y\}) = 2/3$
- $\underline{\mathbb{E}}^m(A) = \overline{\mathbb{E}}^m(A) = 100/3$, $[\underline{\mathbb{E}}^m(B), \overline{\mathbb{E}}^m(B)] = [0, 200/3]$
- $[\underline{\mathbb{E}}^m(C), \overline{\mathbb{E}}^m(C)] = [100/3, 100]$, $\underline{\mathbb{E}}^m(D) = \overline{\mathbb{E}}^m(D) = 200/3$
- Maximin: $D \succ C \approx A \succ B$, Lattice: $A \succ\prec B$, $C \succ\prec D$

An example of belief function, finally!



Set of labellers replying between

- L(ioness)
- P(uma)
- C(at)
- O(celot)

- 20% reply $\{L,P\} \rightarrow m(\{L,P\}) = 0.2$
- 10% reply $\{L,P,O\} \rightarrow m(\{L,P,O\}) = 0.1$
- 15% reply $\{C,O,P\} \rightarrow m(\{C,O,P\}) = 0.15$
- 10% reply $\{L\} \rightarrow m(\{L\}) = 0.1$
- 40% reply $\{L,P,C,O\} \rightarrow m(\mathscr{S}) = 0.4$

Back to our issue: example of some rules (not all)

Is R better than S, or the other way around?

	L	P	C	O
Run	7	−3	3	5
Shout	−10	4	4	11
Do nothing	−5	−5	5	7

- $m(\{L,P\}) = 0.2$
- $m(\{L,P,O\}) = 0.1$
- $m(\{C,O,P\}) = 0.15$
- $m(\{L\}) = 0.1$
- $m(\mathscr{S}) = 0.4$

Evaluation: averaging utility + attitude (pessimist/optimist/robust)

Back to our issue: example of some rules (not all)

Is R better than S, or the other way around?

	L	P	C	O
Run	7	−3	3	5
Shout	−10	4	4	11
Do nothing	−5	−5	5	7

- $m(\{L,P\}) = 0.2$
- $m(\{L,P,O\}) = 0.1$
- $m(\{C,O,P\}) = 0.15$
- $m(\{L\}) = 0.1$
- $m(\mathscr{S}) = 0.4$

Evaluation: averaging utility + attitude (pessimist/optimist/robust)

$$\underline{\mathbb{E}}(Run) = 0.2 \times -3 + 0.1 \times -3 + 0.15 \times -3 + 0.1 \times 7 + 0.4 \times -3 = -2$$
$$\underline{\mathbb{E}}(Shout) = 0.2 \times -10 + 0.1 \times -10 + 0.15 \times 4 + 0.1 \times -10 + 0.4 \times -10 = -7.9$$

Seems better to Run than Shout

Back to our issue: example of some rules (not all)

Is R better than S, or the other way around?

	L	P	C	O
Run	7	−3	3	5
Shout	−10	4	4	11
Do nothing	−5	−5	5	7

- $m(\{L,P\}) = 0.2$
- $m(\{L,P,O\}) = 0.1$
- $m(\{C,O,P\}) = 0.15$
- $m(\{L\}) = 0.1$
- $m(\mathscr{S}) = 0.4$

Evaluation: averaging utility + attitude (pessimist/optimist/robust)

$$\overline{\mathbb{E}}(Run) = 0.2 \times 7 + 0.1 \times 7 + 0.15 \times 5 + 0.1 \times 7 + 0.4 \times 7 = 6.7$$
$$\overline{\mathbb{E}}(Shout) = 0.2 \times 4 + 0.1 \times 11 + 0.15 \times 11 + 0.1 \times -10 + 0.4 \times 11 = 6.95$$

Seems better to Shout than Run

Back to our issue: example of some rules (not all)

Is R better than S, or the other way around?

	L	P	C	O
Run	7	−3	3	5
Shout	−10	4	4	11
Do nothing	−5	−5	5	7

- $m(\{L,P\}) = 0.2$
- $m(\{L,P,O\}) = 0.1$
- $m(\{C,O,P\}) = 0.15$
- $m(\{L\}) = 0.1$
- $m(\mathscr{S}) = 0.4$

Evaluation: averaging utility + attitude (pessimist/optimist/robust)

$$[\underline{\mathbb{E}}(Run), \overline{\mathbb{E}}(Run)] = [-2, 6.7]$$
$$[\underline{\mathbb{E}}(Shout), \overline{\mathbb{E}}(Shout)] = [-7.9, 6.95]$$

Not entirely clear what is the best (as Puma and Lion are both likely)

A generalised parametric rule

- Consider for each (focal) set E values $\beta^E, \gamma^E \in [0,1]$ with

$$\beta^E \leq \gamma^E$$

as pessimism/optimist degrees

- Compute

$$\underline{\mathbb{E}}^{\beta}(a) = \sum_E m(E)\beta_E \overline{\mathbb{E}}^E(a) + (1-\beta_E)\underline{\mathbb{E}}^E(a)$$

$$\overline{\mathbb{E}}^{\gamma}(a) = \sum_E m(E)\gamma_E \overline{\mathbb{E}}^E(a) + (1-\gamma_E)\underline{\mathbb{E}}^E(a)$$

- $a \succeq^{\beta,\gamma} b$ if and only if

$$\underline{\mathbb{E}}^{\beta}(b) \leq \underline{\mathbb{E}}^{\beta}(a) \vee \overline{\mathbb{E}}^{\gamma}(b) \leq \overline{\mathbb{E}}^{\gamma}(a)$$

Retrieving standard rules

This form includes as extremes the following rules:

- Hurwicz with parameter α: $\beta = \gamma = \alpha$
- Lattice: $\beta = 0, \gamma = 1$
- Difference dominance: minimax on act difference

Any theoretical insight?

This rule or similar instances have been axiomatised in different works:

- by Jaffray with a VNM-style argument [19]
- by Shenoy and Denoeux with a VNM-style argument [8] using Dempster's rule
- by Zhou and colleagues using an Anscombe-Aumann setting [30]

A (very) short note on Pignistic rule

- Take the pignistic transform *Betp* of m
- Use $\mathbb{E}^{Betp}$ to make your choice

Initially proposed by Smets [25], it is also a specific case of the axiomatisation proposed by Eichberge and Pasichnichenko [12] in economy, and it is consistent with some previous rules:

- Will not be dominated by Difference dominance
- Corresponds to a specific choice of $\beta^E = \gamma^E$ returning for each focal set the average expected utility

Outline

- Decision with extreme cases
- Decision with belief functions
- Overview of other (advanced) topics
  - Other decision rules
  - Sequential decisions
  - Decisions with multiple agents
  - Decisions with multiple criteria
  - Decision in combinatorial spaces

Outline

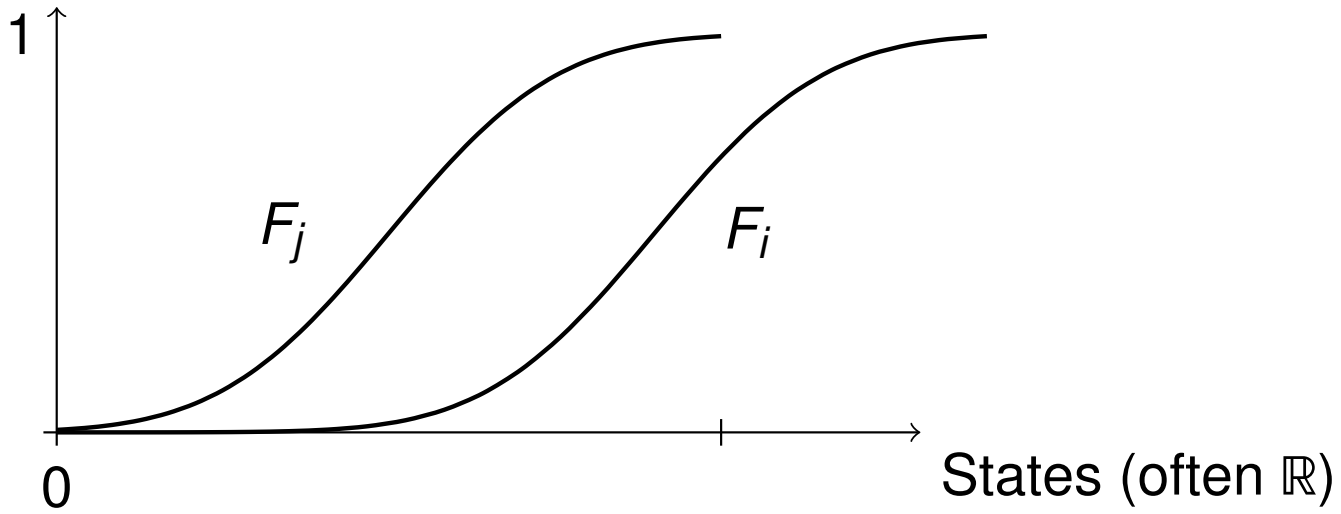
- Decision with extreme cases
- Decision with belief functions
- Overview of other (advanced) topics
  - Other decision rules
  - Sequential decisions
  - Decisions with multiple agents
  - Decisions with multiple criteria
  - Decision in combinatorial spaces

Stochastic dominance

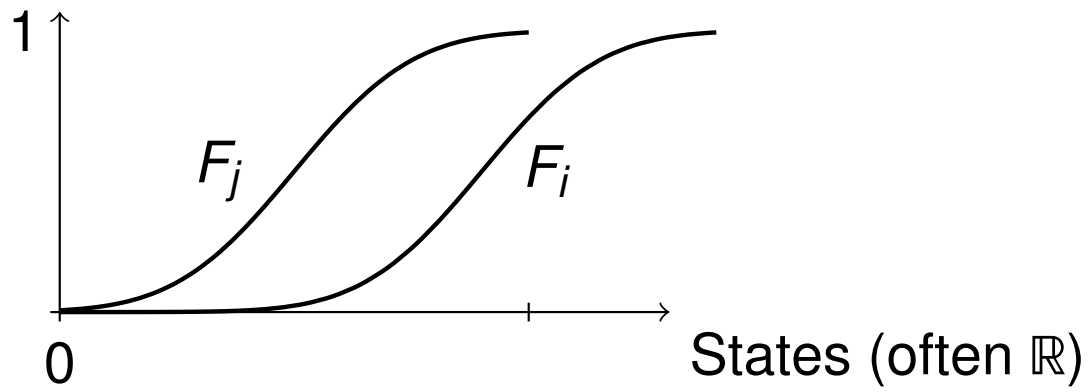
Assumptions and definitions

- States can be arranged from "worst" to "best" linearly
- For action a_i and each state x, we can estimate $F_i(x) = P(]-\infty, x])$

Then a_i stochastically dominates a_j ($a_i \succ_{SD} a_j$) if $F_i(x) \leq F_j(x)$



Why stochastic dominance?



Interests

- $a_i \succ_{SD} a_j \implies \mathbb{E}(a_i) \geq \mathbb{E}(a_j)$ for any increasing utility over $\mathscr{X}$
- Can be seen as a robust version of expected utility
- Only require to be able to rank states (no numbers necessary)

For extensions to belief functions and imprecise CDF, look at [5, 22, 3]

Statistical preference

Assumptions and definitions

- Utilities can be compared with each other (ordered)
- Each action/state is given a (comparable) utility grade

Then a_i is statistically preferred to a_j if $P(a_i \geq a_j) \geq 0.5$

$$p(x_3) = 0.1, p(x_2) = 0.7, p(x_1) = 0.2$$

	x_1	x_2	x_3
a_i	*high*	*high*	*low*
a_j	*low*	*very high*	*very low*
$\mathbb{I}_{a_i \geq a_j}$	1	0	1

$$P(a_j \geq a_i) = 0.7 \implies a_j \text{ stat. preferred to } a_j$$

BF extension to be studied? Yet works on making it imprecise [21, 2]

Statistical preference and transitivity

Two players play the following game. Player 1 erases the spots from the faces of three fair dice and writes one number from 1, 2,. . . , 18 to each face. Each of them risks 1 euro, chooses one dice, they throw the dice, and the one having the bigger number on top of his dice wins the 2 euros. Since Player 1 puts the numbers to the dice it seems fair to let Player 2 choose his dice first. Of course, Player 2 tries to choose the best dice. Despite this, Player 1 wins in the long run.

How can that be?

The dice transitivity problem [4]

Consider the following numbering of the fair Dices A, B, C

- $A = \{1, 3, 4, 15, 16, 17\}$
- $B = \{2, 10, 11, 12, 13, 14\}$
- $C = \{5, 6, 7, 8, 9, 18\}$

Assuming that dices are thrown independently, show that we have

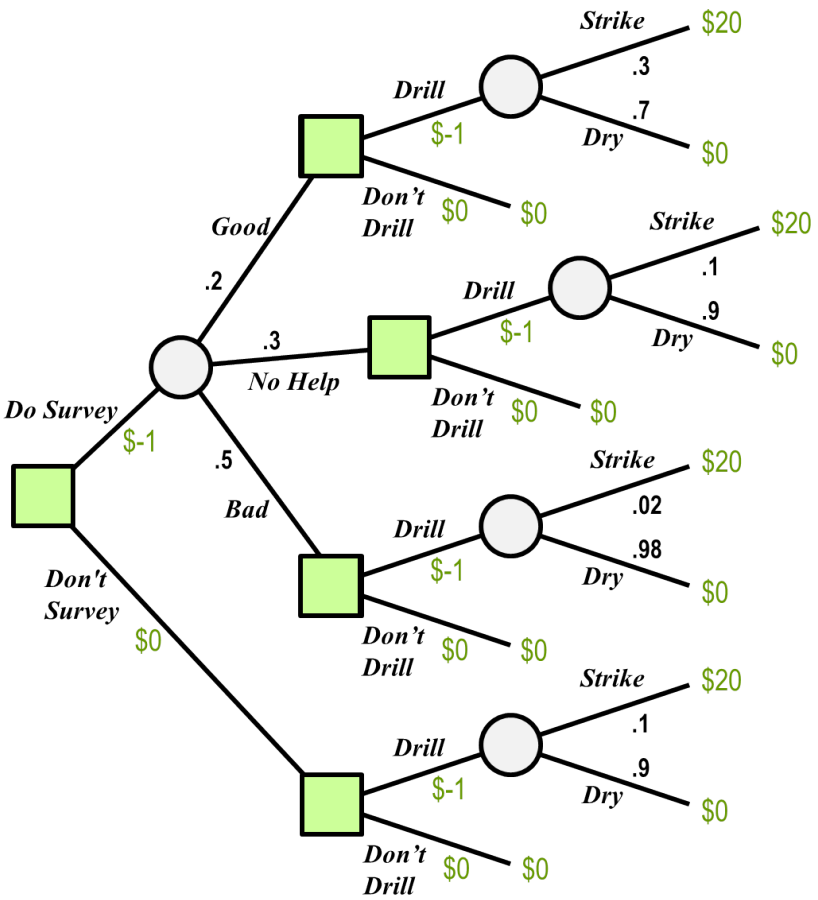
$$P(A > B) > \frac{1}{2}, \quad P(B > C) > \frac{1}{2}, \quad P(C > A) > \frac{1}{2},$$

or, in other words, a cyclic statistical preference.

Outline

- Decision with extreme cases
- Decision with belief functions
- Overview of other (advanced) topics
  - Other decision rules
  - Sequential decisions
  - Decisions with multiple agents
  - Decisions with multiple criteria
  - Decision in combinatorial spaces

Some elements



Some additional issues:

- Dynamical consistency [15]
- Action=policy → potentially huge space to explore

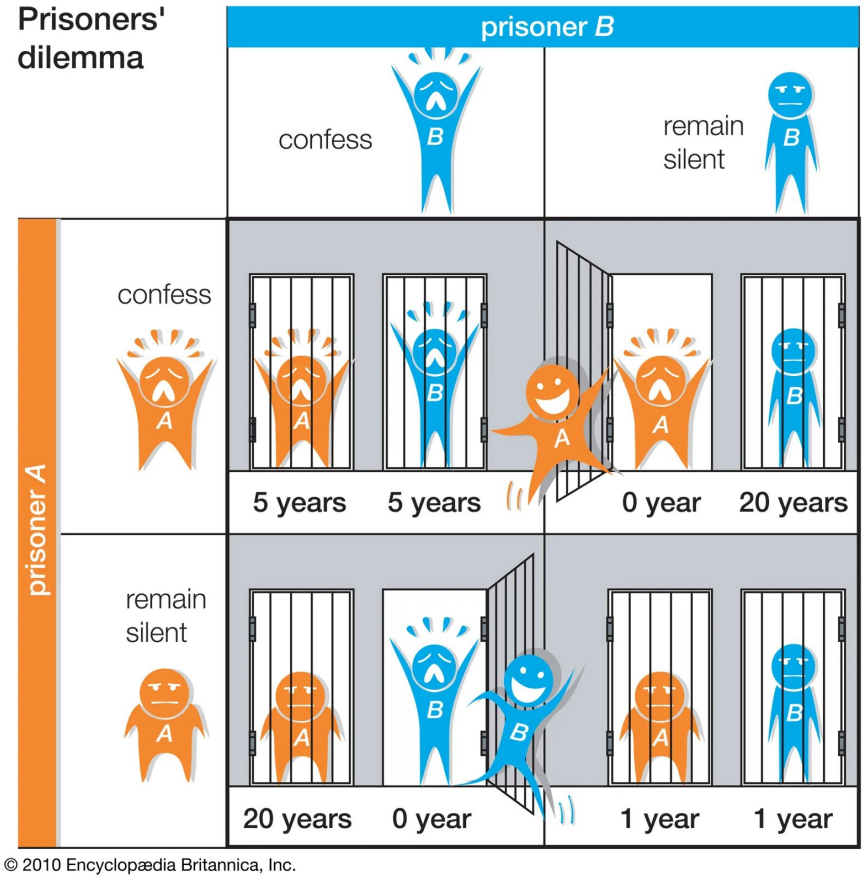
Not that "much" research on belief function + sequential problems, at least algorithmically[a]?

[a] N.B.: Evidential deep reinforcement learning is usually NOT about belief function

Outline

- Decision with extreme cases
- Decision with belief functions
- Overview of other (advanced) topics
  - Other decision rules
  - Sequential decisions
  - Decisions with multiple agents
  - Decisions with multiple criteria
  - Decision in combinatorial spaces

Game-theory and belief functions



Some works touching upon that:

- Generic game-theory [11, 23]
- Coalitional/cooperative games [17]
- Other kind of games (Stackelberg, . . .) → work to be done?

Outline

- Decision with extreme cases
- Decision with belief functions
- Overview of other (advanced) topics
  - Other decision rules
  - Sequential decisions
  - Decisions with multiple agents
  - Decisions with multiple criteria
  - Decision in combinatorial spaces

Multi-criteria decision

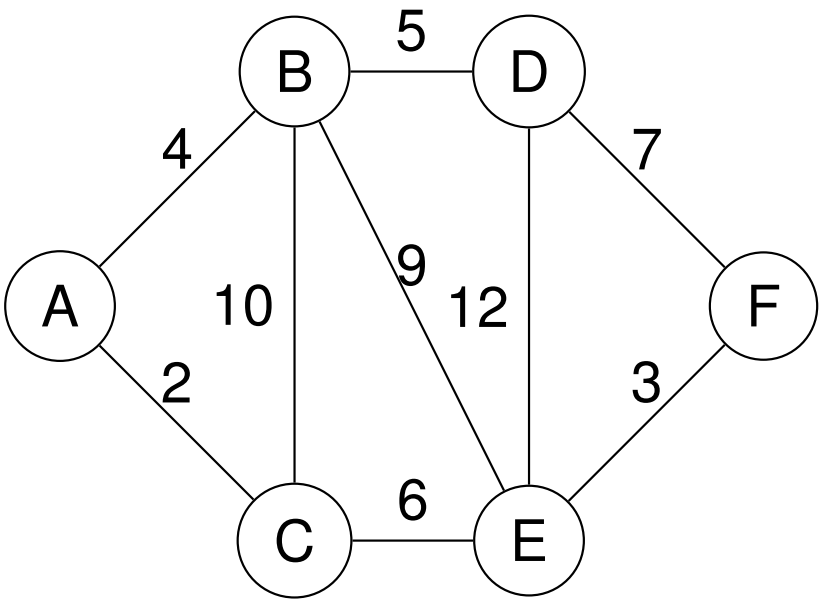
		Non Beneficial	Beneficial	Beneficial	Beneficial
	Attribute Or Criteria	Price or Cost	Storage Space	Camera	Looks
Alternative	Mobile 1	0.8	0.5	0.75	1
	Mobile 2	1	0.5	0.5	0.6
	Mobile 3	0.667	1	1	0.8
	Mobile 4	0.727	1	0.5	0.8
	Mobile 5	0.889	0.5	1	0.4

- Many works on combining BF and TOPSIS, AHP, . . . [14, 10]
- Some generic works on preferences [7, 9]
- Largely unexplored topics: eliciting preference models with uncertain opinions, preference models using logic-based language (CP-nets, weighted logic)
- Also connected to multi-objective optimisation [20]

Outline

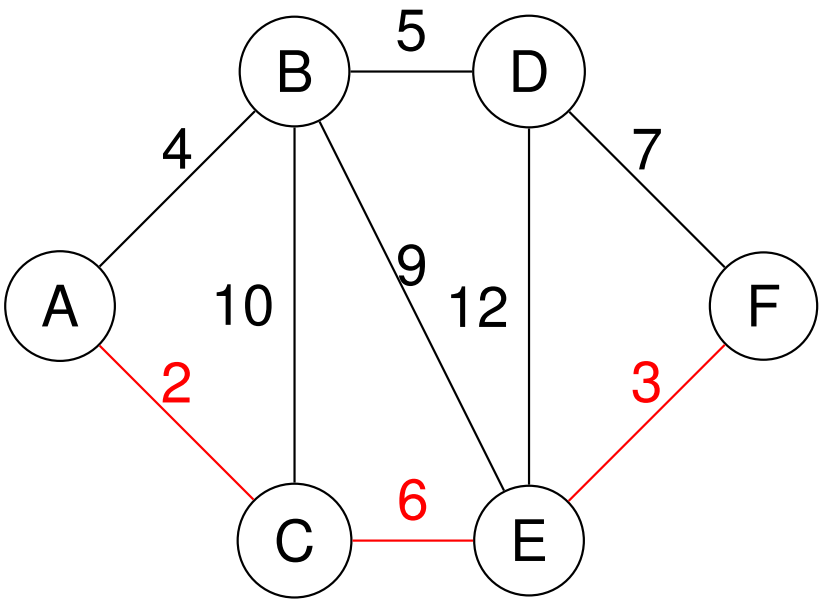
- Decision with extreme cases
- Decision with belief functions
- Overview of other (advanced) topics
  - Other decision rules
  - Sequential decisions
  - Decisions with multiple agents
  - Decisions with multiple criteria
  - Decision in combinatorial spaces

Combinatorial optimisation



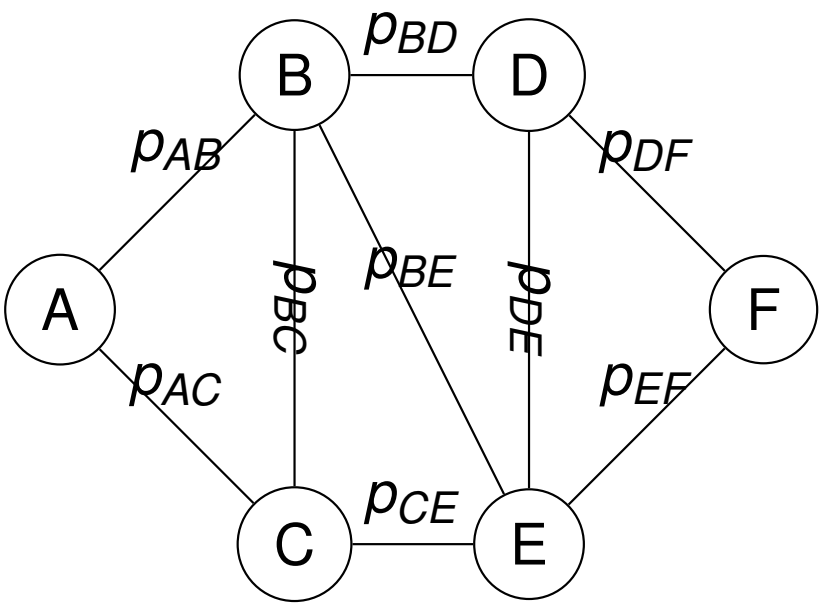
Optimisation and uncertainty

Combinatorial optimisation



Optimisation and uncertainty

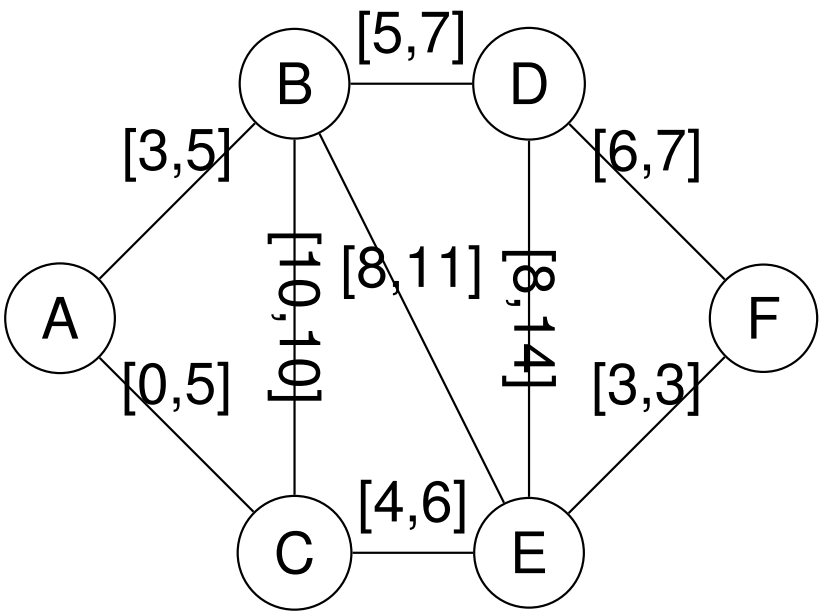
Combinatorial optimisation



Optimisation and uncertainty

- Probabilities=stochastic optimisation

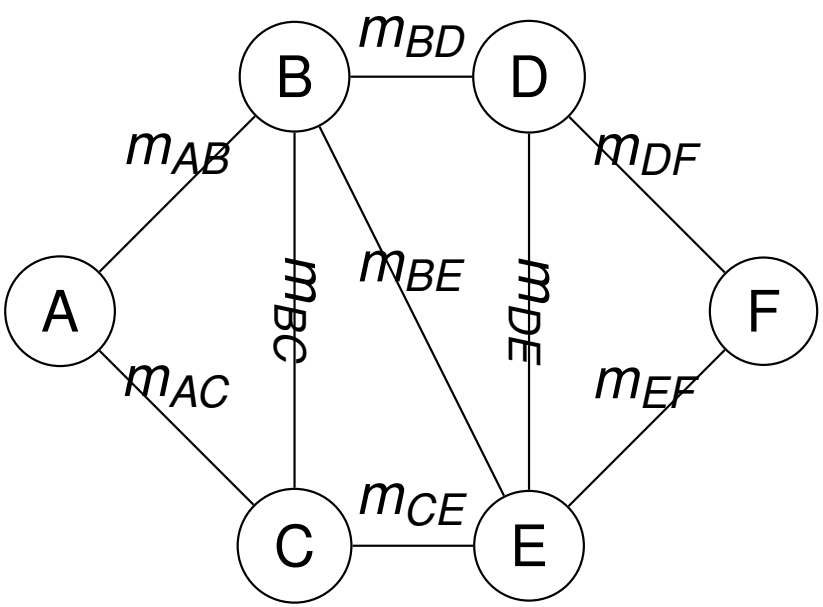
Combinatorial optimisation



Optimisation and uncertainty

- Probabilities=stochastic optimisation
- Interval + Minmax=robust optimisaton

Combinatorial optimisation



- Alternative= one feasible solutions
- In theory, any rule applicable
- But listing alternatives impossible in real-life size problems (combinatorial explosion)

Optimisation and uncertainty

- Probabilities=stochastic optimisation
- Interval + Minmax=robust optimisaton
- Belief functions=evidential optimisation [28, 18, 27]

References I

[1] K. J. Arrow and L. Hurwicz.
An optimality criterion for decision-making under ignorance.
Uncertainty and expectations in economics, pages 1–11, 1972.

[2] I. Couso, S. Moral, and L. Sánchez.
The behavioral meaning of the median.
Information Sciences, 294:127–138, 2015.

[3] I. Couso and L. Sánchez.
Generalized stochastic orderings applied to the study of performance of machine learning algorithms for low quality data.
In *2015 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology (IFSA-EUSFLAT-15)*, pages 1534–1541. Atlantis Press, 2015.

[4] B. De Schuymer, H. De Meyer, B. De Baets, and S. Jenei.
On the cycle-transitivity of the dice model.
Theory and Decision, 54(3):261–285, 2003.

[5] T. Denœux.
Extending stochastic ordering to belief functions on the real line.
Information Sciences, 179(9):1362–1376, 2009.

[6] T. Denoeux.
Decision-making with belief functions: A review.
International Journal of Approximate Reasoning, 109:87–110, 2019.

[7] T. Denœux and M.-H. Masson.
Evidential reasoning in large partially ordered sets: application to multi-label classification, ensemble clustering and preference aggregation.
Annals of Operations Research, 195(1):135–161, 2012.

References II

[8] T. Denoeux and P. P. Shenoy.
An interval-valued utility theory for decision making with dempster-shafer belief functions.
International Journal of Approximate Reasoning, 124:194–216, 2020.

[9] S. Destercke.
A generic framework to include belief functions in preference handling and multi-criteria decision.
International Journal of Approximate Reasoning, 98:62–77, 2018.

[10] J. Dezert, D. Han, and J.-M. Tacnet.
Multi-criteria decision-making with imprecise scores and bf-topsis.
In *2017 20th International Conference on Information Fusion (Fusion)*, pages 1–8. IEEE, 2017.

[11] J. Eichberger and D. Kelsey.
Non-additive beliefs and game theory.
1994.

[12] J. Eichberger and I. Pasichnichenko.
Decision-making with partial information.
Journal of Economic Theory, 198:105369, 2021.

[13] D. Ellsberg.
Risk, ambiguity, and the savage axioms.
The quarterly journal of economics, pages 643–669, 1961.

[14] A. Ennaceur, Z. Elouedi, and E. Lefevre.
Belief ahp method—ahp method with the belief function framework.
International Journal of Information Technology & Decision Making, 15(03):553–573, 2016.

References III

[15] H. Fargier and R. Guillaume.
Decision with belief functions and generalized independence: Two impossibility theorems.
International Journal of Approximate Reasoning, 175:109283, 2024.

[16] B. Finetti.
Theory of probability, volume 1-2.
Wiley, NY, 1974.
Translation of 1970 book.

[17] M. Grabisch and S. Lorenzini.
Bel coalitional games.
In *International Conference on Scalable Uncertainty Management*, pages 186–200. Springer, 2024.

[18] N. Helal, F. Pichon, D. Porumbel, D. Mercier, and É. Lefevre.
The capacitated vehicle routing problem with evidential demands: a belief-constrained programming approach.
In *International Conference On Belief Functions*, pages 212–221. Springer, 2016.

[19] J.-Y. Jaffray.
Linear utility theory for belief functions.
Operations Research Letters, 8(2):107–112, 1989.

[20] P. Limbourg.
Multi-objective optimization of problems with epistemic uncertainty.
In *International conference on evolutionary multi-criterion optimization*, pages 413–427. Springer, 2005.

[21] I. Montes, E. Miranda, and S. Montes.
Decision making with imprecise probabilities and utilities by means of statistical preference and stochastic dominance.
European Journal of Operational Research, 234(1):209–220, 2014.

References IV

[22] I. Montes, E. Miranda, and S. Montes.
Stochastic dominance with imprecise information.
Computational Statistics & Data Analysis, 71:868–886, 2014.

[23] P. Pomeret-Coquot, H. Fargier, and É. Martin-Dorel.
Games of incomplete information: A framework based on belief functions.
International Journal of Approximate Reasoning, 151:182–204, 2022.

[24] L. J. Savage.
The theory of statistical decision.
Journal of the American Statistical association, 46(253):55–67, 1951.

[25] P. Smets.
Decision making in the tbm: the necessity of the pignistic transformation.
International journal of approximate reasoning, 38(2):133–147, 2005.

[26] J. von Neumann and O. Morgenstern.
Theory of Games and Economic Behavior.
Princeton University Press, 1944.

[27] T.-A. Vu, S. Afifi, E. Lefèvre, and F. Pichon.
Optimization problems with evidential linear objective.
International Journal of Approximate Reasoning, 161:108987, 2023.

[28] T.-A. Vu, S. Afifi, E. Lefèvre, and F. Pichon.
Optimization problems with uncertain objective coefficients using capacities.
Annals of Operations Research, 344(1):383–412, 2025.

References V

[29] A. Wald.
Statistical decision functions.
1950.

[30] C. Zhou, B. Qin, D. Li, Z. Xu, and X. Du.
Basic utility theory for belief functions.
In *ECAI*, pages 2648–2655, 2020.