Preference Handling

Alexis Tsoukiàs

LAMSADE - CNRS, Université Paris-Dauphine
tsoukias@lamsade.dauphine.fr
http://www.lamsade.dauphine.fr/∼tsoukias

Napoli 31/05/2011

Outline

1. Problem Setting
2. Basics
3. Preference Learning
4. Preference Modeling
5. Preference Measurement
6. References

Preferences

- Preferences are “rational” desires.
- Preferences are at the basis of any decision aiding activity.
- There are no decisions without preferences.
- Preferences, Values, Objectives, Desires, Utilities, Beliefs, ...

Decision Aiding

A client		An analyst
	A problem situation $\langle \mathcal{A}, \mathcal{O}, \mathcal{S} \rangle$	
	A problem formulation $\langle \mathbb{A}, \mathbb{V}, \Pi \rangle$	
	An evaluation model $\langle A, D, E, H, U, R \rangle$	
	A final recommendation	

Basic information

- *A*: a set of alternatives (enumerative, combinatorial, product space ...)
- *D*: a set of dimensions (attributes) describing *A*.
- *E*: the "scales" used for the attributes in *D*.
- *H*: **Preferential Information**
- *U*: uncertainties ....
- *R*: algorithms, procedures, protocols etc ...

What are the problems?

- How to learn preferences?
- How to model preferences?
- How to aggregate preferences?
- How to use preferences for recommending?

Binary relations

- $\succeq$: binary relation on a set (A).
- $\succeq \subseteq A \times A$ or $A \times P \cup P \times A$.
- $\succeq$ is reflexive.

What is that?

If $x \succeq y$ stands for x is at least as good as y, then the asymmetric part of $\succeq$ ($\succ$: $x \succeq y \ \wedge \ \neg(y \succeq x)$) stands for strict preference. The symmetric part stands for indifference ($\sim_1$: $x \succeq y \ \wedge \ y \succeq x$) or incomparability ($\sim_2$: $\neg(x \succeq y) \ \wedge \ \neg(y \succeq x)$).

More binary relations

- We can further separate the asymmetric (symmetric) part in more relations representing hesitation or intensity of preference.

$$\succ = \succ_1 \cup \succ_2 \cdots \succ_n$$

- We can get rid of the symmetric part since any symmetric relation can be viewed as the union of two asymmetric relations and the identity.
- We can also have valued relations such that: $v(x \succ y) \in [0, 1]$

Binary relations properties

Binary relations have specific properties such as:

- Irreflexive: $\forall x \;\; \neg(x \succ x)$;
- Asymmetric: $\forall x, y \;\; x \succ y \; \rightarrow \; \neg(y \succ x)$;
- Transitive: $\forall x, y, z \;\; x \succ y \; \wedge \; y \succ z \; \rightarrow \; x \succ z$;
- Ferrers; $\forall x, y, z, w \;\; x \succ y \; \wedge \; z \succ w \; \rightarrow \; x \succ w \; \vee \; z \succ y$;

Numbers

$$x \succeq y \quad \Leftrightarrow \quad \Phi(x, y) \geq 0$$

where:

$\Phi : A \times A \mapsto \mathbb{R}$. Simple case $\Phi(x, y) = f(x) - f(y)$; $\;f : A \mapsto \mathbb{R}$

N.B.

Likelihoods can also be expressed under form of binary relations and their numerical representations ($\omega_1 \succeq \omega_2$: event 1 is likely to occur at least as much as event 2).

Numbers

$$x \succeq y \quad \Leftrightarrow \quad \Phi(x, y) \geq 0$$

where:

$\Phi : A \times A \mapsto \mathbb{R}$. Simple case $\Phi(x, y) = f(x) - f(y); \quad f : A \mapsto \mathbb{R}$

N.B.

Likelihoods can also be expressed under form of binary relations and their numerical representations ($\omega_1 \succeq \omega_2$: event 1 is likely to occur at least as much as event 2).

Consider sentences of the type:

- I like red shoes.
- I do not like brown sugar.
- I prefer Obama to McCain.
- I do not want tea with milk.
- Cost is more important than safety.
- I prefer flying to Athens than having a suite at Istanbul.

What do we learn out of such sentences?

- Basic hypotheses about the structure of the evaluation model.
- Binary relations.
- Numerical values (exact or imprecise).
- Importance Parameters.
- Inconsistencies.

Preference Structures

A preference structure

is a collection of binary relations $\sim_1, \cdots \sim_m, \succ_1, \cdots \succ_n$ such that:

- they are pair-disjoint;
- $\sim_1 \cup \cdots \sim_m \cup \succ_1 \cup \cdots \succ_n = A \times A$;
- $\sim_i$ are symmetric and $\succ_j$ are asymmetric;
- possibly they are identified by their properties.

$\sim_1, \sim_2, \succ$ Preference Structures

Independently from the nature of the set A (enumerated, combinatorial etc.), consider $x, y \in A$ as whole elements. Then:

If $\succeq$ is a weak order then:

$\succ$ is a strict partial order, $\sim_1$ is an equivalence relation and $\sim_2$ is empty.

If $\succeq$ is an interval order then:

$\succ$ is a partial order of dimension two, $\sim_1$ is not transitive and $\sim_2$ is empty.

$\sim_1, \sim_2, \succ$ Preference Structures

Independently from the nature of the set A (enumerated, combinatorial etc.), consider $x, y \in A$ as whole elements. Then:

If $\succeq$ is a weak order then:

$\succ$ is a strict partial order, $\sim_1$ is an equivalence relation and $\sim_2$ is empty.

If $\succeq$ is an interval order then:

$\succ$ is a partial order of dimension two, $\sim_1$ is not transitive and $\sim_2$ is empty.

$\sim_1, \sim_2, \succ_1 \succ_2$ Preference Structures

If $\succeq$ is a *PQI* interval order then:

$\succ_1$ is transitive, $\succ_2$ is quasi transitive, $\sim_1$ is asymmetrically transitive and $\sim_2$ is empty.

If $\succeq$ is a pseudo order then:

$\succ_1$ is transitive, $\succ_2$ is quasi transitive, $\sim_1$ is non transitive and $\sim_2$ is empty.

$\sim_1, \sim_2, \succ_1 \succ_2$ Preference Structures

If $\succeq$ is a *PQI* interval order then:

$\succ_1$ is transitive, $\succ_2$ is quasi transitive, $\sim_1$ is asymmetrically transitive and $\sim_2$ is empty.

If $\succeq$ is a pseudo order then:

$\succ_1$ is transitive, $\succ_2$ is quasi transitive, $\sim_1$ is non transitive and $\sim_2$ is empty.

What characterises such structures?

Characteristic Properties

Weak Orders are complete and transitive relations.
Interval Orders are complete and Ferrers relations.

Numerical Representations

w.o. $\Leftrightarrow$ $\exists f : A \mapsto \mathbb{R} : x \succeq y \leftrightarrow f(x) \geq f(y)$
i.o. $\Leftrightarrow$ $\exists f, g : A \mapsto \mathbb{R} : f(x) > g(x); x \succeq y \leftrightarrow f(x) \geq g(y)$

What characterises such structures?

Characteristic Properties

Weak Orders are complete and transitive relations.
Interval Orders are complete and Ferrers relations.

Numerical Representations

w.o. $\Leftrightarrow \quad \exists f : A \mapsto \mathbb{R} : x \succeq y \;\leftrightarrow\; f(x) \geq f(y)$

i.o. $\Leftrightarrow \quad \exists f, g : A \mapsto \mathbb{R} : \; f(x) > g(x); x \succeq y \;\leftrightarrow\; f(x) \geq g(y)$

More about structures

Characteristic Properties

PQI Interval Orders are complete and generalised Ferrers relations.
Pseudo Orders are coherent bi-orders.

Numerical Representations

PQI i.o. $\Leftrightarrow \exists f, g : A \mapsto \mathbb{R} : f(x) > g(x); x \succ_1 y \leftrightarrow g(x) > f(y); x \succ_2 y \leftrightarrow f(x) > f(y) > g(x)$
p.o. $\Leftrightarrow \exists f, t, g : A \mapsto \mathbb{R} : f(x) > t(x) > g(x); x \succ_1 y \leftrightarrow g(x) > f(y); x \succ_2 y \leftrightarrow g(x) > t(y)$

More about structures

Characteristic Properties

PQI Interval Orders are complete and generalised Ferrers relations.
Pseudo Orders are coherent bi-orders.

Numerical Representations

PQI i.o. $\Leftrightarrow \exists f, g : A \mapsto \mathbb{R} : f(x) > g(x); x \succ_1 y \leftrightarrow g(x) > f(y); x \succ_2 y \leftrightarrow f(x) > f(y) > g(x)$
p.o. $\Leftrightarrow \exists f, t, g : A \mapsto \mathbb{R} : f(x) > t(x) > g(x); x \succ_1 y \leftrightarrow g(x) > f(y); x \succ_2 y \leftrightarrow g(x) > t(y)$

What if A is multi-attribute described?

$$x = \langle x_1 \cdots x_n \rangle \quad y = \langle y_1 \cdots y_n \rangle$$

$$x \succeq y \quad \Leftrightarrow \quad \Phi([u_1(x_1) \cdots u_n(_n)], [u_1(y_1) \cdots u_n(y_n)] \geq 0$$

A special case is when Φ is increasing to its first n arguments and decreasing to the following n arguments: it then can be an additive function. See more in conjoint measurement theory.

What if A is multi-attribute described?

$$x = \langle x_1 \cdots x_n \rangle \quad y = \langle y_1 \cdots y_n \rangle$$

$$x \succeq y \quad \Leftrightarrow \quad \Phi([u_1(x_1) \cdots u_n(_n)], [u_1(y_1) \cdots u_n(y_n)] \geq 0$$

A special case is when Φ is increasing to its first n arguments and decreasing to the following n arguments: it then can be an additive function. See more in conjoint measurement theory.

What if A is multi-attribute described?

$$x = \langle x_1 \cdots x_n \rangle \quad y = \langle y_1 \cdots y_n \rangle$$

$$x \succeq y \quad \Leftrightarrow \quad \Phi([u_1(x_1) \cdots u_n(_n)], [u_1(y_1) \cdots u_n(y_n)] \geq 0$$

A special case is when Φ is increasing to its first n arguments and decreasing to the following n arguments: it then can be an additive function. See more in conjoint measurement theory.

What is measuring?

Constructing a function from a set of “objects” to a set of “measures”.

Objects come from the real world.

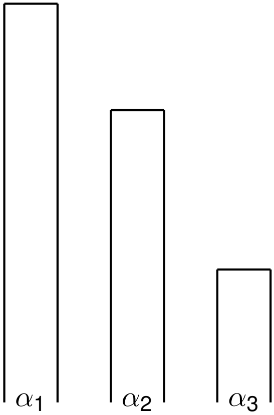
Measures come from empirical observations on some attributes of the objects.

The problem is: how to construct the function out from such observations?

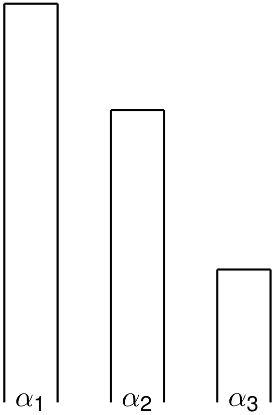
Measurement

1. Real objects $(x, y, \cdots)$.
2. Empirical evidence comparing objects $(x \succeq y, \cdots)$.
3. First numerical representation $(\Phi(x, y) \geq 0)$.
4. Repeat observations in a standard sequence $(x \circ y \succeq z \circ w)$.
5. Enhanced numerical representation $(\Phi(x, y) = \Phi(x) - \Phi(y))$.

Example



Example



$$\alpha_1 \succ \alpha_2 \succ \alpha_3$$

α_1	α_2	α_3
10	8	6
97	32	12
3	2	1

Any of the above could be a numerical representation of this empirical evidence.

<u>Ordinal Scale: any increasing transformation of the numerical representation is compatible with the EE.</u>

Further Example

Consider putting together objects and observing:

$$\alpha_1 \circ \alpha_5 > \alpha_3 \circ \alpha_4 > \alpha_1 \circ \alpha_2 > \alpha_5 > \alpha_4 > \alpha_3 > \alpha_2 > \alpha_1$$

Consider now the following numerical representations:

	L_1	L_2	L_3
α_1	14	10	14
α_2	15	91	16
α_3	20	92	17
α_4	21	93	18
α_5	28	99	29

L_1, L_2 and L_3 capture the simple order among α_{1-5}, but L_2 fails to capture the order among the combinations of objects.

Further Example

NB

For L_1 we get that $\alpha_2 \circ \alpha_3 \sim \alpha_1 \circ \alpha_4$
while for L_3 we get that $\alpha_2 \circ \alpha_3 > \alpha_1 \circ \alpha_4$.
We need to fix a “standard sequence”.

Length

If we fix a “standard” length, a unit of measure, then all objects will be expressed as multiples of that unit.

Further Example

NB

For L_1 we get that $\alpha_2 \circ \alpha_3 \sim \alpha_1 \circ \alpha_4$
while for L_3 we get that $\alpha_2 \circ \alpha_3 > \alpha_1 \circ \alpha_4$.
We need to fix a “standard sequence”.

Length

If we fix a “standard” length, a unit of measure, then all objects will be expressed as multiples of that unit.

Scales

Ratio Scales

All proportional transformations (of the type αx) will deliver the same information. We only fix the unit of measure.

Interval Scales

All affine transformations (of the type $\alpha x + \beta$) will deliver the same information. Besides the unit of measure we fix an origin.

Scales

Ratio Scales

All proportional transformations (of the type αx) will deliver the same information. We only fix the unit of measure.

Interval Scales

All affine transformations (of the type $\alpha x + \beta$) will deliver the same information. Besides the unit of measure we fix an origin.

More complicated

Consider a Multi-attribute space:

$$X = X_1 \times \cdot X_n$$

to each attribute we associate an ordered set of values:

$$X_j = \langle x_j^1 \cdots x_j^m \rangle$$

An object x will thus be a vector:

$$x = \langle x_1^l \cdots x_n^k \rangle$$

Generally speaking ...

$$x \succeq y$$

$$\Longleftrightarrow$$

$$\langle x_1^l \cdots x_n^k \rangle \succeq \langle y_1^i \cdots y_n^j \rangle$$

$$\Longleftrightarrow$$

$$\Phi(f(x_1^l \cdots x_n^k), f(y_1^i \cdots y_n^j)) \geq 0$$

What that means?

	Commuting Time	Clients Exposure	Services	Size	Costs
a	20	70	C	500	1500

What that means?

	Commuting Time	Clients Exposure	Services	Size	Costs
a	20	70	C	500	1500
a_1	25	$70+\delta_1$	C	500	1500

What that means?

	Commuting Time	Clients Exposure	Services	Size	Costs
a	20	70	C	500	1500
a_1	25	$70+\delta_1$	C	500	1500

For what value of δ_1 a and a_1 are indifferent?

What that means?

	Commuting Time	Clients Exposure	Services	Size	Costs
a	20	70	C	500	1500
a_1	25	80	C	500	1500

What that means?

	Commuting Time	Clients Exposure	Services	Size	Costs
a	20	70	C	500	1500
a_1	25	80	C	500	1500
a_2	25	80	C	700	$1500+\delta_2$

What that means?

	Commuting Time	Clients Exposure	Services	Size	Costs
a	20	70	C	500	1500
a_1	25	80	C	500	1500
a_2	25	80	C	700	$1500+\delta_2$

For what value of δ_2 a_1 and a_2 are indifferent?

What that means?

	Commuting Time	Clients Exposure	Services	Size	Costs
a	20	70	C	500	1500
a_1	25	80	C	500	1500
a_2	25	80	C	700	1700

What that means?

	Commuting Time	Clients Exposure	Services	Size	Costs
a	20	70	C	500	1500
a_1	25	80	C	500	1500
a_2	25	80	C	700	1700

The trade-offs introduced with δ_1 and δ_2 allow to get $a \sim a_1 \sim a_2$

What do we get?

Standard Sequences

Length: objects having the same length allow to define a unit of length;

Value: objects being indifferent can be considered as having the same value and thus allow to define a “unit of value”.

Remark 1: indifference is obtained through trade-offs.
Remark 2: separability among attributes is the minimum requirement.

What do we get?

Standard Sequences

Length: objects having the same length allow to define a unit of length;

Value: objects being indifferent can be considered as having the same value and thus allow to define a "unit of value".

Remark 1: indifference is obtained through trade-offs.
Remark 2: separability among attributes is the minimum requirement.

What do we get?

Standard Sequences

Length: objects having the same length allow to define a unit of length;

Value: objects being indifferent can be considered as having the same value and thus allow to define a "unit of value".

Remark 1: indifference is obtained through trade-offs.

Remark 2: separability among attributes is the minimum requirement.

What do we get?

Standard Sequences

Length: objects having the same length allow to define a unit of length;

Value: objects being indifferent can be considered as having the same value and thus allow to define a "unit of value".

Remark 1: indifference is obtained through trade-offs.
Remark 2: separability among attributes is the minimum requirement.

The easy case

IF

1. restricted solvability holds;
2. at least three attributes are essential;
3. $\succeq$ is a weak order satisfying the Archimedean condition $\forall x, y \in \mathbb{R}, \exists n \in \mathbb{N} : ny > x$.

THEN

$$x \succeq y \Leftrightarrow \sum_j u_j(x) \geq \sum_j u_j(y)$$

General Usage

The above ideas apply also in

- Economics (comparison of bundle of goods);
- Decision under uncertainty (comparing consequences under multiple states of the nature);
- Inter-temporal decision (comparing consequences on several time instances);
- Social Fairness (comparing welfare distributions among individuals).

References

- Roberts F.S, *Measurement theory, with applications to Decision Making, Utility and the Social Sciences*, Addison-Wesley, Boston, 1979.
- Roubens M., Vincke Ph., *Preference Modeling*, Springer Verlag, Berlin, 1985.
- Fishburn P.C., *Interval Orders and Interval Graphs*, J. Wiley, New York, 1985.
- Fodor J., Roubens M., *Fuzzy preference modelling and multicriteria decision support*, Kluwer Academic, Dordrecht, 1994.
- Pirlot M., Vincke Ph., *Semi Orders*, Kluwer Academic, Dordrecht, 1997.
- Fishburn P.C., "Preference structures and their numerical representations", *Theoretical Computer Science*, vol. 217, 359-383, 1999.
- Öztürk M., Tsoukiàs A., Vincke Ph., "Preference Modelling", in M. Ehrgott, S. Greco, J. Figueira (eds.), *State of the Art in Multiple Criteria Decision Analysis*, Springer Verlag, Berlin, 27 - 72, 2005.