A brief introduction on Preference Handling

Weiwei Cheng
DKE Group
Uni-Magdeburg
Outline

- Why Preference Handling?
- What is “preference”?
- On the playground of DBS
- On the playground of AI
- End of this presentation
Why Preference Handling?
The **BOOM** of e-services.

The revolution from “**Browser**” to “**Search**”.
The BOOM of e-services.
The revolution from “Browser” to “Search”.

The 3rd quarter of 2006: 27.5 billion $!
New Challenges

How can we enjoy the BOOM?
New Challenges

How can we enjoy the BOOM?

How can we help one attract customers by studying their preferences?
New Challenges

How can we enjoy the BOOM?

How can we help one attract customers by studying their preferences?
What is “preference”?
Partial order:

A (*weak* or *reflexive*) partial order is a binary relation $\leq$ over a set $P$, that for all $a, b$ and $c$ in $P$, we have:

- $a \leq a$ (reflexive)
- If $a \leq b$ and $b \leq a$ then $a = b$ (antisymmetric)
- If $a \leq b$ and $b \leq c$ then $a \leq c$ (transitive)
Preference is ...

Let $P$ be the set of all “packages” of goods and services. Then $\leq$ is a preference relation on $P$ if it is a binary relation on $P$ such that $a \leq b$ if and only if $b$ is at least as preferable as $a$. If $a \leq b$ but not $b \leq a$, then the consumer strictly prefers $b$ to $a$, which is written $a < b$. If $a \leq b$ and $b \leq a$ then the consumer is indifferent between $a$ and $b$. 
Preference is ...

Let $P$ be the set of all "packages" of goods and services. Then $\leq$ is a preference relation on $P$ if it is a binary relation on $P$ such that $a \leq b$ if and only if $b$ is at least as preferable as $a$. If $a \leq b$ but not $b \leq a$, then the consumer strictly prefers $b$ to $a$, which is written $a < b$. If $a \leq b$ and $b \leq a$ then the consumer is indifferent between $a$ and $b$. Preferences of entities are modeled with preference relations.
Preference is...

Let $P$ be the set of all "packages" of goods and services. Then $\leq$ is a preference relation on $P$ if it is a binary relation on $P$ such that $a \leq b$ if and only if $b$ is at least as preferable as $a$. If $a \leq b$ but not $b \leq a$, then the consumer strictly prefers $b$ to $a$, which is written $a < b$. If $a \leq b$ and $b \leq a$ then the consumer is indifferent between $a$ and $b$. Microeconomics read preferences into "choices".
Who are playing the game?
Who are playing the game?

Database System
Who are playing the game?
Who are playing the game?

- Human-computer Interaction
- Database System
- Artificial Intelligence
In computer science, the subfields of preference handling typically *(not only)* include:

- Preference Elicitation
- Preference Representation/Modeling
- Properties and Semantics of Preference
- Preference Management and Repositories
- Algorithms for Preference Handling
- Applications of Preference Handling
On the playground of DBS
Exact world vs. Pref. world

In the “exact world”:

– User’s queries are characterized by “hard constraint”. Wishes can be satisfied completely or not at all.

– A bundle of successful technologies are available, e.g. SQL, E/R-modeling, XML.
Exact world vs. Pref. world

In the “exact world”:

- User’s queries are characterized by a “hard constraint”. Wishes can be satisfied completely or not at all.

- A bundle of successful technologies are available, e.g. SQL, E/R-modeling, XML.
Exact world vs. Pref. world

In the “preference world”:

- No guarantee, that wishes can be always satisfied.

- Compromise is acceptable.
Paradigm Shift

Exact match

Best possible match-making
Paradigm Shift

Treat preferences as *soft constraints*. 

Exact match

Best possible match-making
Kießling’s Preference Model

- Preferences as *partial orders*.
- Unifies non-numerical and numerical ranking
- Features various preference constructors
  - *Pareto accumulation*
  - *Prioritized accumulation*
- Brand new preference algebra

See more on [3]
Relevant Applications

- Preference SQL
- Preference XPATH
- Skyline operator
An Example

The user says:

“I want to buy an Opel, the most important properties are category, price and power. It should be a roadster. If no roadster is available, any other type is acceptable except passenger. The price should around €40000 and the power should be maximized. When these requests can be satisfied, the red color is preferred. At last, the minimal mileage should be applied sequentially to the result.”
An Example

The user says:

“I want to buy an Opel, the most important properties are category, price and power. It should be a roadster. If no roadster is available, any other type is acceptable except passenger. The price should around 15000 and the power should be maximized. When these requests can be satisfied, the red color is preferred. At last, the minimal mileage should be applied sequentially to the result.”
Level 1.

Level 2.  $\sim$40000\€  ✓ Roadster  ❌ Passenger  🤚 Power

Level 3.  Color RED

Level 4.  🕵 Mileage
SELECT * FROM car
WHERE mark = ‘Opel’
PREFERRING (category = ‘Roadster’ ELSE category <> ‘Passenger’ AND price AROUND 40000 AND HIGHEST(power))
CASCADE color = ‘red’
CASCADE LOWEST(mileage);
SELECT * FROM car
WHERE mark = 'Opel'
PREFERRING (category = 'Roadster' ELSE category <> 'Passenger' AND price AROUND 40000 AND HIGHEST(power))
CASCADE color = 'red'
CASCADE LOWEST(mileage);

- SQL is upgraded by **PREFERRING** clause.
SELECT * FROM car
WHERE mark = 'Opel'

PREFERRING (category = 'Roadster' ELSE category <> 'Passenger') AND price AROUND 40000
AND HIGHEST(power))
CASCADE color = 'red'
CASCADE LOWEST(mileage);

- SQL is upgraded by `PREFERRING` clause.
- AND Pareto accumulation.
SELECT * FROM car
WHERE mark = 'Opel'

**PREFERRING** (category = 'Roadster' ELSE category <> 'Passenger' **AND**
price AROUND 40000 **AND** HIGHEST(power))
**CASCADE** color = 'red'
**CASCADE** LOWEST(mileage);

- SQL is upgraded by **PREFERRING** clause.
- **AND** ➔ Pareto accumulation.
- **CASCADE** ➔ Prioritized accumulation.
Preference XPATH

Q1: /cars/car #[@fuel_economy] maximal and (@horsepower) maximal]#

Q2: /cars/car #[@color] in ("black","white") prior to (@price) around 10000]# #[@mileage] minimal]#

Skyline operator

SELECT * FROM Hotels
WHERE city = 'Marburg'
SKYLINE OF price MIN, distance MIN;
Preference XPATH

Q1: /cars/car #[@fuel_economy] maximal and (@horsepower) maximal#

Q2: /cars/car #[@color] in ("black","white") prior to (@price) around 10000# #[@mileage] minimal#

Skyline operator

SELECT * FROM Hotels
WHERE city = ‘Marburg’
SKYLINE OF price MIN, distance MIN;
On the playground of AI
On the playground of AI

- Behavior of a rationally acting agent is always driven by a underlying preference model.

- The task of AI: to provide a recommending decision, which reflect the preferences properly.
1. Collect and aggregate
2. Classify and predict
CP semantics & CP-net

- Totalitarian semantics vs. CP semantics.

- CP (Ceteris paribus) means all else being equal.
I want a TV, a cheap one...

Sure. Here we go...

Faint...

$ $ &*@$%$
I meant any cheaper TV is better, when its other characteristics are not significantly different!

Oh, that is CP semantics. Do it again in CP.

That is much better.
A more concrete example

<table>
<thead>
<tr>
<th></th>
<th>Brand</th>
<th>Ext-color</th>
<th>Int-color</th>
</tr>
</thead>
<tbody>
<tr>
<td>t₁</td>
<td>Mini</td>
<td>red</td>
<td>bright</td>
</tr>
<tr>
<td>t₂</td>
<td>Mini</td>
<td>red</td>
<td>dark</td>
</tr>
<tr>
<td>t₃</td>
<td>Mini</td>
<td>white</td>
<td>bright</td>
</tr>
<tr>
<td>t₄</td>
<td>Mini</td>
<td>white</td>
<td>dark</td>
</tr>
<tr>
<td>t₅</td>
<td>Smart</td>
<td>red</td>
<td>bright</td>
</tr>
<tr>
<td>t₆</td>
<td>Smart</td>
<td>red</td>
<td>dark</td>
</tr>
<tr>
<td>t₇</td>
<td>Smart</td>
<td>white</td>
<td>bright</td>
</tr>
<tr>
<td>t₈</td>
<td>Smart</td>
<td>white</td>
<td>dark</td>
</tr>
</tbody>
</table>

I prefer red Mini to white Smart.

Under *totalitarian* semantics \{ t₇<t₁, t₇<t₂, t₈<t₁, t₈<t₂ \}

Under *CP* semantics \{ t₇<t₁, t₈<t₂ \}
Vantage

**Totalitarian**
- Fewer “optimal” tuples
- Attractive computational properties.
- Implicitly favored by DB community

**CP**
- More common to the nature of preferences.
- Uniformly favored by philosophers, economists and AI researchers
CP-nets

- Boutilier et al. introduce CP-nets in [11]

- A qualitative graphical representing and reasoning tool of preferences under CP semantics.

- Acyclic CP-nets always induce strict partial preference orders.
I prefer red Mini to white Mini.

I prefer white Smart to red Smart.

In white cars I prefer a dark interior.

In red cars I prefer a bright interior.

I prefer Mini to Smart.
<table>
<thead>
<tr>
<th>Brand</th>
<th>Ext-color</th>
<th>Int-color</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$</td>
<td>Mini</td>
<td>red</td>
</tr>
<tr>
<td>$t_2$</td>
<td>Mini</td>
<td>red</td>
</tr>
<tr>
<td>$t_3$</td>
<td>Mini</td>
<td>white</td>
</tr>
<tr>
<td>$t_4$</td>
<td>Mini</td>
<td>white</td>
</tr>
<tr>
<td>$t_5$</td>
<td>Smart</td>
<td>red</td>
</tr>
<tr>
<td>$t_6$</td>
<td>Smart</td>
<td>red</td>
</tr>
<tr>
<td>$t_7$</td>
<td>Smart</td>
<td>white</td>
</tr>
<tr>
<td>$t_8$</td>
<td>Smart</td>
<td>white</td>
</tr>
</tbody>
</table>

S1 I prefer red Mini to white Mini.
S2 I prefer white Smart to red Smart.
S3 In white cars I prefer a dark interior.
S4 In red cars I prefer a bright interior.
S5 **I prefer Mini to Smart.**
S1 I prefer red Mini to white Mini.
S2 I prefer white Smart to red Smart.
S3 In white cars I prefer a dark interior.
S4 In red cars I prefer a bright interior.
S5 I prefer Mini to Smart.

<table>
<thead>
<tr>
<th></th>
<th>Brand</th>
<th>Ext-color</th>
<th>Int-color</th>
</tr>
</thead>
<tbody>
<tr>
<td>t₁</td>
<td>Mini</td>
<td>red</td>
<td>bright</td>
</tr>
<tr>
<td>t₂</td>
<td>Mini</td>
<td>red</td>
<td>dark</td>
</tr>
<tr>
<td>t₃</td>
<td>Mini</td>
<td>white</td>
<td>bright</td>
</tr>
<tr>
<td>t₄</td>
<td>Mini</td>
<td>white</td>
<td>dark</td>
</tr>
<tr>
<td>t₅</td>
<td>Smart</td>
<td>red</td>
<td>bright</td>
</tr>
<tr>
<td>t₆</td>
<td>Smart</td>
<td>red</td>
<td>dark</td>
</tr>
<tr>
<td>t₇</td>
<td>Smart</td>
<td>white</td>
<td>bright</td>
</tr>
<tr>
<td>t₈</td>
<td>Smart</td>
<td>white</td>
<td>dark</td>
</tr>
</tbody>
</table>
S1 I prefer red Mini to white Mini.
S2 I prefer white Smart to red Smart.
S3 In white cars I prefer a dark interior.
S4 In red cars I prefer a bright interior.
S5 I prefer Mini to Smart.

<table>
<thead>
<tr>
<th>t</th>
<th>Brand</th>
<th>Ext-color</th>
<th>Int-color</th>
</tr>
</thead>
<tbody>
<tr>
<td>t1</td>
<td>Mini</td>
<td>red</td>
<td>bright</td>
</tr>
<tr>
<td>t2</td>
<td>Mini</td>
<td>red</td>
<td>dark</td>
</tr>
<tr>
<td>t3</td>
<td>Mini</td>
<td>white</td>
<td>bright</td>
</tr>
<tr>
<td>t4</td>
<td>Mini</td>
<td>white</td>
<td>dark</td>
</tr>
<tr>
<td>t5</td>
<td>Smart</td>
<td>red</td>
<td>bright</td>
</tr>
<tr>
<td>t6</td>
<td>Smart</td>
<td>red</td>
<td>dark</td>
</tr>
<tr>
<td>t7</td>
<td>Smart</td>
<td>white</td>
<td>bright</td>
</tr>
<tr>
<td>t8</td>
<td>Smart</td>
<td>white</td>
<td>dark</td>
</tr>
</tbody>
</table>
BSmart < BMini
BMini : Ew < Er
BSmart < Er
BMini : Ew < Eb
Ew : Ib < Id
Er : Id < Ib
So far so good...

Two vexed problems on machine learning

1. Black box property
2. Data collecting vs. user’s privacy.
Then A Miracle Occurs (Black box)

- Little interrelation between input and output.
- Serious problem in certain area.

“I think you should be more explicit in step two.”
Data collecting vs. User’s privacy

- User sometimes is against the data collecting.
- Tradeoff between performance and privacy must be made.

“Nobody knows you’re a dog on Internet.”
Almost the END
We are 12 billion light-years from the edge.

That's a guess — no-one can ever say it's true.

But I know that I will always be with you.
We are 13.7 billion light-years from the edge of the observable universe.

That's a good estimate with well-defined error bars and with the available information,

I predict that I will always be with you.
References

9. R.Brafman, C.Domshlak. *Database Preference Queries Revisited (Extended abstract)*.