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*".

New Challenges

?

How can we enjoy the BOOM?

New Challenges

How can we enjoy the BOOM?

How can we help one attract customers by study their preferences?

New Challenges

How can we enjoy the BOOM?

How can we help one attract customers by study their preferences?

What is “preference”?

Just in Case…

Partial order:

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

- $a \triangleleft a$ (reflexive)
- If $a \triangleleft b$ and $b \triangleleft a$ then $a=b$ (antisymmetric)
- If $a \triangleleft b$ and $b \triangleleft c$ then $a \triangleleft c$ (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 …

Preferences of entities are modeled with preference relations.

L… ll "packages" of goods a… a preference re… binary relation on P … d only if b is at least as prefera… 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 …

I… s" of goods
a…
re… on on P
… s at least as
prefera… $b \le a$, then
the con…umer s… prefers b to a, which
is written $a < b$. If $a \le b$ and $b \le a$ then the
consumer is indifferent between a and b.

Prefer…
are …
prefere…

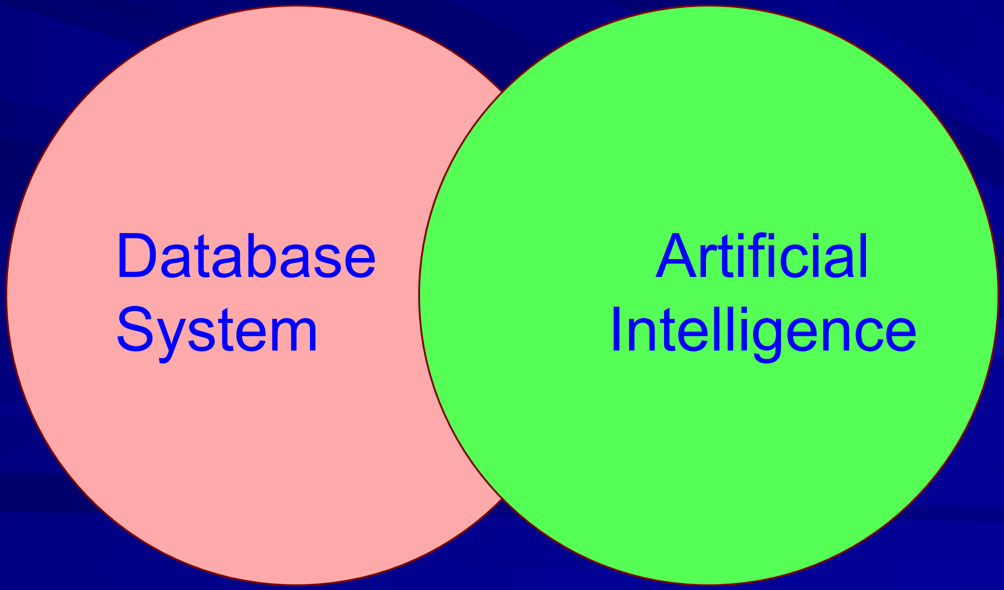
Microeconomics read
preferences into
"**choices**".

Who are playing the game?

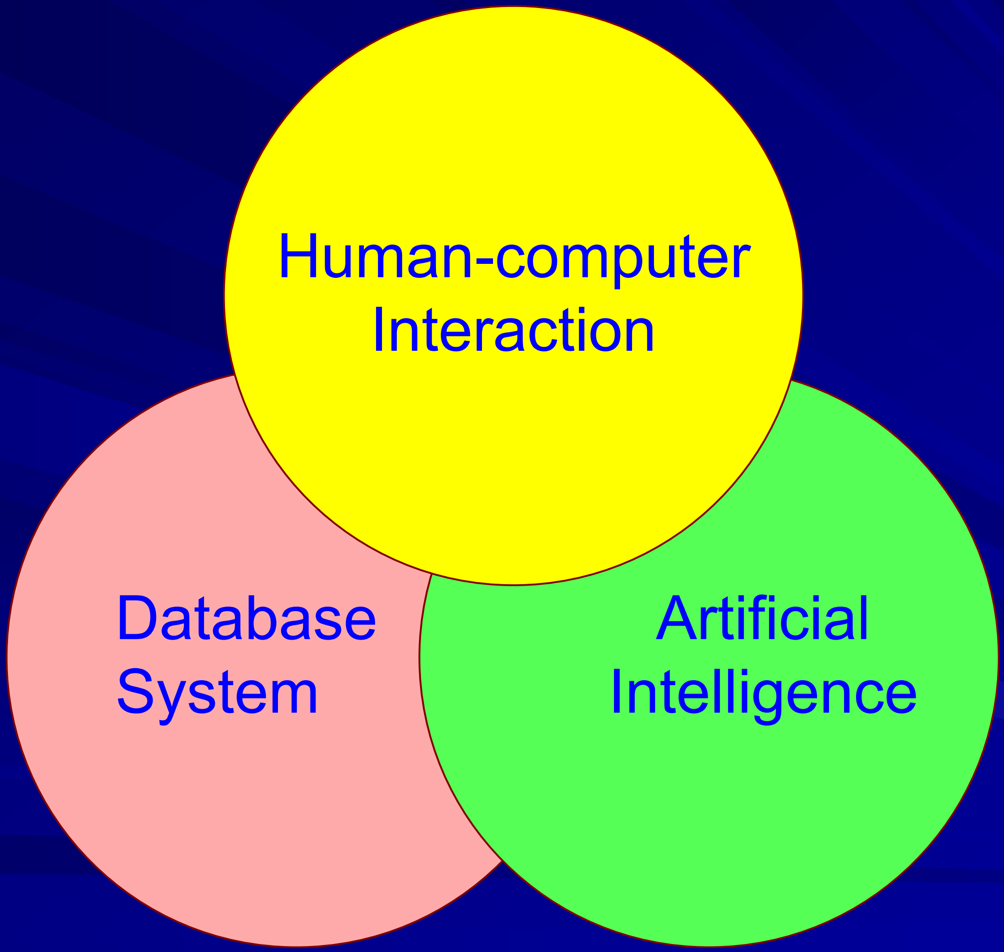
Who are playing the game?

Database
System

Who are playing the game?



Who are playing the game?



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 ch "hard constraint". W fied *completely* or *not at all*.

DBS still lives here ☺

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

Exact world vs. Pref. world

In the "*preference world*":

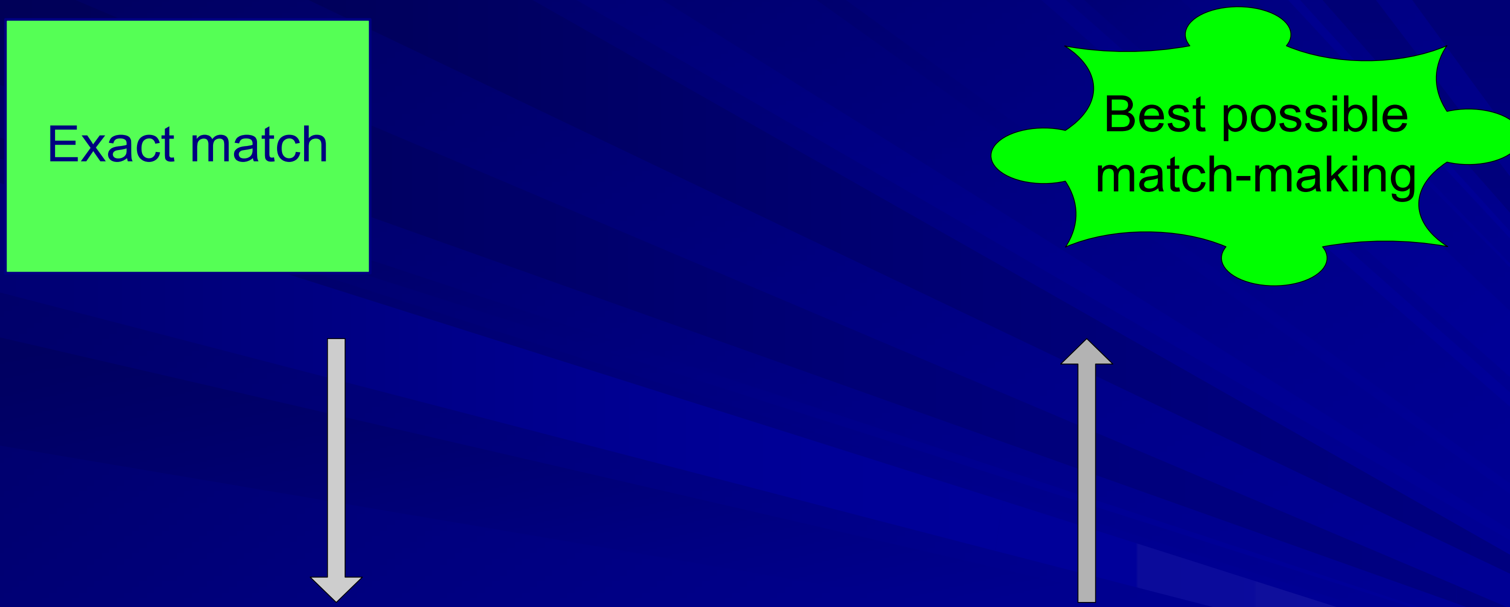
- No guarantee, that wishes can be always satisfied.
- Compromise is acceptable.

Paradigm Shift

Exact match

Best possible match-making

Paradigm Shift



Treat preferences as *soft constraints*.

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 O… …rtant properties are categ… …n. It should be a roadste… …s available, any other type is a… …assenger. The price should around …000 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."

Stop Reading!

Level 1.

Level 2.

~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**

  A restricted form of Pareto accumulation

  SELECT * FROM Hotels

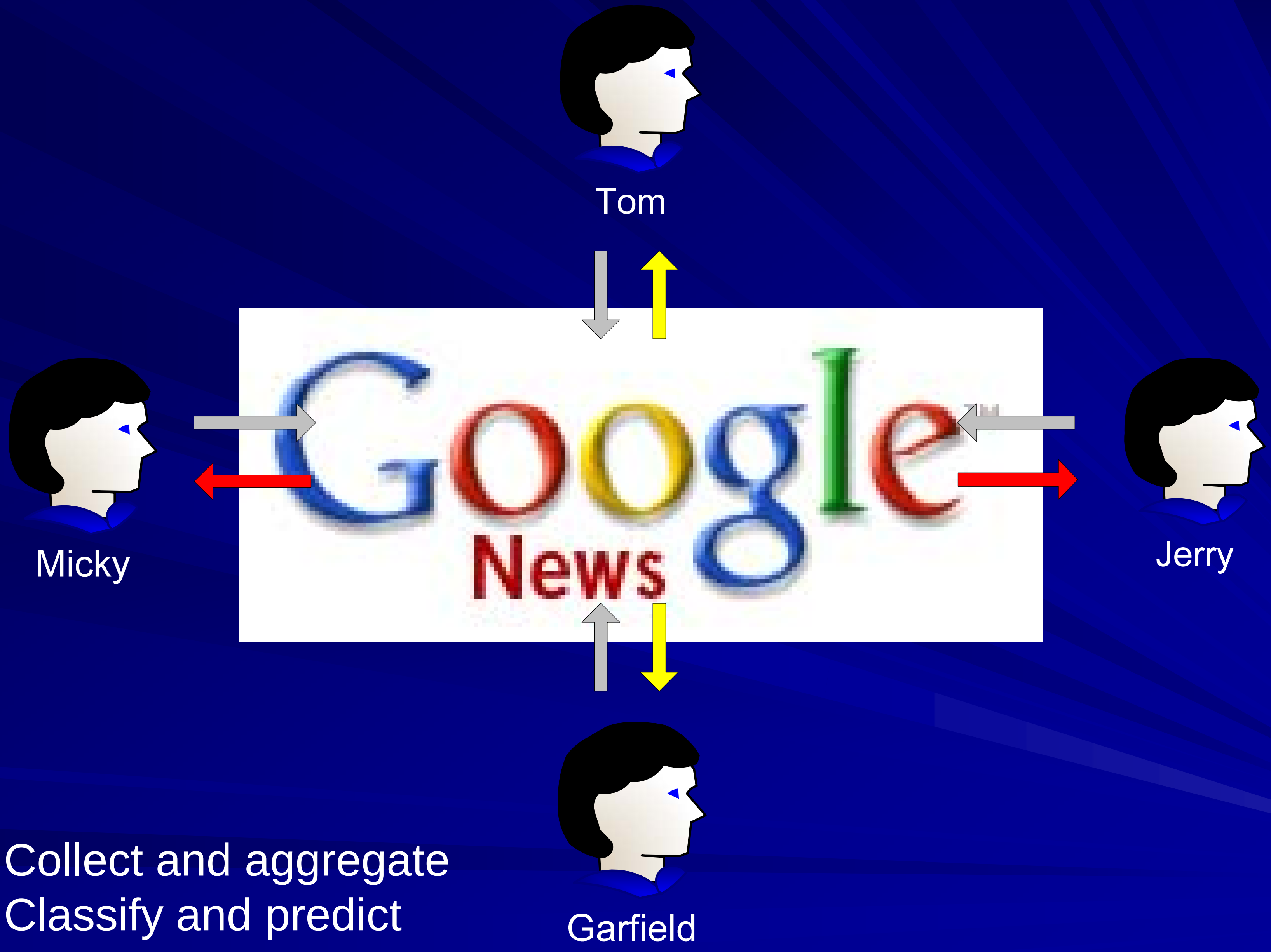
  WHERE city = ‘Marburg’

  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

	Brand	Ext-color	Int-color
t_1	Mini	red	bright
t_2	Mini	red	dark
t_3	Mini	white	bright
t_4	Mini	white	dark
t_5	Smart	red	bright
t_6	Smart	red	dark
t_7	Smart	white	bright
t_8	Smart	white	dark

I prefer red Mini to white Smart.

Under *totalitarian* semantics { $t_7<t_1$, $t_7<t_2$, $t_8<t_1$, $t_8<t_2$ }

Under *CP* semantics { $t_7<t_1$, $t_8<t_2$ }

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

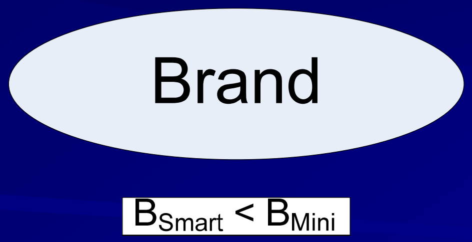
- Boutilier et al. introduce CP-nets in [11]
- A qualitative graphical *representing* and *reasoning* tool of preferences under CP semantics.
- Acyclic CP-nets always induce strict partial preference orders.

	Brand	Ext-color	Int-color
t_1	Mini	red	bright
t_2	Mini	red	dark
t_3	Mini	white	bright
t_4	Mini	white	dark
t_5	Smart	red	bright
t_6	Smart	red	dark
t_7	Smart	white	bright
t_8	Smart	white	dark

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.

	Brand	Ext-color	Int-color
t_1	Mini	red	bright
t_2	Mini	red	dark
t_3	Mini	white	bright
t_4	Mini	white	dark
t_5	Smart	red	bright
t_6	Smart	red	dark
t_7	Smart	white	bright
t_8	Smart	white	dark

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.



	Brand	Ext-color	Int-color
t_1	Mini	red	bright
t_2	Mini	red	dark
t_3	Mini	white	bright
t_4	Mini	white	dark
t_5	Smart	red	bright
t_6	Smart	red	dark
t_7	Smart	white	bright
t_8	Smart	white	dark

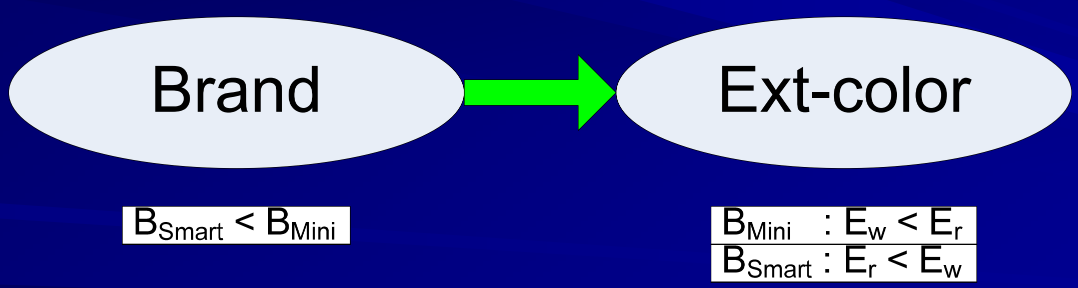
S1 I prefer red Mini to white Mini.

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.



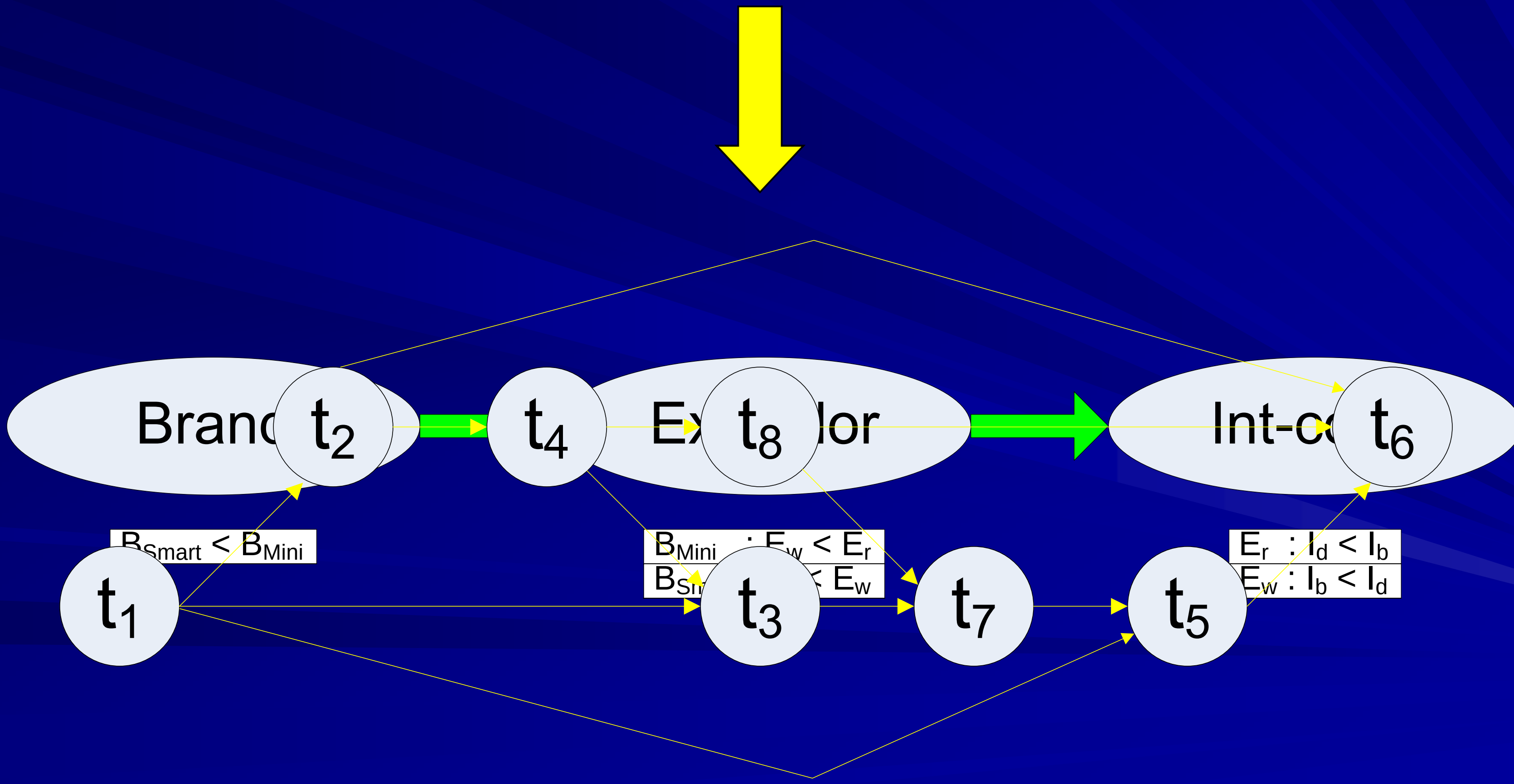
	Brand	Ext-color	Int-color
t_1	Mini	red	bright
t_2	Mini	red	dark
t_3	Mini	white	bright
t_4	Mini	white	dark
t_5	Smart	red	bright
t_6	Smart	red	dark
t_7	Smart	white	bright
t_8	Smart	white	dark

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

$B_{Smart} < B_{Mini}$

$B_{Mini} : E_w < E_r$
$B_{Smart} : E_r < E_w$

$E_r : I_d < I_b$
$E_w : I_b < I_d$


t1
t2
t3
t4
t5
t6
t7
t8
B_Smart < B_Mini
E_r : I_d < I_b
E_w : I_b < I_d

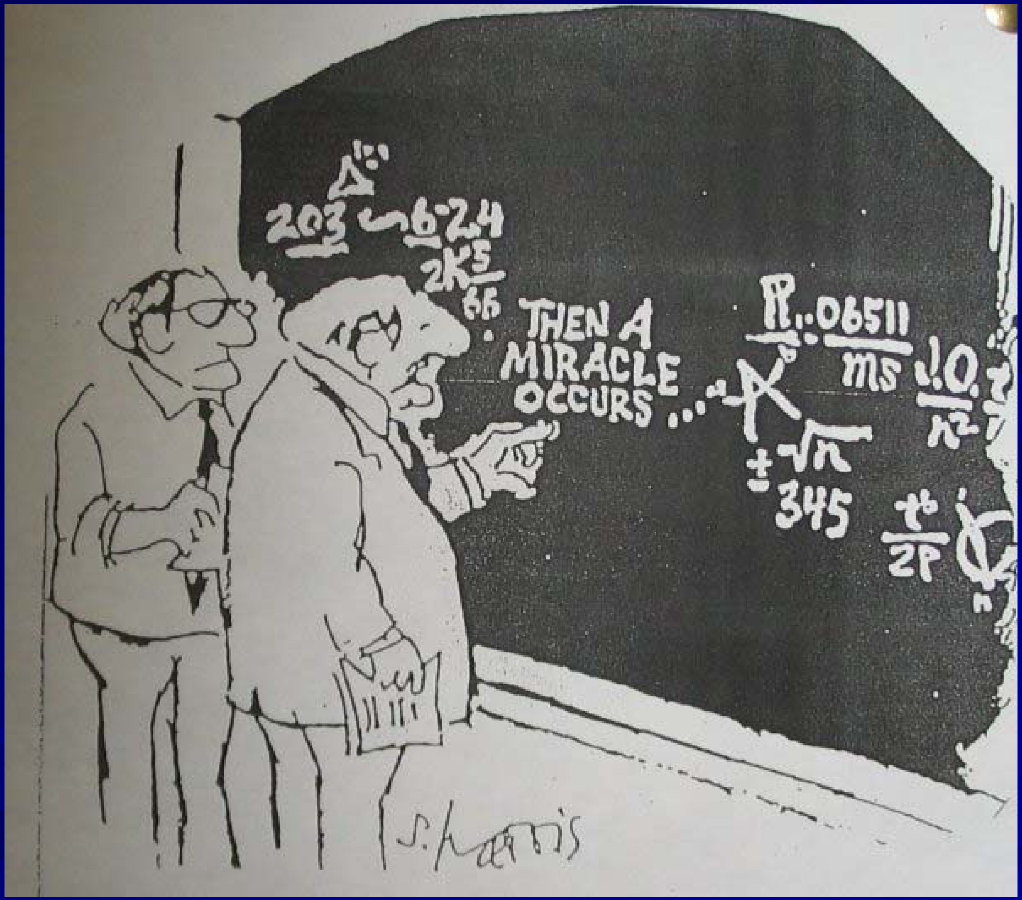
So far so good…

Two vexed problems on machine learning

1. Black box property

2. Data collecting vs. user’s privacy.

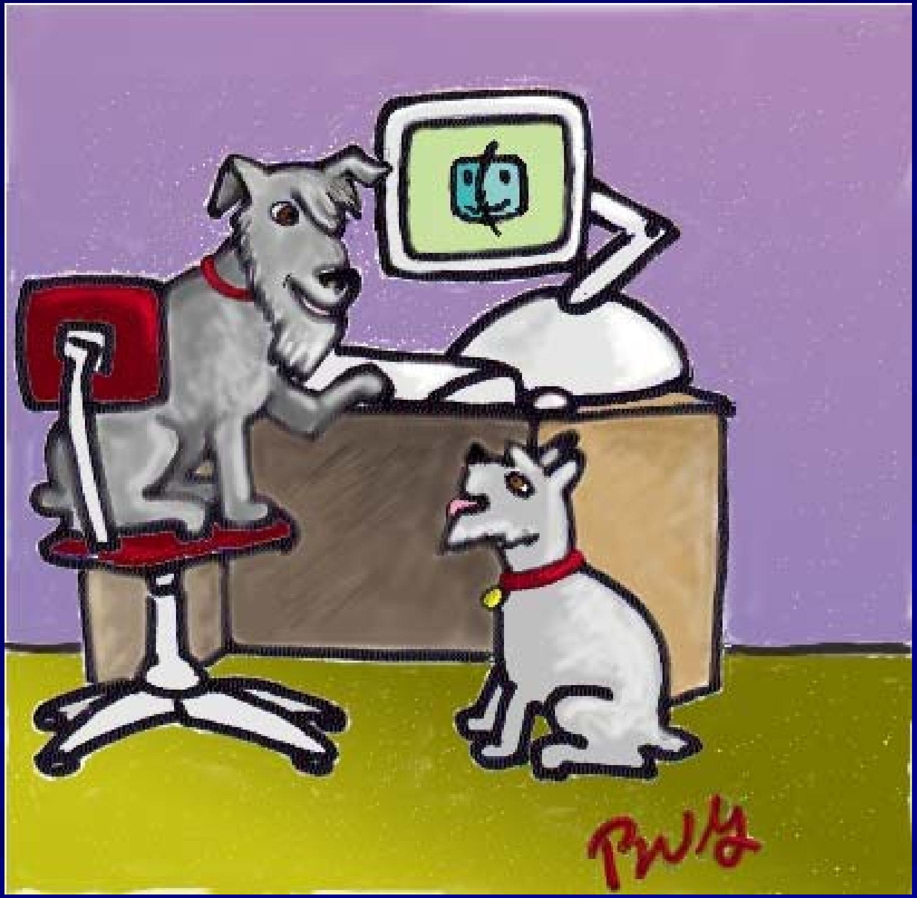
Then A Miracle Occurs (Black box)



“I think you should be more explicit in step two.”

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

Data collecting vs. User's privacy



"Nobody knows you're a dog on Internet."

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

Almost the END

Nine Million Bicycles (origin version)

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.

Nine Million Bicycles (scientific version!!)

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

1. W. Cheng. *A Brief Introduction on Preference Handling*. 2007.
2. U.S.Department of Commerce. *U.S.Census Bureau News CB06-167*. 2006.
3. W.Kießling. *Foundations of Preferences in Database Systems*. 2001.
4. W.Kießling, G.Köstler. *Preference SQL – Design, Implementation, Experiences*. 1999.
5. W.Kießling, B.Hafenrichter, S.Fischer, S.Holland. *Preference XPath: A Query Language for E-Commerce*. 2001.
6. S.Borzsonyi, D.Kossmann, K.Stocker. *The Skyline Operator*. 2001.
7. J.Fürnkranz, E.Hüllermeier. *Preference Learning*. 2005
8. Google News. *http://news.google.com/*
9. R.Brafman, C.Domshlak. *Database Preference Queries Revisited (Extended abstract).*
10. S.Hansson. *What is Ceteris Paribus Preference*. 1996.
11. C.Boutilier, R.Brafman, C.Domshlak, H.Hoos, D.Poole. *CP-nets: A Tool for Representing and Reasoning with Conditional Ceteris Paribus Preference Statements*. 2004.
12. Katie Melua on Wikipedia. *http://en.wikipedia.org/wiki/Katie_Melua*