Algebraic optimization of relational queries with various kinds of preference

Radim Nedbal

Czech Technical University in Prague,
Faculty of Nuclear Sciences and Physical Engineering
&
Institute of Computer Science
Academy of Sciences of the Czech Republic

Nový Smokovec, 21st January 2008

The importance of preference

- Preference is ubiquitous in everyday life.
- Preference has been studied by various scientific communities: psychologists, philosophers, economists, logicians: von Wright (1963)
- Their importance has been addressed by DB community:
  - Lacroix and Lavency (1987)
  - Börzsönyi, Kossmann, Stocker: Skyline queries
  - Chomicki (Buffalo): project 'Preference Queries' (2003-2008), Kießling (Augsburg): program 'It's a Preference World'

Preference – a soft constraint

$$A >_{\text{pref}} B$$

"I like A better than B".

- NOT a hard constraint – a personalized wish
- may come from different, even conflicting sources, may be very complex
  - NOT necessarily a total order – incomparable items (conflict, missing information?)
  - "better than" can be defined quantitatively or qualitatively

Examples of $A >_{\text{pref}} B$ ("I like A better than B.")

- $A :=$ *playing tenis*
- $B :=$ *playing golf*

Do I like playing tennis in the rain better than playing golf on a sunny day?

Examples of $A >_{pref} B$ ("I like A better than B.")

- $A :=$ *playing tenis*
- $B :=$ *playing golf*

Do I like playing tennis in the rain better than playing golf on a sunny day?

- $A :=$ *water-skiing*
- $B :=$ *skiing*

- in winter or in summer??
- water-skiing in summer and skiing in winter?

CEP and the holistic nature of preference

Conjunctive expansion principle

When I have neither A or B, I favor an acquisition of A over B. Similarly, if I have both A an B, I favor loosing B over loosing A.

$$A >_{\text{pref}} B \equiv A \wedge \neg B >_{\text{pref}} \neg A \wedge B \quad .$$

CEP and the holistic nature of preference

Conjunctive expansion principle

When I have neither A or B, I favor an acquisition of A over B. Similarly, if I have both A an B, I favor loosing B over loosing A.

$$A >_{\text{pref}} B \equiv A \wedge \neg B >_{\text{pref}} \neg A \wedge B \quad .$$

A ***world*** w **(of states of affairs S):** $w \in W = 2^S$

$w = \{$playing tenis, not playing golf, sunny day, not rainy day$\}$ is a possible world.

CEP and the holistic nature of preference

Conjunctive expansion principle

When I have neither A or B, I favor an acquisition of A over B. Similarly, if I have both A an B, I favor loosing B over loosing A.

$$A >_{\text{pref}} B \equiv A \wedge \neg B >_{\text{pref}} \neg A \wedge B \quad .$$

A ***world*** w (of states of affairs S): $\quad w \in W = 2^S$

$w = \{$playing tenis, not playing golf, sunny day, not rainy day$\}$ is a possible world.

$A >_{\text{pref}} B \Rightarrow$ preference over worlds

$$A >_{\text{pref}} B \Rightarrow w_{A \wedge \neg B} >_{\text{pref}} w_{\neg A \wedge B} \quad .$$

Transferring preference from $A >_{\text{pref}} B$ to worlds

$A >_{\text{pref}} B \Rightarrow \forall w_{A\wedge\neg B} \forall w_{\neg A\wedge B} : w_{A\wedge\neg B} >_{\text{pref}} w_{\neg A\wedge B}$

- $\forall w_{A\wedge\neg B}\ \forall w_{\neg A\wedge B} : \quad w_{A\wedge\neg B} >_{\text{pref}} w_{\neg A\wedge B}$
- $\exists w_{A\wedge\neg B}\ \exists w_{\neg A\wedge B} : \quad w_{A\wedge\neg B} >_{\text{pref}} w_{\neg A\wedge B}$
- $\exists w_{A\wedge\neg B}\ \forall w_{\neg A\wedge B} : \quad w_{A\wedge\neg B} >_{\text{pref}} w_{\neg A\wedge B}$
- $\forall w_{A\wedge\neg B}\ \exists w_{\neg A\wedge B} : \quad w_{A\wedge\neg B} >_{\text{pref}} w_{\neg A\wedge B}$

Transferring preference from $A >_{pref} B$ to worlds

$A >_{pref} B \Rightarrow \forall w_{A \wedge \neg B} \forall w_{\neg A \wedge B} : w_{A \wedge \neg B} >_{pref} w_{\neg A \wedge B}$

- $\forall w_{A \wedge \neg B}\ \forall w_{\neg A \wedge B} : \quad w_{A \wedge \neg B} >_{pref} w_{\neg A \wedge B}$
- $\exists w_{A \wedge \neg B}\ \exists w_{\neg A \wedge B} : \quad w_{A \wedge \neg B} >_{pref} w_{\neg A \wedge B}$
- $\exists w_{A \wedge \neg B}\ \forall w_{\neg A \wedge B} : \quad w_{A \wedge \neg B} >_{pref} w_{\neg A \wedge B}$
- $\forall w_{A \wedge \neg B}\ \exists w_{\neg A \wedge B} : \quad w_{A \wedge \neg B} >_{pref} w_{\neg A \wedge B}$

$A >_{pref} B \Rightarrow w_{A \wedge \neg B}, w_{\neg A \wedge B} \in W_i : w_{A \wedge \neg B} >_{pref} w_{\neg A \wedge B}$

Does it make sense to compare all worlds?
..contextual equivalence classes $W/\equiv\ = \{W_1, \ldots, W_n\}$.

Transferring preference from $A >_{\text{pref}} B$ to worlds

$A >_{\text{pref}} B \Rightarrow \forall w_{A\wedge\neg B} \forall w_{\neg A\wedge B} : w_{A\wedge\neg B} >_{\text{pref}} w_{\neg A\wedge B}$

- $\forall w_{A\wedge\neg B}\ \forall w_{\neg A\wedge B} : \quad w_{A\wedge\neg B} >_{\text{pref}} w_{\neg A\wedge B}$
- $\exists w_{A\wedge\neg B}\ \exists w_{\neg A\wedge B} : \quad w_{A\wedge\neg B} >_{\text{pref}} w_{\neg A\wedge B}$
- $\exists w_{A\wedge\neg B}\ \forall w_{\neg A\wedge B} : \quad w_{A\wedge\neg B} >_{\text{pref}} w_{\neg A\wedge B}$
- $\forall w_{A\wedge\neg B}\ \exists w_{\neg A\wedge B} : \quad w_{A\wedge\neg B} >_{\text{pref}} w_{\neg A\wedge B}$

$A >_{\text{pref}} B \Rightarrow w_{A\wedge\neg B}, w_{\neg A\wedge B} \in W_i : w_{A\wedge\neg B} >_{\text{pref}} w_{\neg A\wedge B}$

Does it make sense to compare all worlds?
..contextual equivalence classes $W/\equiv\ = \{W_1, \ldots, W_n\}$.

$A \geq_{\text{pref}} B \Rightarrow w_{A\wedge\neg B} \geq_{\text{pref}} w_{\neg A\wedge B}$

Strict and non-strict preference.

Preference model $\mathcal{M} = \langle W, \succeq \rangle$

$$A >_{\text{pref}} B \Rightarrow \forall w_{A\wedge\neg B}\ \forall w_{\neg A\wedge B} \in W : w_{A\wedge\neg B} >_{\text{pref}} w_{\neg A\wedge B}$$

Kaci a Torre, 2005: Nonconflicting preference

Given a set S of states of affairs and a set $W \subseteq 2^S$ of possible words, then a preference model is a totaly ordered set $\mathcal{M}$ s.t.

$$\forall w_1, w_2 \in W : \quad w_1 >_{\text{pref}} w_2 \iff w_1 \succ w_2$$

Preference model $\mathcal{M} = \langle W, \succeq \rangle$

$$A >_{\text{pref}} B \Rightarrow \forall w_{A \wedge \neg B} \, \forall w_{\neg A \wedge B} \in W : w_{A \wedge \neg B} >_{\text{pref}} w_{\neg A \wedge B}$$

Kaci a Torre, 2005: Nonconflicting preference

Given a set S of states of affairs and a set $W \subseteq 2^S$ of possible words, then a preference model is a totaly ordered set $\mathcal{M}$ s.t.

$$\forall w_1, w_2 \in W : \quad w_1 >_{\text{pref}} w_2 \iff w_1 \succ w_2$$

Conflicting preference

$$\forall w_1, w_2 \in W : \quad w_1 >_{\text{pref}} w_2 \Rightarrow w_2 \not\succ w_1$$

Computing a preference model

- Preference can be rewritten to a disjunctive logic program (DLP).
- Semantics of the DLP defines the semantics of the preference.
- Semantics of a DLP based on *Optimal model* (Leone, Scarcello, Subrahmanian, 2004)
  - The optimal model can be defined so that the corresponding partial order resembles a total order as much as possible.

Hard constraints versus soft constraints

RQL's: a **hard constraint** – logical condition!

Filtering out of bad results! --→ Deficiencies:

- Not fulfilled – no perfect match --→ the **empty result**.
- Too loose selection condition --→ the **flooding effect**.

A **soft constraint** – preference (a wish)!

Not every wish can become true! --→ Filtering out of worse results:

- No perfect match --→ deliver best-matching alternatives!
- Never the empty result!

Preference operator versus preference operator

Selection operator

- The parameter is a logical condition!
- It returns the **perfect match, if present** in the DB. **Otherwise**, it delivers **empty result**!

Preference operator

- The parameter is preference!
- It returns the **perfect match, if present** in the DB. **Otherwise**, it delivers **best-matching alternatives**, but nothing worse!

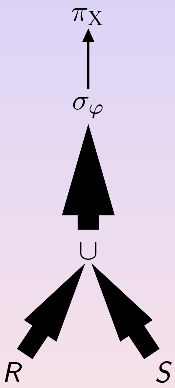
Better expressivity and higher performance!

Adaptive AND/OR-like filter effect

- Implicit query relaxation.
- On-the-fly filtering of worse results.

Optimization techniques for the preference operator

Example of push selection algebraic optimization strategy

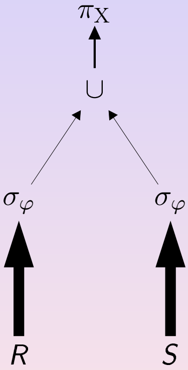


(SELECT **X** FROM **R** WHERE φ)
UNION
(SELECT **X** FROM **S** WHERE φ)

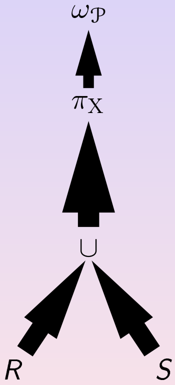
$$\pi_X(\sigma_\varphi(R \cup S)) \equiv \pi_X(\sigma_\varphi(R) \cup \sigma_\varphi(S))$$

Algebraic optimization strategies:

- *push selection*
- *push projection*

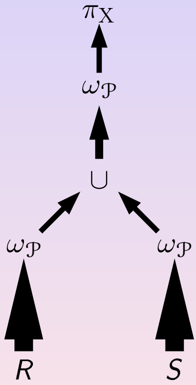


Push preference strategy



Algebraic laws involving preference operator:

- commutativity with selection,
- commutativity with projection,
- distributivity over cartesian product
- distributivity over union



Highlights:

Incorporating preference in RQL by means of preference operator

- with 1 parameter: preference
  - of various kinds,
  - including possible conflicts,
  - between elements or sets of elements.
- returning best possible result
- semantics of minimizing conflicts

Eliminated empty result effect!

Effective algebraic optimization (*push preference* strategy)!

To be done:

efficiency : another rewriting rules involving preference operator
⇢ novel **optimization strategies**:

expressivity : preference constructors to better eliminate the *flooding effect*

- pareto composition,
- lexicographic composition,
- prioritized preference

That's all.

Thank you for your attention!!