1/32

Information Access based on User Preferences

By: Nicolas Spyratos
Laboratoire de Recherche en Informatique
Université Paris sud 11, Orsay, France

Presented by:
M.Elias and A.Anjum
Université Paris sud 11, Orsay, France

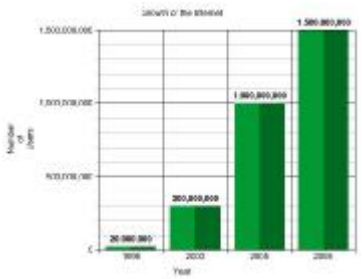
Supervised by: Emmanuel Waller

2/32

Outline

3/32

- The World Wide Web is exploited by more and more people



Internet World Stats. (2008). *Internet Growth Statistics*. Retrieved August 7, 2008, from http://www.internetworldstats.com/emarketing.htm

4/32

General Problem

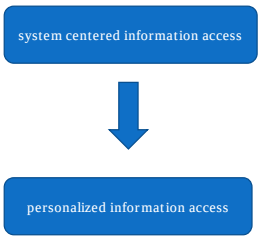
- Internet and Web provides huge amount of information
- Internet and Web casual users are increasing
- Information is useful to user when:

  - can be retrieved from an information source easily

  - retrieved results can be presented in an understandable way to user

5/32

Approach



6/32

Outline

Personalization

- How to achieve personalization?
  - By incorporating user preferences in the dialogue between the user and the information system

Preferences

- How can a user express preferences?
- Preferences are distinguished by:
  - Nature:
    - Quantitative: 'I like BMWs 80%'
    - Qualitative: 'I like BMWs more than VWs'
  - Duration:
    - Long term: system mining query logs
    - Short term: expressed explicitly by the user
- This work focuses on short term, qualitative preferences

Information System

- An information system extends the database with a couple of software tools for querying, presenting, transforming and analyzing the data.
- Tabular data: tables contain items and records or rows and columns

Outline

Information System Problems

- Tables contains thousands of rows, so the answer set of the query can be:
  - Very small, thus unsatisfactory for the user
    - We have to enlarge the query
  - Very large, thus difficult to exploit by the user
    - We have to narrow the query
- Our paper addresses the problem of very large answer sets by rewriting the query and using user's preferences

Preferences

- How can preferences be used to rewrite the user query?
- We embody the rewriting in the query language to generate preference query
- Preference query: is an ordinary query together with a set of preferences expressed by the user online.

Information Access based on User Preferences

Preference Query

- The user submits a query Q and his preferences online
- The system rewrites Q into a sequence of sub-queries Q1-> Q2->...->Qn whose answers contain the information retrieved in decreasing order of preferences
- User can stop the presentation of sub-queries results at any time he decides to

Example: Electronic Catalog for Cars

- preference query <Q, P.A>
  - Q is an ordinary query
  - P.A is a preference relation, we say " t is preferred to t' " if t.A->t'.A is in P.A
- P.{Model, Color}: Red^VW->Yellow^BMW, Black^VW->Yellow^BMW
- The answer to the preference query is:

Serial	Model	Colour	Mileage	Price	Year
1	BMW	Black	15000	5000	2002
2	BMW	Black	60000	7000	2000
3	VW	Red	25000	6000	2002
4	Honda	Blue	61000	2000	2001
5	VW	White	25000	2000	2003
6	BMW	Green	15000	7000	2002
7	Toyota	Black	60000	6000	2000
8	VW	Red	24000	6000	2002
9	VW	Yellow	11000	7000	2007

Q0=Q^[(Red^VW)˅(Black^VW)]
Q1=Q^(Yellow^BMW)
Q2=Q^¬(Q0^Q1)

- where top blocks of are more close to user interests

Outline

- Introduction
- Problem
- Solution
- Implementation
- Conclusions and Perspectives

Multiple Preferences and Priorities

- More than one preferences e.g.
  P.Model: BMW **à** V W, BMW **à** Honda, V W **à** Honda
- Preference on another attribute e.g.
  P.Colour : Red **à** Yellow, Black **à** White, White **à** Yellow
- Combination of two attributes e.g.
  P.{Model, Colour}: Red and BMW **à** Yellow and V W
- Preferences over more than one attributes separately e.g.
  P.Model: BMW **à** VW, BMW **à** Honda
  P.Color : Red **à** Yellow, Black **à** White
- **User might define priorities on preferences in the above case e.g.**
  Priorities: P.Model à P.Color

Example: Electronic Catalog for Cars

- Q= [(Model= BMW) Or (Model= VW)] And (Mileage≤40000 Km)
- Result:
  ans(Q) =
  ({1, 2, 6} U {3, 5, 8, 9})
  ∩{1, 3, 5, 6, 7, 8, 9}
  = {1, 3, 5, 6, 8, 9}

Serial	Model	Colour	Mileage	Price	Year
1	BMW	Black	15000	5000	2002
2	BMW	Black	60000	7000	2000
3	VW	Red	25000	6000	2002
4	Honda	Blue	61000	2000	2001
5	VW	White	25000	2000	2003
6	BMW	Green	15000	7000	2002
7	Toyota	Black	60000	6000	2000
8	VW	Red	24000	6000	2002
9	VW	Yellow	11000	7000	2007

- Problem:
  Result set could be very large and cars of the user's interest could be lost
- Our solution is to present the results cars in descending order, according to user preferences.

Interface

- Input of Queries, Preferences and priorities
- Control of Presentation of answer set by two buttons "Next" and "Stop"

Information Access based on User Preferences

Interface

- Searching the table: User enters ordinary query Q i.e. Where Attribute = value
- Preferences: Selection of attribute by menu, then declares pairs of values of that attribute
- Priorities: If more than one preference relation declared, user is asked to declare priorities

Control of Presentation

- Controlling the Presentation: By two buttons "Next" and "Stop"
- Next button pressed when user wants to see the next block
- Stop will stop further execution i.e. user has found his interest
- Records of User interest will be shown in descending order thus no need to execute all the sub queries in the answer to the preference query

Outline

Conclusion and Perspectives

- Conclusion
  - § Keep the task of user as simple as possible
  - § Preference queries are different from Order-By Instruction of SQL
- Perspectives
  - § Creating intervals over attributes e.g. Mileage or Price

Further Explanation

Order By vs Preference Queries

- Preference Queries differ from Order By instructions
- Order-By sorts the data (Ascending or Descending) following the predefined order of some attribute domain
  - § E.g. the predefined order over the domain of attribute price is that of integers

Information Access based on User Preferences

Order By vs Preference Queries

- In preference queries User inputs an order for attribute domain (this order may conflict the order of attribute domain)
- Some attribute domain have no predefined order (domains of Model and color are un- ordered), Order By doesn't work here
- Preference queries can be applied on any attribute without considering ordering of domain

Example: Two attributes preference relation

- P.{Model, Color}:
  Red^VW->Yellow^BMW, Black^VW->Yellow^BMW
- The answer to the preference query is the following sequence:
  Q0=Q^[(Red^VW)˅(Black^VW)]
  Q1=Q^(Yellow^BMW)
  Q2=Q^¬(Q0^Q1)

Further Explanation

- Order By vs Preference Queries
- Explanation of Algorithm

Evaluation of Preference Queries

Algorithm
Input: A preference query <Q, P.A> such that the graph G(P.A) is acyclic.
Output: The sequence Q0, Q1, .., Qm+1 answering the preference query
Method:
1/ Ordered-partition(G(P.A)) {the output is a sequence B0, B1, .., Bm of sets of values of A}
2/ For each i=0, 1, .., m do
begin
Q'i := conjunction of all values of A in Bi;
Qi := Q^Q'i;
output Qi
end
3/ Qm+1 := Q^ ¬ (Q0vQ1v…vQm);
output Qm+1