Universität Augsburg

Cosima^T - Preference Based Search Technology in Tourism

Sven Döring

Report 2006-24 November 2006

Institut für Informatik
D-86135 Augsburg

COSIMA^T – Preference Based Search Technology in Tourism

Sven Döring
Chair for Databases and Information Systems
University of Augsburg, 86135 Augsburg, Germany
doering@forsip.de

Abstract
Today there are still a lot of people preferring to consult a human employee in a travel agency instead of using the internet for booking or organizing a travel. Technical problems, incomprehensible interfaces, and insufficient search engines are part of that problem. With the interplay and adjustment of several advanced, preference driven middleware components we achieve to automate skills that so far could be executed only by a human employee in a travel agency. Our search engines *Preference XPath* and *Preference SQL* deliver best alternatives, if there is no perfect match. It is possible to distinguish between hard and soft constraints like a human vendor would do. The *Preference Presenter* implements a smart and sales psychology based presentation of search results, supporting various human sales strategies; the *Preference Repository* provides the management of situated long-term preferences. A novel query rewriting approach enables the smart combination of single preferences, e.g. the category of a hotel, with global preferences like the overall prize for the whole journey enabling a deep personalized packaging of travels. The key technologies for this breakthrough are based on preferences modeled as strict partial orders. Our first advanced prototype COSIMA^T is promising.

1. Introduction

E-commerce has a major impact on the tourist and travel industry. For example low-cost airlines like Ryanair[1] or EasyJet[2] gained huge significance within the last years. They strongly rely on their internet portal in order to sell tickets. Nevertheless, lots of people prefer to consult a human employee in a travel agency instead of using the internet for booking or organizing a journey.
The overall tourism information space is huge [28]. Average consumers are often overstrained arranging a vacation in the internet due to technical problems, incomprehensible interfaces and insufficient search engines. This might be one reason that there is still little evidence of electronic markets leading invariably to lower search costs [24]. If a customer specifies a lot of search preferences existing online booking engines will often return no solution [33]. This is called the *empty result effect*. The user has to start another search with a subset of his preferences until the system returns a solution. This can make the planning process very tedious. Thus, new and customer friendly search technologies are necessary [25]. E.g. the customer's needs and constraints could not only be noticed top-down. Instead they could be constructed in a bottom-up way [30]. Unfortunately, this can be a time-consuming process too and even good recommendations are useless, if the site navigation is confusing or the item descriptions are too terse. A personalized search and product presentation is necessary in order to convince also inexperienced consumers of the advantages of internet travel portals and booking engines.
Existing online systems like Expedia[3] are sufficient for simple problems like inquiring a flight from A to B with specific hard constraints. However, complex problems involving knowledge about a specific traveler require the modeling of that knowledge as well as more intelligent systems [34]. The selection process of matching products and packages is quite challenging from the database point of view. There are preferences regarding single aspects of the journey, for example the flight, and there are preferences about global constraints like the overall prize of the whole journey. In addition, some preferences strongly depend on the situation of the consumer, e.g. if it is a private or business travel. There are wishes which **should be** fulfilled and constraints which **must be** fulfilled, respectively.
A preference driven search might avoid or reduce a lot of typical problems as already shown for e-commerce [5]. In section 2 we will present some of our preference driven middleware components.

[1] www.ryanair.com
[2] www.easyjet.com
[3] www.expedia.com

We will shortly introduce the personalized search *Preference XPath* and *Preference SQL*, the *Preference Presenter* implementing a sales psychology based presentation of search results, and the *Preference Repository* responsible for the management of situated long-term preferences. They are generic in the sense that they can be customized for different application scenarios. Furthermore, an overview about related work in the tourism domain is given. In section 3 the main part of our work will present our complex customization process for the travel and tourist domain. A novel approach for the usage of our components dealing with single as well as global preferences is shown. In addition, our first advanced prototype COSIMA^T will be presented in section 4. Finally, a conclusion and outlook are pointed out.

2. Preliminary Work

In this chapter we will introduce our preliminary work compromising a preference driven search process. Though firstly, a survey of related work from the tourist domain is given.

2.1 Related Work in Tourism

The internet offers a huge amount of data for travelers. For instance it is possible to find information about the weather, the currency, the language, possible activities, and so on. Within the *ESPRIT* project MIRO-Web ([10]) a set of middleware components was developed in order to provide transparent access from standard web browsers to multiple heterogeneous data sources. MIRO-Web is based on a three-tier architecture with a Data Source Adapter Layer, a Mediation Layer and a Client Layer. Heterogeneous data are transformed into a structured format by wrappers. Thereafter they are integrated and combined by mediators before they can be shown to users. In another work ([1]) mediator components based on the Resource Description Framework RDF[4] are proposed. Again based on RDF a mediation facility responsible for the integration of heterogeneous data from hotel suppliers is presented ([15]). In *TheaterLoc* ([2]) an entire virtual application based on wrappers and mediators was implemented that allows users to get information about theaters and restaurants for a lot of cities in the USA. The core of the *Harmonise* project ([11], [22]) is a shared, conceptual reference schema, the so called Interoperable Minimum Harmonise Ontology (IMHO). Stakeholders in the tourism industry do not have to change their own data format. Mediators transform the local data format into a representation based on IMHO and vice versa.

The major impact of e-commerce to the tourist and travel industry is shown in [35]. Affects to typical market players like tourists or travel agents are figured out. Furthermore this work points out the importance of attentive user interfaces, personalization, and the user modeling process. In [8] an observational study delivered six different decision styles of users mainly influencing the decision process, e.g. while *highly pre-defined* users are quite sure about their destination, the *recommendation-oriented* user does not have any fixed feature of the trip in mind. In [30] two groups of factors influencing the destination choice are identified: personal features like age, education, experience, personality etc. and travel features e.g. travel purpose, travel-party size, and distance.

With *reality* a travel planer is presented that uses a conversational model instead of a sequential one in order to deliver best solutions [33]. In a kind of dialog the user can react to suggested solutions by adding, modifying or removing some preferences. After each modification the corresponding new best solutions are automatically computed and displayed. A study [27] analyzing the records of caller dialog indicates that travelers mainly search by location or by interests.

The vague Query System (VQS) uses multidimensional concepts and so called Numeric-Coordinate-Representation-Tables to carry out similarity searches ([25], [26]). Using a computed total distance, VQS always tries to deliver best matches. Thus, the annoying empty result effect is avoided. Unfortunately, the result is presented as ranked list together with the value for the total distance. Often such a numerical presentation is not intuitively comprehensive by human beings [18]. Another study proposes an interactive query management which can deliver results by the relaxation of query constraints in case of an empty result [31]. Only one constraint is changed in time. In a kind of dialog the user can decide which search attribute he is willing to relax or to skip. In case of categorical data no reasonable relaxation can be given. Thus, in such a case the whole constraint is discarded.

[4] http://www.w3.org/RDF

Search term associations are used in order to support users' keyword search with appropriate additional search terms either narrowing or expanding the original keyword search query [9]. The search term associations are constructed by mining user interactions with the system. The approach assumes that there is a semantic relationship between a keyword and the web pages' abstract which provoked the user to visit a site. Search pattern of users are examined in a study using log files from Visiteuropean-cities.info [23].
Presenting a map with points of interests, e.g. hotels, can support users in their decision process. In [26] and [29] a geographic search is presented. Maps with an integrated view of geographic and tourist information are shown to the user. However, the calculation of distance is based on the Euclidean distance, this means the distance is only the beeline between two geographic locations.
Since users may find different information relevant a search engine should work accordingly. Taking the preferences of users into account is a promising approach for the personalization of database queries [3]. Experts as well as inexperienced users prefer search results adjusted to their preferences. In [21] a model for representing and storing preferences is proposed. Furthermore, algorithms for the generation of personalized search results are presented. Numerical values between -1 and 1 are used to express the interest, i.e. the preference, of a user. However, this seems not an intuitive understandable model. In the next section preferences are modeled in a more natural fashioned way.

2.2 Preference Based Middleware

Advanced applications of the travel and tourism industry require a high level of personalization and situation awareness in order to provide individual recommendations and custom-tailored travel packages. Within the research program "It's a Preference World" at the University of Augsburg preferences are treated as significant factor for e-services. Preferences are modeled as strict partial orders with intuitive comprehensible "A is better than B" semantics (see [16], [17]).
In the following we will demonstrate the usage as well as the benefits of our flexible middleware components within the tourism domain. Therefore let us consider a common example. Mr. Black is a business traveler. He has a business meeting in Munich and is looking for a hotel for an overnight stay. Please note, that the formulation (**should be**) indicates a preference instead of a hard constraint (**must be**). He expresses his interests:

"I would like to have an accommodation for one night. It ***should be*** *a* ***double room.*** *The* ***prize must be less or equal 70 EUR*** *according to the travel policy of our company."*

	Name	Category/Stars	Meals	Room	Location	Prize_per_night (€)
t_{H1}	IBIS	3	full	single	downtown	85.--
t_{H2}	ARCOR	3	breakfast	single	city	70.--
t_{H3}	PLAZA	4	breakfast	single	outside	75.--
t_{H4}	GREEN	3	none	double	city	75.--
t_{H5}	SUNSHINE	3	breakfast	double	downtown	80.--
t_{H6}	HILTON	5	breakfast	double	city	100.--
t_{H7}	KING	4	breakfast	double	downtown	90.--
t_{H8}	TOKIO	2	none	single	city	60.--

Table 1: Hotel database

Preference Search Engine:
Considering the hotel database in Table 1 above there would be no perfect match treating all search preferences of the example as hard constraints which leads to the annoying empty result effect. When interpreting the constraints as or-conditions, the flooding effect with a lot of irrelevant results occurs. Another approach is to iteratively ask the customer to soften his or her search criterions, which is a very frustrating and time-intensive process. There is a need for a search engine delivering the best alternatives when there is no perfect match. With preference search engines like *Preference SQL* respectively *Preference XPath* wishes can be easily expressed within one statement (see [18], [20]).

They are fully compatible with standard XPath and SQL, respectively. Therefore, it enables the combination of common hard selection conditions as well as soft selection conditions. There, instead of the SQL keyword for hard conditions “WHERE” the keyword “PREFERRING” is used, analogously in Preference XPath “#[]#” instead of “[]”:

```
SELECT * FROM HOTEL WHERE Prize_per_night <= 70
      PREFERRING Room = 'double'                                    Q1

/HOTEL [Prize_per_night <= 70] #[Room is 'double']#                 Q2
```

Furthermore, preference search engines reduce the flooding effect with lots of irrelevant results by filtering objects that are subsumed by better ones, delivering best matches only (BMO, see [16]). Using the queries Q1 and Q2 tuple t_{H2} and t_{H8} of the hotel database would be delivered as best matches. The preference for a double room could not be achieved, but both tuples match the hard constraint regarding the prize limit.

Preference Repository:

Usually, customers expect their long-term preferences to be considered automatically by a good and familiar appointee in a travel agency. Such preferences can be gained from log files by preference mining algorithms (see [13]). The Preference Repository provides an XML based storage structure for preferences.

```
<PreferenceRepository>
 <UserIdentifier>
   <Name xml:lang="en">Black</Name>
 </UserIdentifier>
 <PreferenceData name="meal_business">
  <Situation>
   <Conditionkey="role" value="business"/>
  </Situation>
  <Preference>
   <POS att="Meals">
     <Value val="breakfast"/>
   </POS>
  </Preference>
 </PreferenceData>
 <PreferenceData name="meal_private">
  <Situation>
   <Conditionkey="role" value="private"/>
  </Situation>
  <Preference>
   <POS att="Meals">
     <Value val="none"/>
   </POS>
  </Preference>
 </PreferenceData>
 …
</PreferenceRepository>
```

Figure 1: Preference Repository excerpt

Moreover, relevant information about the situational context of a preference can be managed, e.g. a situation may be specified by the role of the customer. Considering the example above, Mr. Black may have the preference to get breakfast included when traveling for business while preferring no meals in hotels during a private journey. An excerpt of the appropriated storage structure of the *Preference Repository* is shown in Figure 1. Obviously, Mr. Black's long-term preferences should be considered by a personalized application of the tourism and travel domain. Inserting this information the search queries Q1 and Q2 would be adjusted accordingly. Since Mr. Black intends to do a business travel his preference for the breakfast should be included:

```
SELECT * FROM HOTEL WHERE Prize_per_night <= 70
      PREFERRING Room = 'double' and Meals = 'breakfast'                Q3
```

```
/HOTEL [Prize_per_night <= 70]
      #[Room is 'double' and Meals is 'breakfast']#                     Q4
```

The new result set includes tuple t_{H2} of the hotel database only. Tuple t_{H8} is dominated because it does not include the breakfast.

Preference Presenter:

For a successful deal the product presentation is a decisive factor. Unlike to human appointees in a travel agency, travel online-portals do not consider principles of sales psychology when presenting the search results. In real world a sales agent has to find a way to satisfy his or her own preferences and the preferences of the customer as good as possible, which is a challenging act. In order to convince the customer from his or her offered goods a decisive factor is to argue about the quality of the presented products with respect to the search preferences of the customer. This is stated by well-known models of user behavior (see [14]) within sales scenarios. Some of today's search engines use similarity searches ([25], [26]) in order to compute alternatives in case of a missing perfect match, but they are not able to provide semantic information about the quality of the search result. Our *Preference Presenter* component can automatically deliver intuitively comprehensible quality information for each single preference as well as a situated and personalized overall valuation of the quality of each tuple in the result set (see [7]). Instead of using numerical scores we claim that using linguistic terms (see [36]) is an appropriate choice. Empirical psychological studies support that an ordered linguistic domain with about five terms is a reasonable way for many applications. For the scope of online travel portals we decided on this choice and their respective ordering: 'sufficient' < 'acceptable' < 'good' < 'very good' < 'perfect'. The result of Mr. Black's search can be automatically presented by the following natural language dialog:

> *"Overall this hotel fits your preferences* ***very good****, because it* ***perfectly*** *hits your company's prize limit and it* ***perfectly*** *fulfills your preference for breakfast during business trips. Thus, I assume the single room instead of the double room is acceptable."*

3. Preference Driven Applications in the Tourism Domain

E-commerce has a major impact on the tourist and travel industry. Low-cost airlines like Ryanair or EasyJet gained huge significance within the last years. They strongly rely on their internet portals in order to sell ticket. Thus, attentive user interfaces, personalization, and the user modeling process are extremely important [35]. Using our previous knowledge (section 2.2) and expertise on personalization technologies in e-commerce we examined an approach to enhance the search process enabling a better, personalized product search and result presentation in travel and tourism applications.

3.1 Problem Definition

As confirmed by studies (e.g. [24]) consumers are often unable to arrange vacations in the internet by themselves because of technical problems or insufficient search engines and interfaces. Online systems like Expedia are sufficient for simple problems like inquiring a flight from A to B with specific hard constraints. However, complex problems involving knowledge about a specific traveler require more intelligent systems [34]. Of course, those problems are not limited to tourism. They also occur in a lot of e-commerce applications (see [5], [18]). But due to the complex nature of the decision process in tourism these problems are even worse. Thus, the selection process of fitting products and packages is quite challenging from the database point of view.

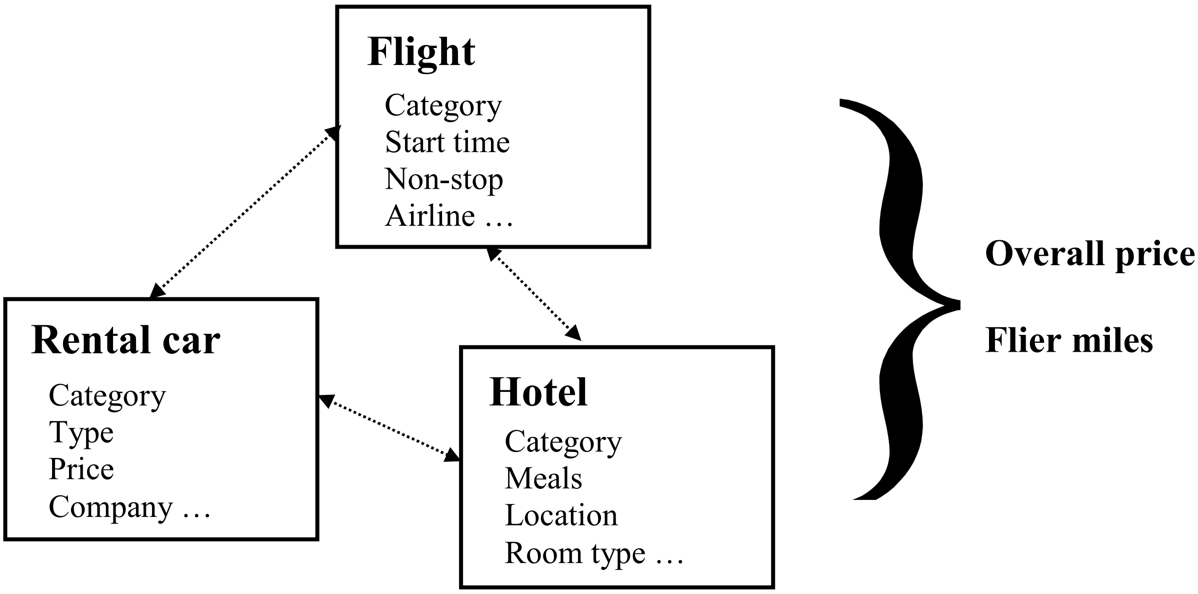

Figure 2: Complex preferences in tourism

There are preferences regarding single aspects of the journey, for example the flight, and there are preferences about global constraints (see Figure 2). In addition, some preferences strongly depend on the situation of the consumer, e.g. if it is a private or business travel. There are wishes which **should be** fulfilled and constraints which **must be** fulfilled, respectively. In the next section a novel approach using preference technology is presented in order to address these problems.

3.2 Preference Based Search in Tourism

Some problems are already addressed by our unique preference search. It is not only possible to distinguish between hard and soft constraints, but also to incorporate the situational context of preferences (see section 2.2. and [5], [12], [18]). However, the problem of a reasonable treatment of global and single preferences is still a huge challenge. We meet this problem by usage of a query rewriting process. Search queries for single parts of the travel, e.g. hotels or flights, are smartly rewritten and extended in order to address the global constraint. Subsequently, this approach is presented in detail. For illustration let us consider an example including three common parts of a travel: flight, hotel, and rental car (Figure 2). A small database exists for each of those parts. Table 1 (above on page 3) represents the data set for hotels, Table 2 for rental cars, and Table 3 for flights:

	Name	Category	Type	Prize_per_day (€)
t_{C1}	Audi A4	medium car	limousine	58.--
t_{C2}	Opel Zafira	medium car	van	73.--
t_{C3}	Renault Megane	compact car	limousine	41.--
t_{C4}	BMW 7	luxury car	limousine	80.--
t_{C5}	Ford Fiesta	small car	limousine	25.--
t_{C6}	Skoda Superb	medium car	limousine	67.--

Table 2: Rental car database

	Category	Airline	Prize (€)
t_{A1}	economy	Lufthansa	400.--
t_{A2}	economy	Air Spain	300.--
t_{A3}	economy	Ryan Air	200.--
t_{A4}	business	Air Spain	500.--
t_{A5}	business	Air France	700.--

Table 3: Flight database

Mr. Black intends to travel to Barcelona for a week. He is able to express his preferences as follows:

*"The flight's airline **should be Lufthansa** and the seat **should be** in the **business category**. Furthermore, the hotel **should have 3 stars** and **should be located in downtown**. I need a rental car, since I have to be mobile. It **should be a compact car**. The whole trip **must not** cost more than **900 EUR.** "*

After feeding his wishes into an intuitive comprehensible search mask (e.g. Figure 4 on page 12) search queries are created with respect to the user's preferences.

1. Query composition:
Using the information gained from the search mask preference queries can easily be composed in Preference SQL or Preference XPath. Of course long-term preferences of the customer, for example stored in the Preference Repository, can also be added. Since we used Preference XPath (see [19],[18],[4]) in our prototype the following queries are formulated in this way.

`/FLIGHT #[Airline is 'Lufthansa' and Category is 'business']#` **Q5**

`/HOTEL #[Location is 'downtown' and Category around 3]#` **Q6**

`/CAR #[Category layered ('compact car', 'medium car', 'small car')]#` **Q7**

While Q7 states that a compact car is preferred it also expresses that a medium car is a good alternative and better than a small car or the rest of the cars, respectively. Such knowledge can be gained from an advanced search mask, the *Preference Repository*, domain knowledge or by advanced data mining techniques like the Preference Miner ([12]).

2. Query rewriting and extension:
The queries Q5 till Q7 could be started, already. However, let us consider global preferences like the overall prize. People often tend to underestimate or overestimate aggregated values. In addition, they often try *to play* with the system using unrealistic values. This especially applies for prizes and costs. A human vendor would suggest an alternative if there is no travel available matching the customer's prize constraints. We extend our queries to do so likewise. If a customer expresses a prize limit for the whole package we add the preference constructor *lowest* for the prize attribute of each query. Therefore, we also get the cheapest products for each part of the travel (e.g. hotel, flight) and it is easily possible to check if the constraint for the overall prize can be matched at all just by adding the prizes of the cheapest products. For illustration, let us consider Mr. Black's preferences from above. If the cheapest flight to Barcelona would cost 800 EUR, the cheapest hotel 400 EUR, and the cheapest car 200 EUR. Then his prize limit of 900 EUR obviously could not be matched even without considering his other preferences, e.g. for the category.

The extended queries are as follows:

```
/FLIGHT #[Airline is 'Lufthansa' and Category is 'business'
      and Prize lowest]#
```
Q8

```
/HOTEL #[Location is 'downtown' and Category around 3
      and Price_per_day lowest]#
```
Q9

```
/CAR #[Category layered ('compact car ','medium car', 'small car')
      and Prize_per_night lowest]#
```
Q10

3. Calculation of results and their quality:
The preference search engine is now able to compute the three result sets:

$\text{BMO-Set}_{Flight} = \{t_{A1}, \mathbf{t_{A3}}, t_{A4}\}$

$\text{BMO-Set}_{Hotel} = \{t_{H2}, t_{H5}, \mathbf{t_{H8}}\}$

$\text{BMO-Set}_{RentalCar} = \{t_{C3}, \mathbf{t_{C5}}\}$

Please note, the tuples t_{A3}, t_{H8} and t_{C5} are only included into the result set because of the search queries' extension.
After this, the quality of the tuples is computed (see section above) and expressed in linguistic terms from perfect to sufficient (see [7]). A single product's prize (e.g. for a flight or hotel) won't be included in the computation of its quality since this extension of the query was made by us. This approach makes sure the tuples will be valuated exactly in terms of the user preferences.

		Quality	Prize (€)[5]
t_{A1}	Flight	good	400.--
t_{A3}	Flight	sufficient	200.--
t_{A4}	Flight	good	500.--
t_{H2}	Hotel	very good	490.--
t_{H5}	Hotel	perfect	560.--
t_{H8}	Hotel	acceptable	420.--
t_{C3}	Rental Car	perfect	287.--
t_{C5}	Rental Car	acceptable	175.--

Table 4: Qualities of the result set tuples

In Table 4 quality and prize of each result tuple are shown. For example, t_{C3} was valuated with *perfect* because it exactly matches Mr. Blacks preference for a compact car while t_{C5} just got an *acceptable* valuation. This is reasonable because it is a cheap alternative to the compact car only included to the result set because of our prize extension to the original preference search query.

4. Composition of the package:
After the valuation of each single product, price and quality of each possible combination of flight, hotel and car are computed.

[5] Please note, each prize is cumulated for the whole trip, which means 7 days/nights in case of the rental car or accommodation.

Combination	Overall quality	Overall prize (€)	Flight	Rental Car	Hotel
1	very good	1347.--	t_{A4}	t_{C3}	t_{H5}
2	very good	1277.--	t_{A4}	t_{C3}	t_{H2}
3	very good	1247.--	t_{A1}	t_{C3}	t_{H5}
4	very good	1147.--	t_{A1}	t_{C3}	t_{H2}
5	very good	1047.--	t_{A3}	t_{C3}	t_{H5}
6	good	1235.--	t_{A4}	t_{C5}	t_{H5}
7	good	1207.--	t_{A4}	t_{C3}	t_{H8}
8	good	1165.--	t_{A4}	t_{C5}	t_{H2}
9	good	1135.--	t_{A1}	t_{C5}	t_{H5}
10	good	1107.--	t_{A1}	t_{C3}	t_{H8}
11	good	1065.--	t_{A1}	t_{C5}	t_{H2}
12	good	977.--	t_{A3}	t_{C3}	t_{H2}
13	good	935.--	t_{A3}	t_{C5}	t_{H5}
14	good	907.--	t_{A3}	t_{C3}	t_{H8}
15	acceptable	1095.--	t_{A4}	t_{C5}	t_{H8}
16	acceptable	995.--	t_{A1}	t_{C5}	t_{H8}
17	acceptable	865.--	t_{A3}	t_{C5}	t_{H2}
18	acceptable	795.--	t_{A3}	t_{C5}	t_{H8}

Table 5: Combinations of the single products to a package

The combinations are listed in Table 5. For the valuation of the overall quality the intuitive comprehensible equidistant linguistic average is used. However, there are other strategies possible like the optimistic or pessimistic valuation (see [6], [7]) depending on situation and customer. Only the last two product combinations match Mr. Black's prize preference. However, the quality of these combinations is just *acceptable*.

5. Presentation of the product

There exists a huge amount of possibilities to present the result to the customer. Combination 17 and 18 (see Table 5) could be presented because they match Mr. Black's prize preference. Furthermore, combination 1 might be shown because of its very good overall quality or combination 14 because of its good ratio between quality and prize. This strongly depends of the preferences of the customer as well as the situation of customer and company.

Importantly, the quality information can be used to convince or influence the customer like a human vendor would do. It is possible to argue about the search result and point out its quality in natural language:

> *"We can offer you two travel packages matching your price preference in **acceptable quality**. Furthermore there is one travel package in **good quality** which exceeds your price preference by less than 1 %"*

We introduced a novel approach for advanced applications of the tourism domain enabling deep personalization and situation awareness. By the usage of preference search engines it is possible to distinguish between hard and soft constraints. Situational knowledge can be modeled and integrated easily. Additionally, a combination of single and global preferences is enabled by our new rewriting approach. Thereby, we avoided or reduced serious problems of current travel portals. In the following chapter we will introduce our prototype COSIMA^T based on our preference technologies.

4. The COSIMA^T Prototype

Advanced applications require a high level of personalization, individualization and situation awareness in order to provide individual product recommendations, custom-tailored travel packages and prize offers. Our prototype is based on existing preference middleware components (section 2.2) like the *Preference Presenter* or the *Preference XPath* search engine offering the so called *P-Services* (see Figure 3).

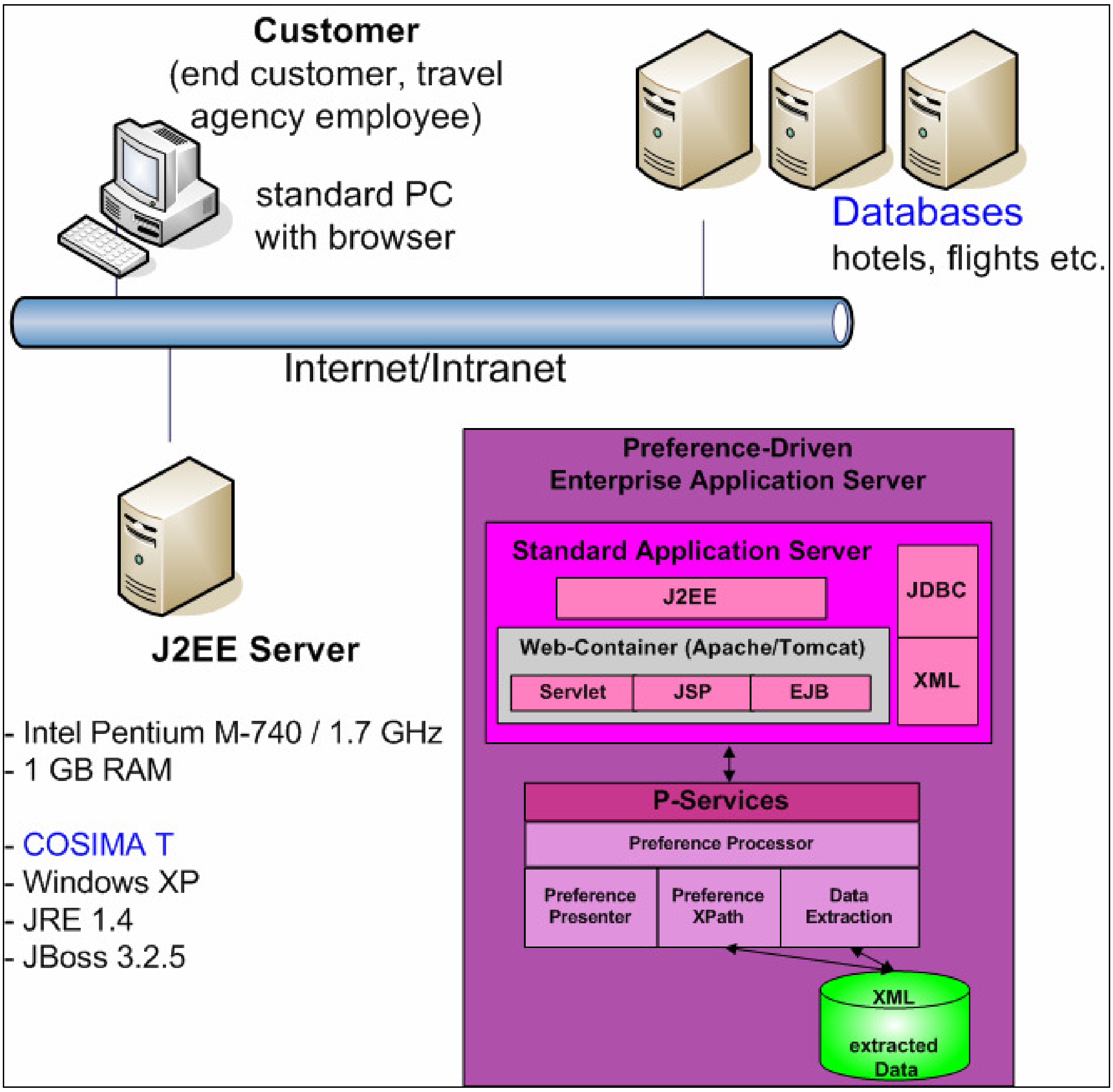

Figure 3: Architecture of COSIMA^T

We properly combined those components adjusted to the tourist and travel domain. Thereby, we only used widespread standard technologies like Java, J2EE, and XML. All components are implemented as Enterprise JavaBeans running on the open source solution JBoss Application Server[6]. Since we used only the specified functionality of J2EE our components are compatible to e.g. IBM Websphere[7], BEA Weblogic[8] and Oracle Application Server[9]. Our components are highly runtime efficient. Hence, the application server is located at a standard PC.
Potential customers or employees of travel agencies are able to access our intuitive comprehensible and compact user interface via internet using a standard PC with web browser. If a request of the customer is received, we extract all necessary data from provider databases for hotels, flights and rental cars. Thereafter, the request is evaluated by preference middleware components using our novel search query rewriting approach and the result is presented with respect to the customer's preferences. In the following the prototype is introduced by a sample tour.

4.1 Prototype Tour

For illustration purposes let us consider Mr. Black again. He still wants to travel to Barcelona for a week:

> *"The flight's airline **should be Lufthansa** and the seat **should be** in the **business category**. Furthermore, the hotel **should have 3 stars** and **should be located in downtown**. I need a rental car, since I have to be mobile. It **should be a compact car**. The whole trip **must not** cost more than **900 EUR.** "*

Mr. Black enters the website of COSIMA^T. He can easily express his preferences in the search mask (see Figure 4). In the part "Your personal travel preferences" wishes about the travel and its single aspects flight, hotel, and rental car can be specified. A global preference about the overall prize can be inserted beneath. The search process (section 2.2 and 3.2) starts after clicking the search button. Actually, the search constraints of Mr. Black cannot be fulfilled considering the sample databases of our prototype. Though, the *Preference Search* automatically delivers best alternatives if there is no perfect match. The search result is presented in the next mask (see Figure 5) offering human understandable information about the quality of the search result with respect to the customer's preferences.

[6] www.jboss.org
[7] www.websphere.com
[8] www.bea.com
[9] www.oracle.com

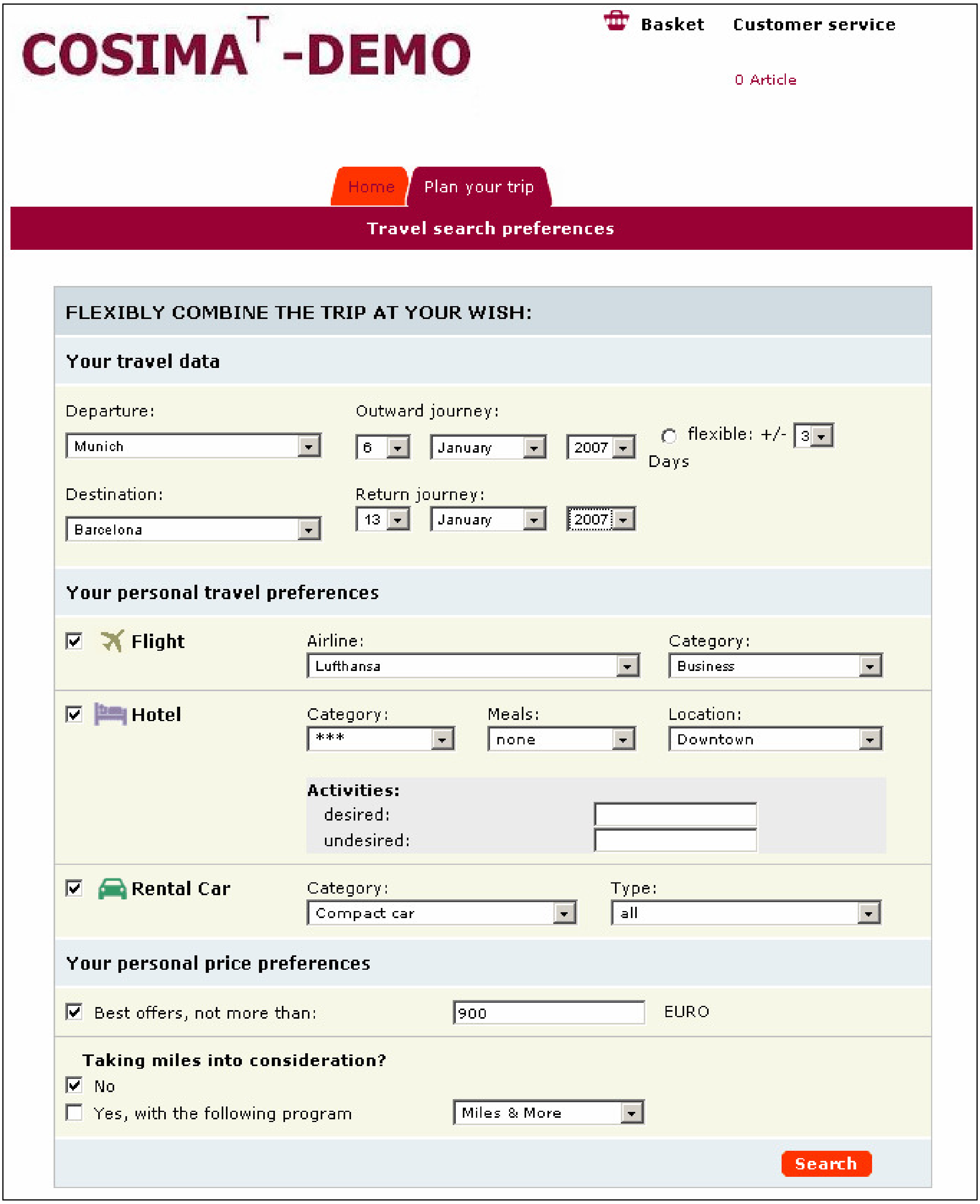

Figure 4: Search mask of COSIMA^{T}

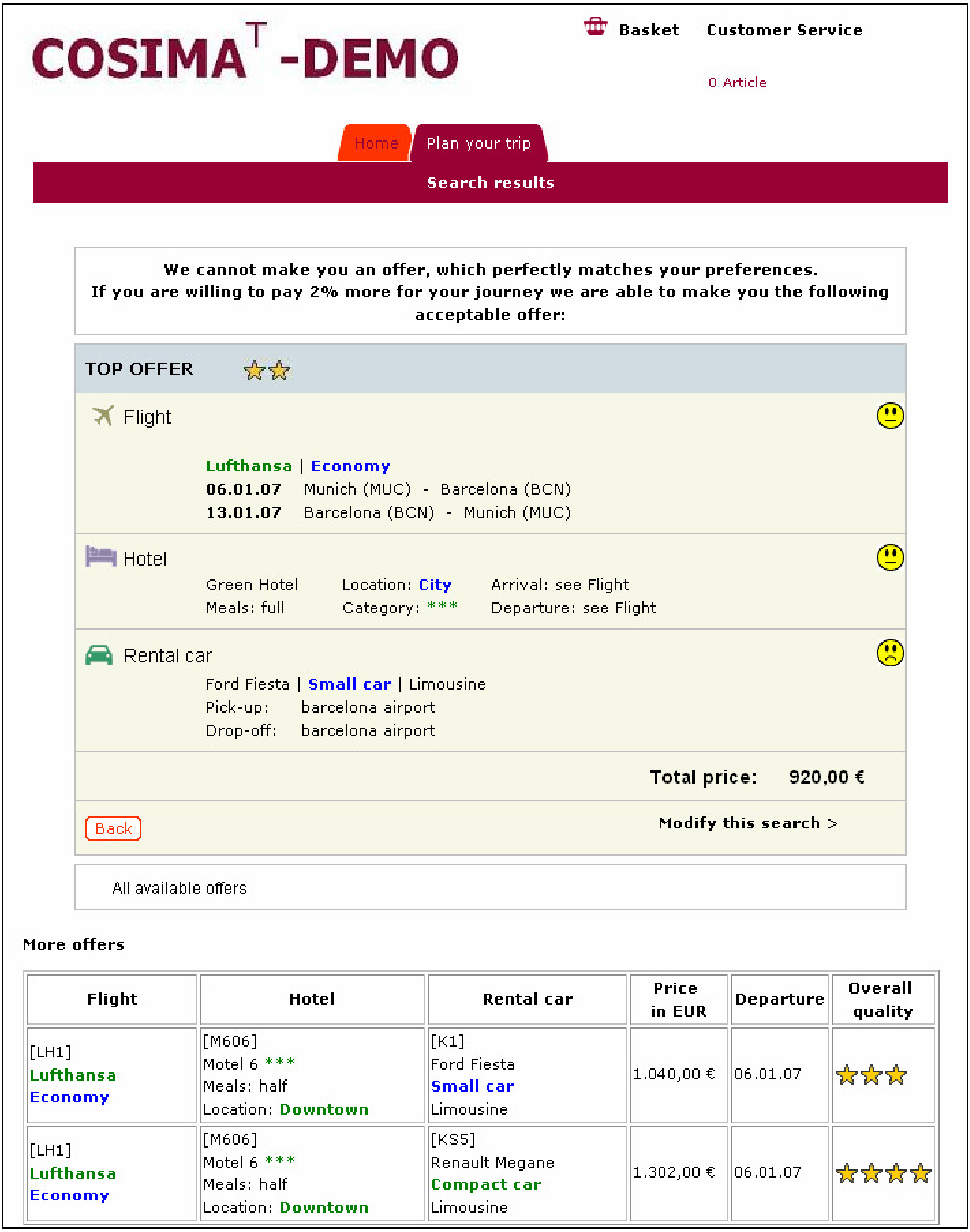

Figure 5: Search result presentation of COSIMA^T

The search result of Mr. Black's request is presented in Figure 5. Since there was no perfect match best alternatives have to be presented. The *Preference Presenter* component (see section 2.2) provides a smart result presentation offering human understandable information about the quality of the search result with respect to the customer's preferences. The quality of each single preference as well as the overall quality is evaluated in linguistic terms (see section 2.2). Since stars are already known in the travel and tourist domain as indication of the hotels' quality; we decided to use this symbol too. For an overview regarding our indication of the quality see Figure 6.

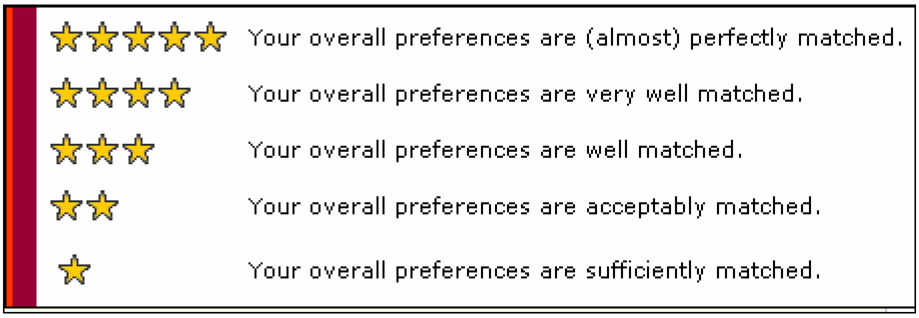

Figure 6: Evaluation of the quality

Of course, there are different strategies to present products and to argue about their quality (see [14], [6]). A vendor will be enabled to apply personalized sales strategies. Furthermore a top result has to be chosen. In our example the top offer was selected because it almost matches the prize preference of Mr. Black requesting an accumulated prize of not more than 900 EUR. Considering the above mentioned example COSIMA^T is able to argue about the search result and point out the top result as follows:

> *"We cannot make you an offer, which perfectly matches your preferences. If you are willing to pay 2% more for your journey we are able to make you the following **acceptable offer**"*

The top offer (Figure 5) has an acceptable overall quality and is marked with 2 stars. Furthermore, there are more alternatives presented, which have a better overall quality (3 or 4 stars, respectively) wrt. the preferences of Mr. Black. Please note, the offer at the bottom of Figure 5 meets almost all preferences of Mr. Black except his prize preference. The tradeoff between prize and quality is obvious. At last, Mr. Black will choose a travel package matching his preferences best.

5. Conclusion and Outlook

Lots of people prefer to consult a human employee in a travel agency instead of using the internet for booking or organizing a travel. Technical problems, incomprehensible interfaces and lacking search engines are part of that problem. We presented the personalized search *Preference XPath* and *Preference SQL*, the *Preference Presenter* implementing a sales psychology based presentation of search results, and the *Preference Repository* responsible for the management of situated long-term preferences. They are generic in the sense that they can be customized for different application scenarios. In the main part of our work we presented this complex customization process for the travel and tourist domain. An accordingly adjusted *Preference XPath* enables the distinction between hard and soft constraints automatically delivering best alternatives, if necessary. A novel query rewriting approach was shown in order to be able to combine single as well as global preferences in a smart manner. Afterwards, the *Preference Presenter* convincingly presents the search result. Finally, our first advanced prototype COSIMA^T implementing this research knowledge was introduced.

Our research and development for personalized applications will continue along various topics. As one next step we will examine and develop purpose-built search constructors for the tourist and travel domain in order to deliver even better search results and alternatives, in case there is no perfect match. Furthermore, an extension of the *Preference Presenter* using geographical information seems reasonable. Therefore, alternatives, e.g. for the departure airport, could be plausibly presented to the customer. Another important part of our future work will be the modeling of situations influencing the search process. The adaptation of the *Preference Recommender* [32] to the tourism domain could be one more step towards deep personalized and human like applications.

6. Acknowledgements

This work is supported by the Bavarian Research Cooperation for Situated, Individualized and Personalized Human-Computer Interaction (FORSIP 2), funded by the Bavarian State Ministry for Science, Research and Art, Germany.

References

[1] A. Ashari et al.: *A Framework for Integrating Heterogeneous Tourism Information Sources*. In Proceedings of the 8th International Conference on Information and Communication Technologies in Tourism (ENTER 2001), pages 102-111, Montreal, Canada, 2001.

[2] G. Barish et al.: *The TheaterLoc Virtual Application.* In Proceedings of the 7th National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence (AAAI/IAAI 2000), pages 980-987, Austin, Texas, USA, 2000.

[3] J. Chomicki: *Preference Formulas in Relational Queries.* In ACM Transactions on Database Systems 43 (4), pages 427-466, 2003.

[4] S. Döring, A. Eberle, T. Preisinger: *Delivering a Personalized Result Set by the Adaptation of Preference Queries*. In A. B. Cremers et al. (eds.): Informatik 2005, LNI Proceedings, Vol. P-67(1), pages 241-245, Bonn, Germany, 2005.

[5] S. Döring, W. Kießling, T. Preisinger, S. Fischer: *Evaluation and Optimization of the Catalog Search Process of E-Procurement Platforms*. In Journal of Electronic Commerce Research and Applications, Vol. 5, Issue 1, pages 44-56, Elsevier, 2006.

[6] S. Fischer: *Personalized Query Result Presentation and Offer Composition for E-Procurement Applications.* Dissertation, University of Augsburg, Augsburg, Germany, 2004.

[7] S. Fischer, W. Kießling, T. Preisinger: *Preference based Quality Assessment and Presentation of Query Results*. In G. Bordogna, G. Paila (eds.): Flexible Databases Supporting Imprecision and Uncertainty, Springer, 2006.

[8] K. Grabler, A. H. Zins: *Vacation Trip Decision Styles as Basis for an Automated Recommendation System: Lessons from Observational Studies*. In Proceedings of the 9th International Conference on Information and Communication Technologies in Tourism (ENTER 2002), pages 458-469, Innsbruck, Austria, 2002.

[9] U. Gretzel, K. Wöber: *Intelligent Search Support: Building Search Term Associations for Tourism-Specific Search Engines*. In Proceedings of the 11th International Conference on Information and Communication Technologies in Tourism (ENTER 2004), pages 239-248, Cairo, Egypt, 2003.

[10] M. Haller et al.: *Integrating Heterogeneous Tourism Information in TIScover – The MIRO-Web Approach*. In Proceedings of the 7th International Conference on Information and Communication Technologies in Tourism (ENTER 2000), pages 71-80, Barcelona, Spain, 2000.

[11] *HARMONISE*, http://www.harmo-ten.info/, requested on July 28, 2006.

[12] S. Holland: *Preference Mining and Preference Repositories: Design, Algorithms and Personalized Application,* Dissertation, ISBN 3-89821-352-8, ibedim, 2004.

[13] S. Holland, W. Kießling: *Situated Preferences and Preference Repositories for Personalized Database Applications*. In Proceedings of the 23rd International Conference of Conceptual Modeling (ER 2004), pages 511-523, Shanghai, China, 2004.

[14] J. A. Howard, J. N. Sheth: *The Theory of Buyer Behaviour*. John Wiley and Sons Inc, USA, 1969.

[15] M. Kaukal, H. Werthner: *Integration of Heterogeneous Information Sources*. In Proceedings of the 7th International Conference on Information and Communication Technologies in Tourism (ENTER 2000), pages 81-92, Barcelona, Spain, 2000.

[16] W. Kießling: *Foundations of Preferences in Database Systems*. In Proceedings of the 28th International Conference on Very Large Data Bases (VLDB 2002), pages 311-322, Hong Kong, China, 2002.

[17] W. Kießling: *Preference Queries with SV-Semantics*. In Proceedings of the 11th International Conference on Management of Data (COMMAD 2005), pages 15-26, Goa, India, 2005.

[18] W. Kießling, S. Fischer, S. Döring: *COSIMA B2B – Sales Automation for E-Procurement*. In Proceedings of the 6th IEEE Conference on E-Commerce Technology (CEC 2004), pages 59-68, San Diego, USA, 2004.
[19] W. Kießling, B. Hafenrichter, S. Fischer, S. Holland: *Preference XPATH: A Query Language for E-Commerce*. In Information Age Economy. Physika-Verlag, pages 427-440, Augsburg, Germany, 2001.
[20] W. Kießling, G. Köstler: *Preference SQL - Design, Implementation, Experiences*. In Proceedings of the 28th International Conference on Very Large Data Bases (VLDB 2002), pages 990-1001, Hong Kong, China, 2002.
[21] G. Koutrika, Y. Ioannidis: *Personalized Queries under a Generalized Preference Model.* In Proceedings of the 21st International Conference on Data Engineering, pages 841-852, Tokyo, Japan, 2005.
[22] M. Missikoff et al.: *Harmonise – Towards Interoperability in the Tourism Domain*. In Proceedings of the 10th International Conference on Information and Communication Technologies in Tourism (ENTER 2003), pages 58-66, Helsinki, Finland, 2003.
[23] N. Mitsche: *Understanding the Information Search Process within a Tourism Domain-specific Search Engine*. In Proceedings of the 12th International Conference on Information and Communication Technologies in Tourism (ENTER 2005), pages 183-193, Innsbruck, Austria, 2005.
[24] A. Öörni, S. Klein: *Electronic Travel Markets: Elusive Effects on Consumer Behavior*. In Proceedings of the 10th International Conference on Information and Communication Technologies in Tourism (ENTER 2003), pages 29-38, Helsinki, Finland, 2003.
[25] J. Palkolska, A. Dunzendorfer, J. Küng: *Vague Queries in Tourism Information Systems*. In Proceedings of the 7th International Conference on Information and Communication Technologies in Tourism (ENTER 2000), pages 61-70, Barcelona, Spain, 2000.
[26] J. Palkoska et al.: *Advanced Query Mechanisms in Tourism Information Systems*. In Proceedings of the 9th International Conference on Information and Communication Technologies in Tourism (ENTER 2002), pages 438-447, Innsbruck, Austria, 2002.
[27] B. Pan, D. R. Fesenmaier: *Exploring Caller Dialogue: Analyzing Directed Information Search for Tourism Products*. In Proceedings of the 8th International Conference on Information and Communication Technologies in Tourism (ENTER 2001), pages 244-250, Montreal, Canada, 2001.
[28] B. Pan, D. R. Fesenmaier: *Travel Information Search on the Internet: A Preliminary Analysis*. In Proceedings of the 10th International Conference on Information and Communication Technologies in Tourism (ENTER 2003), pages 242-251, Helsinki, Finland, 2003.
[29] F. Pühretmair, H. Rumetshofer, E. Schaumlechner: *Extended Decision Making in Tourism Information Systems.* In Proceedings of the 3th International Conference on E-Commerce and Web Technologies, pages 57-66, Aix-en-Provence, France, 2002.
[30] F. Ricci: *Travel Recommender Systems*. In IEEE Intelligent Systems, pages 55-57, 2002.
[31] F. Ricci et al.: *Intelligent Query Management for Travel Product Selection*. In Proceedings of the 9th International Conference on Information and Communication Technologies in Tourism (ENTER 2002), pages 448-457, Innsbruck, Austria, 2002.
[32] B. Satzger, M. Endres, W. Kießling: *A Preference-Based Recommender System*. In Proceedings of the 7th International Conference on Electronic Commerce and Web Technologies (EC-Web / DEXA 2006), Crakow, Poland, 2006.
[33] M. Torrens et al.: *reality: a Scalable Intelligent Travel Planner.* In Proceedings of the 2003 ACM Symposium on Applied Computing (SAC 2003), pages 623-630, Melbourne, Florida, USA, 2003.
[34] H. Werthner: *Intelligent Systems in Travel and Tourism.* In Proceedings of the 18th International Joint Conference on Artificial Intelligence (IJCAI 03), pages 1620-1628, Acapulco, Mexico, 2003.
[35] H. Werthner, F. Ricci: *E-Commerce and Tourism.* In Communications of the ACM, Vol. 47, Issue 12, pages 101-105, 2004.
[36] L. A. Zadeh: *The Concept of a linguistic variable and its application to approximate reasoning.* Elsevier Pub. Co., New York, 1973.