Comparing the effectiveness of intentional preferences versus preferences over specific choices: a user study

Panagiotis Papadakos*

Institute of Computer Science (ICS),
Foundation for Research and Technology – Hellas (FORTH),
N. Plastira 100, Vassilika Vouton,
GR-700 13 Heraklion, Crete, Greece
Email: papadako@ics.forth.gr
*Corresponding author

Yannis Tzitzikas

Institute of Computer Science (ICS),
Foundation for Research and Technology – Hellas (FORTH),
N. Plastira 100, Vassilika Vouton,
GR-700 13 Heraklion, Crete, Greece
and
Computer Science Department,
Voutes Campus,
University of Crete,
700 13 Heraklion, Crete, Greece
Email: tzitzik@ics.forth.gr

Abstract: Many works have proposed the enrichment of database query languages with preferences, aiding the user to better rank the results. In this work we propose and examine the hypothesis that effective preference specification in many cases presupposes knowledge of the information space and the available choices. Otherwise, the expression of intentional preferences can be a tiresome process, leading to non-optimal results. We designed a user study from an information systems perspective, where participants had to express their preferences for buying a new car (intentional preferences) and then select the most preferred car from a list (preferences over specific choices). The results showed that only 20% of the users expressed intentional preferences that led to the finally selected car. The conducted statistical hypothesis testing supports the results with a 1% error. Consequently, we argue that the ability to gradually express preferences while exploring the available choices is beneficial for effectively and efficiently ranking a set of choices.

Keywords: information systems; multi criteria decision making; preferences; exploratory search; car purchase; user study; preference-enriched query languages.

Reference to this paper should be made as follows: Papadakos, P. and Tzitzikas, Y. (2016) 'Comparing the effectiveness of intentional preferences versus preferences over specific choices: a user study', *Int. J. Information and Decision Sciences*, Vol. 8, No. 4, pp.378–403.

Biographical notes: Panagiotis Papadakos is a postdoctoral scientist of the Information Systems Laboratory at FORTH-ICS (Greece). He received his PhD in Preference-based Interactive Exploration of Multidimensional Information Spaces from the Computer Science Department of the University of Crete, Greece in 2013. His main research interests are in the areas of information retrieval, databases and the semantic web, with a focus on exploratory search and preference-based interaction over faceted and dynamic taxonomies. In the past, he has worked extensively in the areas of virtual reality and 3D user interfaces. He has published more than 15 papers in peer-reviewed conferences and journals, and he has been involved in international and national projects.

Yannis Tzitzikas is currently an Assistant Professor of Information Systems in the Computer Science Department at University of Crete (Greece) and Associate Researcher of the Information Systems Laboratory at FORTH-ICS (Greece). His research interests fall in the intersection of the following areas: information systems, information indexing and retrieval, conceptual modelling, knowledge representation and reasoning, and collaborative distributed applications, with a focus on exploratory searching, semantic data management, and digital preservation. His work has been published in more than 90 papers in refereed international conferences and journals, and he has received two best paper awards (at CIA 2003 and ISWC 2007). Finally he regularly participates to the scientific committees of several international conferences and journals.

1 Introduction

In our everyday lives we express a number of preferences, even contradicting ones, for all different kind of aspects, and take a number of decisions according to them. Such preferences can be independent or might affect each other, and might change over time. In addition they might be satisfied or not, leading even to conflicts between them since they are generally not considered as hard constraints but rather as simple or complicated wishes (Stefanidis et al., 2011). Preferences have been studied in a number of fields, including computer science. Specifically, they have been thoroughly studied in information systems and especially in the database area where a number of approaches have been proposed. Those works make the assumption that the user is able to express his/her preferences using a query language in one shot, without necessarily seeing the available choices.

In this work we examine the hypothesis that effective preference specification presupposes knowledge of the information space and of the available choices. Specifically, we believe that without knowing the available choices, the declarative expression of intentional preferences can be a tiresome and time-consuming process, that leads to ineffective preference specification and non-optimal results regarding the user's preferences. For example consider the following case:

Example: Assume that James wants to change his car. He is interested in a family car, although he preferred sport cars when he was younger. His wife prefers Jeeps but he is reluctant due to the extra parking space required and because the garage of his home is somehow small. He believes that Japanese and German cars are more reliable than French or Korean cars. He likes the fact that hybrid cars consume less, are more

ecological and that the annual taxes are lower for such cars. James lives at the city of Heraklion, so cars owned by persons that are not near the city are less preferred for him due to the travelling time and cost, unless the case is exceptional. In addition he cannot afford an expensive car. Ideally he would like a Porsche with four doors (e.g., Porsche Panamera) and enough space for luggage, hybrid with consumption less than 6 lt/100 Km, bigger than Panamera (to satisfy his wife) but smaller than Cayenne, with less than ten thousands kilometres, in sale by his favourite neighbour and at a very good price (e.g., less than 30K Euros), but this is an utopian desire. James aims at buying one car, but it is probable that he would buy a 'Porsche Carrera 4s' if available at a very good price, and another decent but inexpensive family car to satisfy the rest requirements.

Although lengthy, the above description is by no means complete. There are a lot of other aspects that would determine James final decision (guarantee, grip, airbags, colour, trunk, GPS, sunroof, etc.). What we want to stress with this example is that the specification of preferences is a laborious, cumbersome and time-consuming task, and that the resulting descriptions are in most cases *ineffective*. Pragmatically, decision making is based on complex trade-offs that involve several (certain or uncertain) attributes as well as user's attitude towards risk (Keeney and Raiffa, 1976). Moreover preferences are not stable over time (Doyle, 2004) and users change their preferences even during the inspection of available choices.

As a result, frameworks that capture preferences should not assume that preferences are fixed or can be given in one shot. Specifically, we believe that it is beneficial to provide users with an interaction method in which the preference specification cost is paid *gradually* and *depends on the available choices* (i.e., a 'pay as you go' style). Returning to our example, even if the user managed to express all the above preferences, he should not spend time for expressing complex trade-offs between Porsche models with four or two doors, when no Porsche car is available.

We conducted a user study for evaluating correctness, completeness, validity and effort of intentional preferences (i.e., preferences expressed by the user without knowing the available choices) versus extensional preferences (i.e., preferences over specific choices). The user study was related to the common decision task of car purchase, where initially users expressed their preferences intentionally and then they selected the most preferred car from a list of cars. In this context, preference completeness holds if a clear winner over the set of cars can be induced by the expressed intentional preferences (i.e., user's preference return only one top rated car), while preference correctness holds if preference completeness holds and the selected by the user car is the same as the first car induced by the intentional preferences (i.e., the top rated car is the selected car). The results of the user study and of a theoretical analysis strongly support our hypothesis.

The focus of this work is on decision tasks where the choices are described by *numerous* attributes, their value-set may not be known entirely by the user or may not be known apriori (i.e., values that are dynamically mined), and might not be single-valued. We use Pareto domination, prioritisation and simple rules to aggregate users' preferences and rank the objects, as implemented in the available preference-based frameworks and query languages (Kießling, 2002; Chomicki, 2003; Kießling and Kostler, 2005).

Note that a number of other methods and models have been proposed in the multi-criteria decision making and analysis bibliography, such as utility-based models (MAUT, OWA, Choquet integral), outranking models (ELECTRE or Promethee methods), decision rule methods, AHP, or other hybrid approaches like Sakthivel et al.

(2013) [see Figueira et al. (2005) for a survey]. For example the well known AHP, which is a very successful method in certain applied areas, is based on pairwise comparisons in order to calculate the weights for each criterion and derive the final ranking. But according to Utkin (2014), it is often difficult to validate or to justify the optimal solutions, since it is an adhoc method, and the criteria need to have identical numerical scales. Moreover, AHP does not behave well when only partial information about the weights of the criteria is available, and the time and cost of assessing the weights and scores increase significantly with the number of criteria. On the other hand, Pareto domination has been studied extensively in the database community, where a number of efficient algorithms have been proposed (i.e., skyline algorithms). Such algorithms calculate the Pareto sets, offering a good behaviour regarding recall (since the best objects with respect to multiple monotonic functions can be found among them), while prioritisation offers a way to increase precision by reducing the corresponding sets.

The rest of this paper is organised as follows. Section 2 expresses our hypothesis, and Section 3 analyses theoretically some aspects of this hypothesis. Section 4 describes the scenario of the user study for testing the hypothesis, and then Section 5 presents the results. A statistical hypothesis test over the gathered data of the evaluation is conducted in Section 6. Section 7 discusses the possible consequences of our results for interaction schemes that allow the expression of preferences. Finally, Section 8 concludes the paper and identifies issues for future research.

2 Context and hypothesis

We consider decisions over a number of choices, each described by a number of arithmetic or categorical attributes. Attributes can be either single-valued or multi-valued. The task of selecting the preferred choice over a set of choices is a common decision problem (e.g., products are described by a number of attributes) and the examined task of *car selection* is an instance of such a decision problem.

Let $\mathcal{V}$ be the set of all possible choices that are described by a set of attributes $\mathcal{A}$. If the preferences of a user are expressed *intentionally* over the values of the attributes in $\mathcal{A}$, then they can be useful for ordering according to preference *any* set of choices $\mathcal{C}$, where $\mathcal{C} \subseteq \mathcal{V}$. However, complete and correct intentional preferences may be too hard to obtain. In our specific scenario, a complete and correct intentional expression of the user preferences would mean that this expression would allow to select correctly among any subset of cars the most desired car for the user. Unfortunately, and as we shall see in the user study, this does not hold.

Our main hypothesis, which for short we shall hereafter call *intentional preferences difficulty* (*IPD*), can be expressed as follows.

2.1 IPD hypothesis

Without the ability to view and explore the different aspects of the available choices, the expression of intentional preferences by a human can be tiresome and time-consuming, leading in most cases to non-optimal results according to preference. As a result, the expressed preferences are not sufficient for selecting the most desired option from a particular set of choices. ◊

This hypothesis does not concern decision problems over choices described by a few arithmetic attributes (e.g., a job selection task, where each job is described by the offered salary and the required number of working hours per week). It mainly concerns decision tasks where the choices are described by *numerous* attributes, their value-set may not be known entirely by the user or may not be known apriori (i.e., values that are dynamically mined), and might not be single-valued. This is actually the case for most real-world decision tasks. Below we try to analytically discuss and compare the completeness, correctness and effort of expressing intentional and extensional preferences.

3 Analytical comparison

3.1 Notation

Let $\mathcal{A}_1$, ..., $\mathcal{A}_n$ be the n attributes that describe the set of choices $\mathcal{S}$ (e.g., Table 1 in Section 4 shows different attributes for our scenario including `price`, `horsePower`, `fuel`, etc.) and let $dom(\mathcal{A}_i)$ denote the set of values that $\mathcal{A}_i$ can take $\forall i \in \{1, ..., n\}$ (e.g., $dom(\texttt{fuel}) = \{gas, diesel\}$). We will denote by $\mathcal{V} = X_{i=1}^{n}\ dom(\mathcal{A}_i)$ the Cartesian product of the domains of the n attributes, i.e., $\mathcal{V} = X_{i=1}^{n}\ dom(\mathcal{A}_i) = dom(\mathcal{A}_1) \times ... \times dom(\mathcal{A}_n)$.

A preference over a set of elements E can be expressed as a binary relation over the elements of E. So we hereafter assume that a preference relation is a binary relation $(E, \succ)$. If we have n attributes, then we could define $n + 2$ different preference relations, one relation over the n attributes expressing our priorities, n relations over the domains of these attributes expressing our preferences over the values of a specific attribute, and one relation over the objects we want to rank (e.g., cars). For example `Price` $\succ$ `Manufacturer` denotes a preference (prioritisation) of the `Price` attribute over the `Manufacturer` attribute, while `5 Doors` $\succ$ `3 Doors` denotes a preference over the domain of the attribute `doors`. Finally, `Toyota Yaris` $\succ$ `Fiat Punto` denotes a preference over two particular cars.

In our context we will use the term *intentional preference* to refer to those preferences that a user expresses without viewing and exploring the available choices (i.e., it refers to a preference relation over the whole Cartesian product $\mathcal{V}$). On the other hand, we will use the term *extensional preference* to refer to those preferences that are expressed as the user explores a specific set of choices $\mathcal{C}$. Specifically, we will denote with $(\mathcal{V}, \succ_{\mathcal{I}})$ an intentionally specified preference relation regarding $\succ$ over $\mathcal{V}$, while with $(\mathcal{C}, \succ_{\mathcal{E}})$ we will denote an extensionally specified preference relation regarding $\succ$ over the specific set of choices $\mathcal{C}$, where $\mathcal{C} \subseteq \mathcal{V}$.

3.2 Completeness of intentional and extensional preferences

We can consider that a *complete intentional preference relation* aims at defining a total order of the elements of $\mathcal{V}$ and will be denoted with $(\mathcal{V}, \succ_{\mathcal{I}}^{\mathcal{T}})$. Respectively, a *complete extensional preference relation* aims at defining a total order of the elements of $\mathcal{C}$ and will be denoted with $(\mathcal{C}, \succ_{\mathcal{E}}^{\mathcal{T}})$.

3.3 Correctness and k-correctness of intentional preferences

Consider that we have a set of choices $\mathcal{C}$ and that the user has defined a $(\mathcal{C},\succ_{\mathcal{E}}^{\mathcal{T}})$. Then we can define *Correctness* as:

Definition 1: Correctness

We say that a $(\mathcal{V},\succ_{\mathcal{I}}^{\mathcal{T}})$ is *correct* with respect to $(\mathcal{C},\succ_{\mathcal{E}}^{\mathcal{T}})$, if the restriction of $(\mathcal{V},\succ_{\mathcal{I}}^{\mathcal{T}})$ on $\mathcal{C}$ is equal to $(\mathcal{C},\succ_{\mathcal{E}}^{\mathcal{T}})$, i.e., $(\mathcal{V},\succ_{\mathcal{I}}^{\mathcal{T}})|_{C}=(\mathcal{C},\succ_{\mathcal{E}}^{\mathcal{T}})$. □

If $(\mathcal{C}, \succ^{\mathcal{T}})$ is a preference relation that defines a total order of the elements of $\mathcal{S}$ with respect to $\succ$, then we shall use $(\mathcal{C}, \succ^{\mathcal{T}})$ $[k]$ to denote the first k elements of this total order. We can now define *k-correctness* as:

Definition 2: k-correctness

$(\mathcal{V},\succ_{\mathcal{I}}^{\mathcal{T}})$ is k-Correct with respect to $(\mathcal{C},\succ_{\mathcal{E}}^{\mathcal{T}})$, if the restriction of $(\mathcal{V},\succ_{\mathcal{I}}^{\mathcal{T}})[k]$ on $\mathcal{C}$ is equal to $(\mathcal{C},\succ_{\mathcal{E}}^{\mathcal{T}})[k]$, i.e., $(\mathcal{V},\succ_{\mathcal{I}}^{\mathcal{T}})|_{C}\,[k]=(\mathcal{C},\succ_{\mathcal{E}}^{\mathcal{T}})[k]$. □

Since most of the times a user is interested in only one object (i.e., he wants to buy only one car), we will consider in our work *1-Correctness*, i.e., $k = 1$.

3.4 Effort required for expressing complete preferences

The required effort for expressing complete intentional preferences can be roughly quantified with the amount $|\mathcal{V}|$. Respectively, the effort required for expressing extensional preferences can be quantified with the amount $|\mathcal{C}|$. For instance, if we have only one attribute that takes two values and the collection consists of only two objects, then $|\mathcal{V}| = |\mathcal{C}| = 2$. In this case, it is equally laborious to express preferences either intentionally or extensionally. On the other hand, if we have ten attributes, each having ten possible values and again two objects, then $|\mathcal{C}| = 2$ while $|\mathcal{V}| = 10^{10}$. Since the *IPD* hypothesis of Section 2 concerns decision problems with numerous attributes, which can be either single-valued or multi-valued, we expect that in the previously described decision making tasks $|\mathcal{V}| \gg |\mathcal{C}|$, meaning that it is much more laborious to express preferences intentionally rather than extensionally.

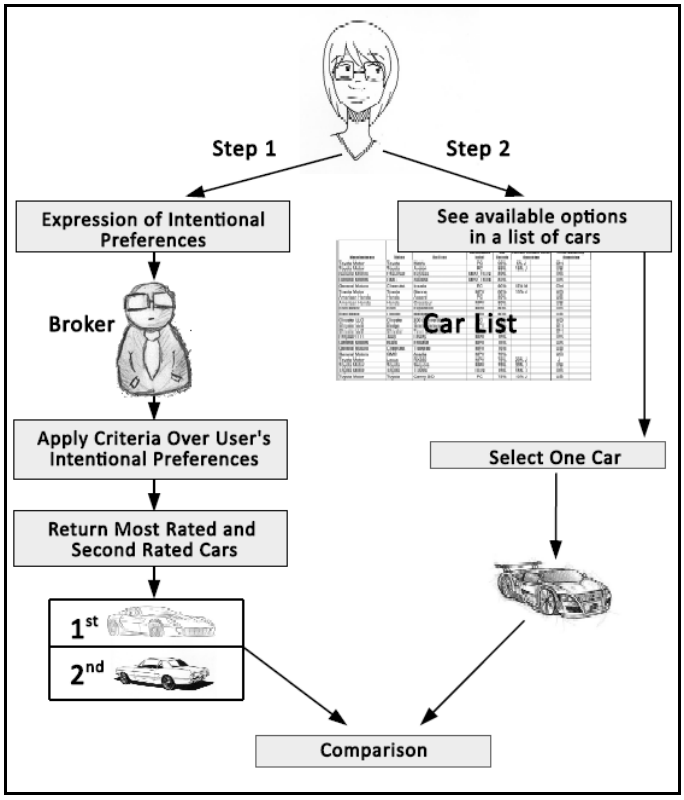
The above specified costs do not aim at being accurate; they aim to capture the main point. One could easily refine the costs according to various aspects. Such an aspect could be the type of the attribute values. Specifically, for a categorical attribute $\mathcal{A}_i$ we can define $Cost(\mathcal{A}_i) = |dom(\mathcal{A}_i)|$. Else, if the domain is arithmetic we can define $Cost(\mathcal{A}_i) = 1$. The latter holds because in arithmetic attributes such as `price`, `horsePower`, or `fuelConsumption`, most of the times the user does not have to inspect the available values. He just has to express whether he prefers either the highest or lowest values, those around a specific value, or finally those values that belong in a user specified interval. On the contrary, in categorical attributes such as `bodyType`, `brand`, or `colour`, the user has to express his/her preference on the specific values of the attributes. Based on the above perspective, the cost for specifying complete intentional preferences could then be defined as $Cost(\mathcal{A}) = Cost(\mathcal{A}_1) \cdot \ldots \cdot Cost(\mathcal{A}_k)$ (note that $Cost(\mathcal{A}) \leq \Pi_{i=1}^{k} \,|\, dom(\mathcal{A}_i) \,|$).

4 Scenario of the user study

In this work we do not focus on defining refined cost models, but on testing the main hypothesis with real users and real tasks. In addition, we should note, that in decision tasks humans mainly have to select the most desired element (i.e., the car to buy, the hotel to book) instead of ordering the entire list of available options. As a result, in our user study we will only consider *1-Correctness*. Finally, in order to be independent from system issues we decided to conduct the user study using paper forms, since the expression of preference-based queries is a quite complicated task for users. To test the *IPD* hypothesis the user study was based on the very common and familiar decision task of *car selection*. This task involves a not so big number of relevant attributes with values quite known to most persons. The study involved 30 persons, 18 male and 12 female, from seven EU countries with ages ranging from 22 to 75 years old.

The study consisted of two parts, *Step 1* and *Step 2*. Each participant had to complete both parts. In *Step 1* users were asked to express their intentional preferences for the car they would buy at that time, while in *Step 2* they were asked to select the most preferred car from a particular list of cars. Figure 1 depicts the whole process of the user study. Below we describe the two steps in more details.

Figure 1 Description of the user study process



4.1 Step 1: formulation of intentional preferences

In the first step, all participants were asked to express their preferences according to the task described below.

4.1.1 Step 1: task

Suppose that you have (it is obligatory) to change your car. You have to select and buy a new one, and of course *you will have to pay it*. Please express your preferences on paper. This paper will be handed to a different person who has at his/her disposal a limited collection of available cars. This person will select one for you based on the available cars and the preferences that you expressed.

You have 30 minutes at most to express your preferences. You are free to express them in any form you like, e.g., in natural language text (e.g., 'I prefer a car with an engine volume between 1,200 and 1,400 cc'), by providing an ordering according to preference, by specifying the preferred (ideal) price or price interval, etc.

4.2 Step 2: extensional preferences – selection from specific choices

Immediately after completing the first step and in order to avoid user preferences alteration, participants continued with the second step. In this step they were given a list of cars and were asked to identify the most preferred car for them. The list consisted of 50 cars (shown in the Appendix). Its schema is described in Table 1.

Table 1 Example of using the schema of the car collection for representing one particular car

Attribute	*Value*	*Attribute (cont.)*	*Value*	*Attribute (cont.)*	*Value*
ID	29	Power (HP)	174	Netto weight kg.	2,128
Model	Nissan	Torque (NM)	403	Weight kg.	2,805
Modification	Navara	Max. speed (km/h)	170	Trunk l.	1,260
Year	2008	Acceleration 0–100 km/h (sec.)	11.5	Fuel tank l.	80
Body type	Pickup	Consumption (city) (l/100 km)	11.1	Fuel	Diesel
Doors	4	Consumption (highway) (l/100 km)	7.1	Price EUR	22,190
Engine volume	2,488				

5 Analysis of the results of the user study

At first, in Section 5.1 we discuss how the results of *Step 1* and *Step 2* were compared regarding the correctness and completeness of the intentionally expressed preferences. In Section 5.2 we provide some general observations regarding the formulation time, the type, the validity, and the applicability of intentionally expressed preferences.

5.1 Correctness and completeness of intentional preferences

The essential objective is to check whether the paper-written intentional preferences of *Step 1* would allow someone to obtain the car selected in *Step 2*. In case the answer is positive then it means that the preference expression on paper was complete and correct in order to select the most preferred car (at least for the specific set of cars of *Step 2*). If the answer is negative then the conclusion would be that the participants did not manage to express their preferences in a sufficient way, in order to get the most desired car from the small list of available cars.

To check the correctness and completeness of the intentionally specified preferences, a *broker* was given the filled forms from *Step 1*. For each participant the broker divided the provided preferences into two categories: *specific* and *general*. The former are preferences that use specific values of the attributes domain or impose a specific ordering, e.g., 'I prefer red to yellow cars' or 'I want a car with a displacement between 1,200 cc and 1,400 cc'. The latter are preferences that do not use specific values and do not impose a specific ordering, e.g., 'I want a cheap car' or 'I want a car that does not pollute the environment'.

To derive the ranking of the cars according to preference, the broker aggregated user's preferences and applied the criteria described below. Firstly, the broker considered only the *specific preferences* (i.e., preferences that induced a specific ordering) and applied (*Criterion 1*), simulating the process of a preference-enabled information system. In case of ties, *general* preferences were transformed into corresponding *specific* preferences and the broker applied the same process in each bucket with tied cars (*Criterion 2*). Finally, if ties were still present, the broker selected the car with the highest number of satisfying preferences (i.e., wins) (*Criterion 3*), returning the cheaper one if two or more cars were found tied (*Criterion 4*). Using the last two criteria the broker could produce a total order of the cars. In more details:

Criterion 1: Specific preferences criterion (SPC)

The broker considers only the specific preferences applicable to the car collection at hand. If the expressed preferences are prioritised (e.g., 'I can only buy a car that costs no more than 10,000 Euros. I would prefer in this price range, a car with a displacement between 1,200 cc and 1,400 cc'), cars are ranked according to the most prioritised preference, then according to the second most prioritised preference, and so on (*prioritised composition*). Preferences having the same priority are considered equal and *Pareto composition* is used to remove those cars dominated by others. □

Criterion 2: General preferences criterion (GPC)

If ties exist after criterion 1, they are resolved by using the available general preferences (if applicable to our collection). Specifically, each *general preference is transformed to a specific preference*. For example the preference 'I want a cheap car', can be transformed to ordering the cars of each bucket according to their price, or 'I want a car that does not pollute the environment' can be transformed to ordering the cars in ascending order according to their fuel consumption. Again priority and Pareto compositions are used to derive the final ordering of the cars. □

Criterion 3: Wins criterion (WPC)

If ties still exist after Criterion 2, cars in each bucket are ordered according to the *number of satisfying preferences*. □

Criterion 4: Cheapest criterion (CC)

If ties still exist after Criterion 3, the broker puts the cars of each bucket *in ascending order according to their price*. □

Table 2 Excerpt from the results table

User										
User information				*Step 1*				*Step 2*		
UID	*Age*	*Gen.*	*Cntr*	$\|P_S\|$	$\|P_G\|$	$\|P\|$	T_1	*CIDu*	*Grade*	T_2
33	23	F	GR	3	8	11	20 m.	46	{0}{×✓}	3 m.
Broker										
Ideal				*2nd Ideal*				*1st vs. 2nd*		
CID_1		*Grade*		CID_2		*Grade*		*Wins in*		
43		{1}{✓✓}		28		{2}{✓o}		SPC		
Overview										
Br vs. Us		*Inconsistent*			*Unused*			*Correct*		
Wins in		$\|P_S\|$	%		$\|P\|$	%		C_1	C_2	
SPC		0	0		5	45.4		×	×	

The detailed results are shown in Table 3. An indicative, and easily readable, excerpt is shown in Table 2 and is described below. The first set of columns of Table 2 is related to user information, and the two steps of the case study. Specifically, the first columns show information about each user (id, age, gender and country). Columns in [*Step* 1] show the number of *specific*, *general* and *total* preferences ([$|P_S|$] = 3, [$|P_G|$] = 8 and [$|P|$] = 11 resp.), as well as the time the user spent in *Step 1* ([T_1] = 20 m.). Columns in [*Step* 2] show the id of the user selected car and the time the user spent in *Step 2* (i.e., [*CID*] = 46 and [T_2] = 3 m. resp.). Additionally, the car has a [*Grade*], a preorder over the different preference grades, indicating which *specific* preferences of a particular user this car satisfies, as well as their priority. Each preference grade can take the following values:

- ✓ if the preference is satisfied for this car (for binary preferences of the form 'I prefer a car with ABS').
- × if the preference is not satisfied for this car (again for binary preferences).
- A number if this car satisfies the corresponding value from an ordered set of preference values (i.e., 1 for the first, 2 for the second, etc.). For example, if *Fiat* $\succ$ *Audi* and *Audi* $\succ$ *Mercedes*, then a car made by *Fiat* has value 1, a car made by *Audi* has value 2, a car made by *Mercedes* has value 3, etc. The rest of the cars made by other manufacturers, which are inactive elements are considered as the worst (value is 0).
- o if the preference is not applicable for this car.

When there is no preference priority, preferences are considered indifferent (i.e., equal), and *Grade* has the form $\{G_1 ... G_n\}$. On the other hand, if a preference P_1 is prioritised over P_2, which is prioritised over the rest preferences, then *Grade* has the form $\{G_1\}\{G_2\}\{G_3 ... G_n\}$. In the example of Table 2, P_1 is prioritised over P_2 and P_3 (P_2 and P_3 have equal priority). Note that for space and readability reasons we provide only grades regarding *specific* preferences only.

The second set of columns concern the process of the broker, who returns the first and second most rated cars according to the intentional preferences of the user and the previously described four criteria. For each of the two cars it shows their *ids* ([*CID*]) and their corresponding grades. The [1*st vs.* 2*nd wins*] column reports in which criterion the first rated car won the second rated car, taking values from {*SPC*, *GPC*, *WC*, *CC*}.

Finally, the last seven columns provide an overview of the results. Column [*Br vs. Us wins in*] reports the criterion in which the broker was able to discriminate the first rated car to the user selected car, and takes values from {*SPC*, *GPC*, *WC*, *CC*, –} ('–' means that the broker and the user selected the same car). Column [*inconsistent*] shows the number ([P_S]) and the percentage ([%]) of the intentionally *specific* preferences expressed in *Step 1* that were not consistent with the user selected car (e.g., choosing a car with a displacement of 1,800 cc when the user had intentionally specified a preference for an engine less than 1,400 cc). Column [*unused*] holds the number and percentage of intentionally *total* preferences that were non-applicable to our collection and thus not used. Notice that some preferences might be applicable to specific cars only, since some cars have incomplete descriptions. Finally, column [*correct*] marks if the user selected the same car as the first ([C_1]) or the second ([C_2]) rated car by the broker respectively.

5.1.1 Discussion of the results

The results support our hypothesis that without exploring the existing choices, the expression of preferences leads to non-optimal preferences, since only six out of the 30 participants (20%) selected the most preferred car in *Step 2*, according to their intentionally expressed preferences in *Step 1*.

Notice that in the 24 cases that there was no agreement between the user selected car and the first car returned by the broker, the broker's car won the user selected car:

a 58.3% during the *specific* preference criterion phase (*SPC*)

b 20.83% during the *general* preference criterion phase (*GPC*)

c 12.5% during the *wins* criterion (*WPC*)

d 8.33% during the *cheaper* criterion phase (*CC*).

As a result the car returned by the broker was clearly preferred according to the user's intentionally specified *specific* and *general* preferences over the user selected car.

Taking into consideration the second ideal car returned by the broker, the number of the participants that selected either the first or the second car returned by the broker raises to 10 (33.3%). Notice though, that 75% of the second ideal cars (three out of four cars) that matched the user selected car, lost from the ideal car during the *specific* preference criteria phase (*SPC*). This shows that again the ideal car is clearly preferred to the second ideal one, according to the expressed *specific* preferences by the user. The rest one car lost during the cheaper criterion phase (*CC*). The above results indicate that the user was

not able to efficiently specify his/her preferences on paper or that the user changed his/her preferences while exploring the available choices. Thus we can conclude that the user is not able to provide *correct* intentional preferences.

Regarding the 1*st* vs. 2*nd wins* column, 33.3% of the times the broker was able to discriminate the ideal from the 2nd ideal car during the *SPC* phase, 40% during the *GPC* phase and 26.6% during the *CC* phase. As a result, if we consider only the *specific* preferences, 33.3% of the users were able to express preferences that could lead in a clear winner in our experiment, while if we consider both the *specific* preferences and the transformations of the *general* preferences to *specific* preferences, the percentage raises to 73.3% of the users. Consequently, it seems that most of the users are able to provide *complete* intentional preferences (i.e., preferences that could provide a clear winner), but as previously shown without being *correct*, meaning that for the set $\mathcal{C}$ of 50 cars used in the evaluation, for 80% of the cases $(\mathcal{V}, \succ_{\mathcal{I}}^{\mathcal{T}})|_{C}[1] \neq (\mathcal{C}, \succ_{\mathcal{E}}^{\mathcal{T}})[1]$.

5.2 Time, validity and applicability of intentional preferences

Here we evaluate the intentionally specified preferences according to various aspects, like time for completing each task, type and number of preference actions, and finally validity and applicability of preference actions.

5.2.1 Time

According to the results shown in Table 3, participants spent on average ten minutes in *Step 1*. The worst case was a user that used the whole 30 minutes time slice. Despite the fact that users spend a lot of time formulating their preferences, this process led to a number of inconsistent and non-effective preferences regarding the very small set of 50 choices. Even for the previously mentioned user that took the whole 30 minute time slice, there was a disagreement between the first car returned by the broker and the user selected car (though this was not true for the second car returned by the broker). On the other hand, participants spent only four minutes in *Step 2*, in order to find the most preferred car from the list of 50 cars. However, if the list contained thousands of cars, the participants would have spent much more time to pick the most preferred car. In such cases it is beneficial to exploit an *information thinning* approach, enriched with preference actions over the attributes of the specific dataset, as the one discussed in Section 7.

Table 3 Results of the hypothesis evaluation

User										
User information				*Step 1*				*Step 2*		
UID	*Age*	*Gen*	*.Cntr*	$\lvert P_S \rvert$	$\lvert P_G \rvert$	$\lvert P \rvert$	T_1	*CIDu*	*Grade*	T_2
1	32	M	GR	6	3	9	4 m	40	{1✓✓o✓✓}	2 m
2	25	F	GR	4	5	9	5 m	46	{oo✓✓}	3 m
3	26	F	GR	7	0	7	7 m	31	{✓4✗✓oo✓}	1 m

Table 3 Results of the hypothesis evaluation (continued)

User										
User information				*Step 1*				*Step 2*		
UID	*Age*	*Gen*	*.Cntr*	$\|P_S\|$	$\|P_G\|$	$\|P\|$	T_1	*CIDu*	*Grade*	T_2
4	27	M	GR	3	4	7	5 m	46	{✓✓✓}	3 m
5	26	M	GR	5	1	6	6 m	46	{✓✗✗✓✗}	2 m
6	30	F	FR	6	1	7	10 m	47	{✗✓✓✗✗✗o}	5 m
7	32	M	AT	8	1	9	10 m	16	{✓✓o✗✗✗✓}	3 m
8	33	F	EE	3	5	8	7 m	48	{o✗✓}	1 m
9	33	M	NO	4	7	11	10 m	45	{✗✓✓✓}	3 m
10	31	F	BG	4	6	10	15 m	34	{✓✓✓✓}	15 m
11	32	M	In	10	0	10	7 m	4	{✓✗}{✗o✓ooo✓o}	5 m
12	31	M	GR	7	3	10	5 m	40	{✓✓}{✓✓✓✓1}	2 m
13	29	M	GR	11	0	11	10 m	34	{oo✗oooo✓✓✓✓}	5 m
14	28	M	GR	10	1	10	10 m	14	{oo✓✓✓✓✓o✗✓}	3 m
15	37	F	GR	15	6	21	13 m	49	{✗✓✓✗ooooooooooo}	4 m
16	23	F	GR	3	8	11	20 m	46	{o✓✓}	3 m
17	76	M	GR	14	1	15	15 m	46	{ooooooo✓✓o✗✓✗✗}	5 m
18	44	M	AT	5	5	10	30 m	17	{o✓✗✓✗}	30 s.
19	25	M	GR	5	6	11	5 m	46	{✓✗✗✓✓}	2 m
20	42	M	GR	10	4	14	14 m	46	{1✓✓✗✗}	18 m
21	47	F	GR	3	2	5	15 m	15	{✓✗✗}	1 m
22	46	M	GR	8	0	8	5 m	30	{✓o✓✓✓✓✓✗✓}	1 m
23	22	M	GR	9	0	9	8 m	21	{o✓✗✓✓✓✗✓✗}	1 m
24	26	F	GR	4	0	4	5 m	30	{o✓✓✓}	5 m
25	30	F	GR	5	12	17	10 m	36	{ooo✗✓}	5 m
26	26	F	GR	9	0	9	7 m	31	{oooo✗✗✗✓✓}	1 m
27	25	M	GR	5	4	9	15 m	46	{✗✓✓✗✗}	5 m
28	32	M	GR	9	0	9	12 m	32	{✓✓✗✗o✓✓✗✗}	5 m
29	26	F	GR	6	4	10	10 m	48	{✗2✗✓✓✗}	5 m
30	27	M	GR	3	3	6	5 m	4	{o✓✓}	5 m
Average values				6.67	3.06	9.73	10 m			4 m

Table 3 Results of the hypothesis evaluation (continued)

Broker				
Ideal		*2nd Ideal*		*1st vs. 2nd*
CID₁	*Grade*	*CID₂*	*Grade*	*Wins in*
40	{1✓✓o✓✓}	4	{3✓✓✓✓✓}	GPC
46	{oo✓✓}	40	{oo✓✓}	GPC
30	{✓4✓✓oo✓}	38	{✓0✓✓oo✓}	SPC
46	{✓✓✓}	30	{✓✓✓}	GPC
15	{✓✗✗✓✓}	16	{✓✗✗✓✓}	CC
10	{✗✓✓✗✓✗o}	32	{✗✓✓✗✓✗o}	CC
34	{✓✓✓o✓✓✓}	31	{✓✓✓o✓✓✓}	CC
48	{o✗✓}	4	{o✗✓}	CC
42	{✗✓✓✓}	41	{✗✓✓✓}	CC
30	{✓✓✓✓}	35	{✓✓✓✓}	GPC
6	{✓✗}{✗o✗ooo✗o}	4	{✓✗}{✗o✓ooo✓o}	SPC
40	{✓✓}{✓✓✓✓1}	42	{✓✓}{✓✓✗✓2}	SPC
11	{oo✓oooo✓✓✓✓}	16	{oo✗oooo✓✓✓✓}	SPC
26	{oo✓✓✓✓✓o✓✓}	43	{oo✓✓✓✓✓o✓✓}	GPC
33	{✓✓✓✓oooooooooo}	50	{✗✓✓✓oooooooooo}	SPC
43	{✓✓✓}	28	{✓✓✓}	GPC
31	{ooooooo✓✓o✗✓✗✓}	46	{ooooooo✓✓o✗✓✗✗}	SPC
37	{o✓✓✓✗}	17	{o✓✗✓✗}	SPC
15	{✓✗✓✗✓}	14	{✓✗✓✗✓}	GPC
20	{1✓✓✓✓}	17	{1✗✓✓✓}	SPC
20	{✗✓✓}	17	{✗✗✓}	SPC
30	{✓o✓✓✓✓✓✗✓}	46	{✓o✗✗o✗✓✗✓}	SPC
47	{o✗✗✓✓✓✓✓✓}	5	{o✓✗✓✓✓✓✓✓}	CC
46	{o✓✓✓}	30	{o✓✓✓}	CC
30	{ooo✓✓}	16	{ooo✓✓}	GPC
16	{oooo✓✓✓✗✗}	30	{oooo✓✗✓✗✓}	CC
16	{✓✓✓✓✓}	32	{✓✓✓✓✓}	GPC
30	{✓✓✗✗o✓✓✗✓}	39	{✓✗✓✗o✓✓✗✗}	GPC
14	{✗0✓✗✓✓}	39	{✗0✓✗✓✓}	GPC
49	{o✓✓}	30	{o✓✓}	GPC

Table 3 Results of the hypothesis evaluation (continued)

Overview						
Br vs. Us	*Inconsistent*		*Unused*		*Correct*	
Wins in	$\|P_S\|$	*%*	$\|P\|$	*%*	C_1	C_2
-	0	0%	0	0%	✓	×
-	0	0%	5	55.5%	✓	×
SPC	1	14.2%	2	28.5%	×	×
-	0	0%	0	0%	✓	×
SPC	4	60%	0	0%	×	×
SPC	3	50%	1	14.2%	×	×
SPC	3	42.8%	1	11%	×	×
-	1	33.3%	5	62.5%	✓	×
CC	1	25%	7	63.5%	×	×
GPC	0	0%	5	50%	×	×
SPC	2	20%	5	50%	×	✓
-	0	0%	0	0%	✓	×
SPC	1	9.09%	6	54.5%	×	×
SPC	1	10%	3	27.2%	×	×
SPC	2	13.3%	14	82.3%	×	×
SPC	0	9%	5	45.5%	×	×
SPC	3	21.4%	7	46.6%	×	✓
SPC	2	40%	4	40%	×	✓
WC	2	40%	2	18.8%	×	×
SPC	2	20%	7	50%	×	×
WC	2	66.6%	0	0%	×	×
-	1	12.5%	1	12.5%	✓	×
GPC	3	33.3%	1	11.1 %	×	×
CC	0	0%	1	25%	×	✓
SPC	1	20%	10	58.8%	×	×
WC	3	37.5%	4	44.4%	×	×
SPC	3	60%	1	11.1%	×	×
GPC	4	44.4%	3	33.3%	×	×
GPC	3	50%	3	30%	×	×
GPC	0	0%	1	16.6%	×	×
Average values	1.6	24.4%	3.46	35.5%		Sum
Total number					*6*	*4*

5.2.2 Type of expressed preferences

Each participant expressed on average 9.73 preferences; 6.67 of them were *specific* and the rest 3.06 were *general* preferences. All users provided *specific* preferences regarding the price attribute. Only two out of the 30 participants provided a prioritised list of preferences and these users were able to select one of the two cars returned by the broker. As a result, it seems that the prioritisation of preferences, although effective, is a difficult task for users when expressing intentional preferences (such a process could again be benefited by letting the user explore the available options as described in Section 7).

5.2.3 Validity and applicability of the expressed preferences

On average, 24.4% of the *specific* preferences were *Inconsistent*, meaning that the user selected a car that did not satisfy this *specific* preference (e.g., selecting a car using diesel, when the user prefers gas to diesel). Moreover, on average 35.5% of the total preferences per user were *unused* (i.e., they were not considered by the broker), since they were either not valid in our collection, or expressed subjective preferences that could not be evaluated (e.g., 'I prefer a beautiful car').

5.2.4 Summary

Overall, the results show that despite the fact participants spent a lot of time expressing intentional preferences that were found *complete* (i.e., provided a clear winner in our collection), they were not *correct*. Specifically, for 80% of the users intentional preferences led to non-optimal results, by either being non-applicable or inconsistent with the final choice of the user. Finally, there is a cue that spending time to express prioritised intentional preferences can lead to more effective intentional preferences. It seems though that such prioritised preferences might be difficult to express without an overall perspective over the available choices (only 6.67% of the users expressed prioritised intentional preferences).

6 Hypothesis statistical testing

We conducted a statistical significance test to check the randomness of our results. A suitable test for our case is a *one-tailed* (*lower-tailed*) *binomial significance test*, since we have dichotomous data, observations are independent from each other, probabilities of success and failure are constant across trials (i.e., all participants can select the ideal car with the same probability) and the critical region falls at one end of the possible values [i.e., we are only interested in the direction that users did not found the ideal car (Griffiths, 2009)]. The null hypothesis H_0 is:

Hypothesis H_0 *Half or more* of the users expressed their preferences without exploring available cars and were returned the most preferred car for them from a car collection.

Notice that we examine a relaxed version of the hypothesis (i.e., more than 50% of the users instead of more than 80% of the users for example) in order to not be biased against it. The alternative hypothesis H_1 is:

Hypothesis H_1 Less than half of the users expressed their preferences without exploring available cars and were returned the most preferred car for them from a car collection.

In our scenario, ideally, the user selected car and the first car returned by the broker should be the same for more than half of the cases. If Y is the number of successes in n trials, then the probability of getting Y successes in n trials is due to the binomial distribution (Griffiths, 2009):

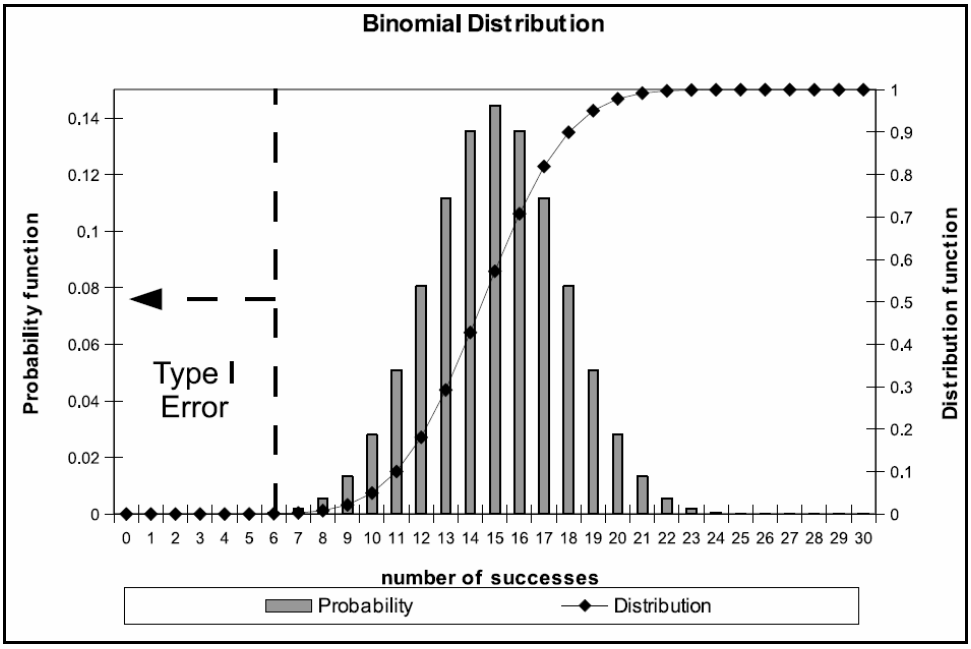
$$P(Y = y) = \binom{n}{y} \cdot p^y \cdot (1-p)^{n-y} = \frac{n!}{y! \cdot (n-y)!} \cdot p^y \cdot (q)^{n-y}$$

where p is the probability of success and $q = 1 - p$ the probability of failure. In our case $n = 30$, $p = 0.5$ and $q = 0.5$. We need to check what is the probability that only six or less participants select the desired car for them, which is calculated by:

$$P(X \le 6) = \sum_{i=0}^{6} P(X = i)$$

This formula provides the *p-value* (the probability of obtaining a test statistic at least as extreme as the one that was actually observed). Figure 2 shows the probabilities of the binomial distribution for different number of successes, the cumulative distribution function for these probabilities and the Type I error area (i.e., rejecting falsely a true null hypothesis). Since the user study results show that only six out of 30 users found the first car returned by the broker, the Type I error area is the sum of the probabilities in the area with less than six number of successes.

Figure 2 Probabilities and distribution function of the binomial distribution



Regarding the significance level, according to Wasserman (2004) an α value of 0.05, which is commonly used in the bibliography, provides a strong evidence against H_0, while a value of less than 0.01 provides a very strong evidence against H_0. In our case we used an α value of 0.01. The α value determines the risk of a Type I error (false positive). We used *R* to calculate the above probability. Specifically,

```
binom.test(6,30,0.5, alternative = "less")
```

returned a *p-value* = 0.0007155 ≤ α = 0.01. This means that we have a *very strong* evidence against H_0. So we can reject the null hypothesis H_0 and we can conclude that: "Less than half of the users can express their preferences without exploring available cars, in such a way that they can be returned the most preferred car for them from a car collection" with a 1% error.

If we relax the significance level to $\alpha = 0.05$ and additionally consider the set of two cars returned by the broker (where the number of successes raises to 10), the statistical test returns a *p-value* = 0.04957 ≤ α = 0.05. As a result, in this case we have a *strong* evidence to reject the H_0 and accept the H_1.

Subsequently, for an α = 0.05 we relaxed the initial null hypothesis to 35.75% or more of the users instead of half or more of the users. The returned p-value was 0.04944, meaning that we have a *strong* evidence to reject this hypothesis.

Concluding this statistical analysis, we have to highlight the following results:

- we can reject with a 1% error risk that at least half of the users expressed preferences without exploring available cars, and were returned the most preferred car for them from the car collection
- we can reject with a 5% error risk that at least half of the users expressed preferences without exploring available cars, and were returned a set of two cars that included the most preferred car for them for the car collection
- we can reject with a 5% error risk that at least 35.75% of the users expressed preferences without exploring available cars, and were returned the most preferred car for them from the car collection.

7 Discussing consequences and recommendations

Preferences are a multi-disciplinary topic. They have been studied in a number of fields, such as philosophy (Hansson, 2001), psychology (Scherer, 2005), decision making (Lichtenstein and Slovic, 2006), and economics (Fishburn, 1999). Moreover, they have been thoroughly studied in computer science fields such as in AI (Wellman and Doyle, 1991), HCI (Linden et al., 1997), and in information systems like in databases (Kießling, 2002; Chomicki, 2003), XML (Kießling et al., 2001), and OLAP (Golfarelli et al., 2011). A number of guidelines and examples for product search and recommender systems have been proposed in Pu and Chen (2008), while a survey of major questions and approaches for preference handling in applications such as recommender systems, personal assistant agents and personalised UIs is given at Peintner et al. (2008).

The last few years works like Kießling and Kostler (2005), Kießling et al. (2011) and Levandoski et al. (2010) have proposed the enrichment of declarative database query languages with preferences [a survey on representation, composition and application of preferences in DBs is given at Stefanidis et al. (2011)]. Our results show that it is

beneficial to offer the ability to express preferences during information exploration. Without providing the user with an overview of the available choices, the user not only will spend more time to express his preferences, but these preferences will probably not be effective. For instance, a widely adopted interaction model for exploratory search is that of *faceted and dynamic taxonomies* (FDT) (Sacco and Tzitzikas, 2009; Ruotsalo et al., 2013), which is used by various popular applications, e.g., booking.com, dblp, ebay, etc. FDT aids the user in getting acquainted with the information space and the available choices, and can be applied to a variety of sources, (e.g., see Papadakos et al., 2009, 2012). This model also supports attributes whose value is not a priori known but can be computed on demand, such as the result of mining tasks applied at query time [i.e., combination of a snippet-based results clustering algorithm (Kopidaki et al., 2009) with FDT].

Most FDT systems, like Flamenco (Hearst et al., 2002), output facets and zoom-points in lexicographical order, or based on the number of indexed documents (Oren et al., 2006). Some other systems like eBay, only present a manually chosen subset of facets to the users, and the zoom-points are again ranked based on the number of indexed documents. In addition in systems like eBay or Amazon, users are able to order the available objects according to simple object ordering operations over one specific attribute (e.g., order objects according to `Price`, or `Price + Shipping`, or `Duration` of auction in ascending or descending order). A number of automatic ways for ranking facets and zoom-points have been proposed in the literature. These approaches are based on *set-coverage* (Dakka et al., 2005), *interestingness* (Dash et al., 2008), *collaborative filtering* (Koren et al., 2008), *minimum-effort* (Roy et al., 2008; Kashyap et al., 2010), *log-based utility* (Pound et al., 2011), and *intuition-based* (Wagner et al., 2011) ranking methods.

None of the above approaches though provide an *incremental* preference specification mode, where the user can express his preferences in a controlled but flexible and unconstrained way. An enrichment of the FDT model with preferences expressed at interaction time has been proposed in Tzitzikas and Papadakos (2013), where the user can define the desired preference relation *gradually* as he/she explores the information base. The framework is appropriate for information spaces comprising resources described by attributes whose values can be hierarchically valued and/or multi-valued, supporting preference inheritance in the hierarchies, automatic conflict resolution, as well as preference composition.

The implementation prototype `Hippalus`, accessible through http://www.ics.forth.gr/Hippalus and described in Papadakos and Tzitzikas (2014), allows browsing and exploring RDF/S sources. Figure 3 shows the interface of `Hippalus` over an information base of 50 cars after the user has restricted his/her focus and expressed a number of preferences. In more details, the left sidebar displays the available and active attribute-values for this specific collection, with which the user can interact in order to restrict his/her focus. Next to each attribute-value the corresponding count numbers are also displayed. The middle part of the UI provides the ranked list of objects according to preference, while the right sidebar displays the history of information thinning and preference actions. Preference actions can be expressed over the attribute-values through HTML 5 context menus. Figure 4 shows how the user can express the relative preference that he/she prefers *Korean* to *European* cars. A video that demonstrates the full functionality of `Hippalus` is available online (http://www.youtube.com/watch?v=Cah-z7KmlXc).

Figure 3 Hippalus: restricted focus with preferences applied (see online version for colours)

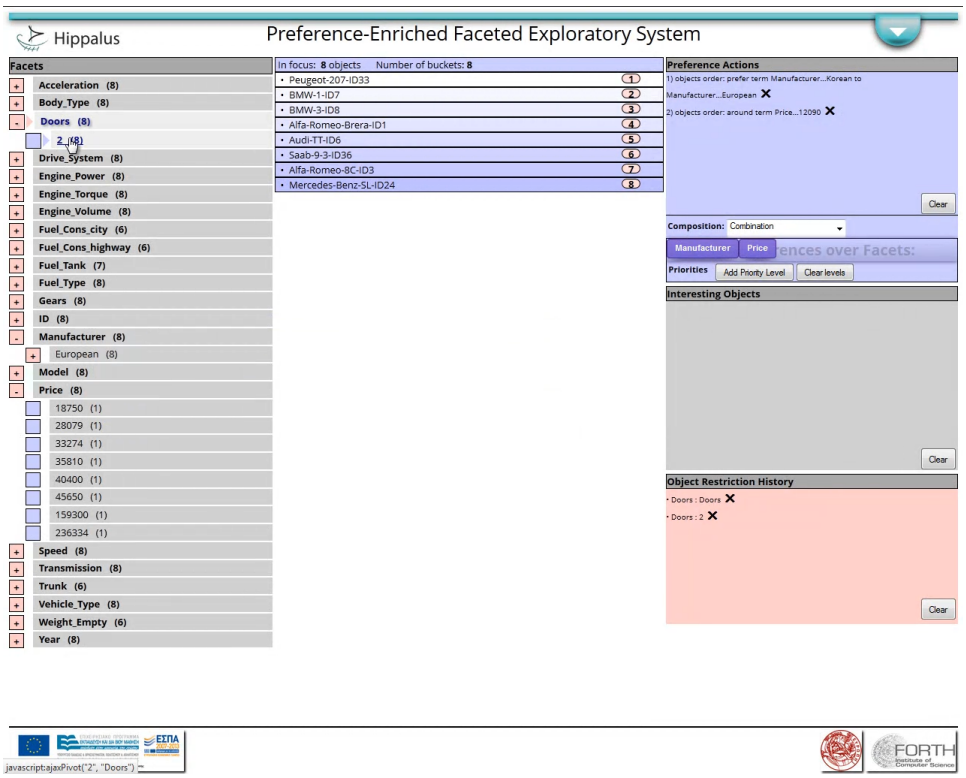
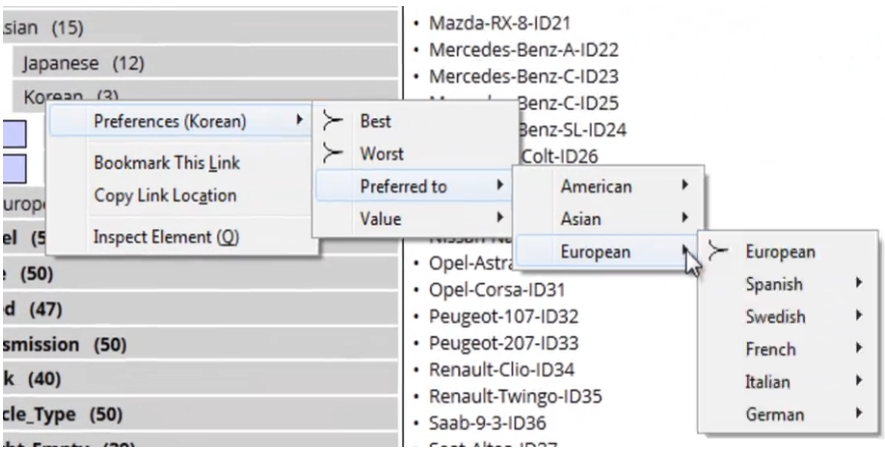


Figure 4 Hippalus: expression of the relative preference *Korean* $\succ$ *European* (see online version for colours)



According to the results reported in this paper, it seems that this direction is expected to be effective for the users. The reason is that such an interaction allows the flexible inspection of the available choices and the easy expression of intentional preferences over those choices as shown in Papadakos and Tzitzikas (2014).

8 Conclusions

In this work we formulated the hypothesis that without knowing the available choices, the declarative expression of preferences is a tiresome and time-consuming process, that does not lead to the desired results. We designed a user study, from an information systems perspective, using a model based on Pareto domination and prioritisation for aggregating preferences. Initially the users had to express their preferences for buying a new car and then they were asked to select the most preferred car from a list of cars. Although most of the users provided preferences that led to a clear winner, we found that only 20% of the users' expressed preferences leading to their selected car. On average users spent at least ten minutes to provide preferences that were either non-applicable or inconsistent to the selected car from a list of cars. The conducted statistical hypothesis test supports the results with a 1% error. As a consequence of the demonstrated results, we argue that it would be beneficial if systems let the users to explore the available choices and at the same time, i.e., during browsing/exploration of the information base, allow them to express their preferences as they are formulated according to the available choices. The current approach adopted by information systems, that restricts the user to give in one shot a complicated preference-based query, seems to be inefficient for supporting multi-criteria decision making over such systems.

In this direction, it is worthy to further research and elaborate on more complex and expressive structures than those currently supported by the preference framework proposed in Tzitzikas and Papadakos (2013). For instance, objects could be described with values accompanied by numbers expressing various quality aspects like accuracy, specificity, certainty, trust, authority, popularity, etc., which will require advancements of both the interaction and the preference framework. In such a way, users will be able to express preferences that take into consideration other qualitative aspects of the information space, and thus be returned a more refined ordering of the objects of interest.

Acknowledgements

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operation Program 'Education and LifeLong Learning' of the National Strategic Reference Framework (NSRF) – Research Funding Program: Herakleitus II. Investing in Knowledge Society through the European Social Fund. The authors would also like to thank the students of the 1st MUMIA Training School 'Building Next Generation Search Systems' for their participation in the user study.

References

Chomicki, J. (2003) 'Preference formulas in relational queries', *ACM Transactions on Database Systems*, Vol. 28, No. 4, pp.427–466.

Dakka, W., Ipeirotis, P. and Wood, K.R. (2005) 'Automatic construction of multifaceted browsing interfaces', in *Procs. of the 14th ACM CIKM '05*, pp.768–775, New York, NY, USA.

Dash, D., Rao, J., Megiddo, N., Ailamaki, A. and Lohman, G. (2008) 'Dynamic faceted search for discovery-driven analysis', in *Procs. of CIKM*.

Doyle, J. (2004) 'Prospects for preferences', *Computational Intelligence*, Vol. 20, No. 2, pp.111–136.

Figueira, J., Greco, S. and Ehrgott, M. (2005) *Multiple Criteria Decision Analysis. State of the Art Surveys*, Vol. 78, Springer Verlag.

Fishburn, P. (1999) 'Preference structures and their numerical representations', *Theoretical Computer Science*, Vol. 217, No. 2, pp.359–383.

Golfarelli, M., Rizzi, P. and Biondi, P. (2011) 'myOLAP: an approach to express and evaluate OLAP preferences', *IEEE Transactions Knowledge and Data Engineering*, Vol. 23, No. 7, pp.1050–1064.

Griffiths, D. (2009) *Head First Statistics*, Head First, O'Reilly, Sebastopol, CA.

Hansson, S.O. (2001) 'Preference logic', in Gabbay, D. and Guenthner, F. (Eds.): *Handbook of Philosophical Logic*, Chapter 4, Vol. 4, pp.319–393, Kluwer.

Hearst, M., Elliott, A., English, J., Sinha, R., Swearingen, K. and Yee, K-P. (2002) 'Finding the flow in web site search', *Commun. ACM*, Vol. 45, No. 9, pp.42–49.

Kashyap, A., Hristidis, V. and Petropoulos, M. (2010) 'FACeTOR: cost-driven exploration of faceted query results', in *Procs. of the 19th ACM International Conference on Information and Knowledge Management*, *CIKM '10*, pp.719–728, ACM, New York, NY, USA.

Keeney, R.L. and Raiffa, H. (1976) *Decisions with Multiple Objectives. Preferences and Value Tradeoffs*, John Wiley & Sons, New York.

Kießling, W. (2002) 'Foundations of preferences in database systems', in *VLDB '02*, *Procs. of the 28th International Conference on Very Large Data Bases*, pp.311–322, VLDB Endowment.

Kießling, W. and Kostler, G. (2005) 'Preference SQL – design, implementation, experiences', in *Procs. of the 28th International Conference on Very Large Data Bases* (*VLDB*), pp.990–1001, Hong Kong, China.

Kießling, W., Endres, M. and Wenzel, F. (2011) 'The preference SQL system – an overview', *IEEE Data Engineering Bulletin*, Vol. 34, No. 2, pp.11–18.

Kießling, W., Hafenrichter, B., Fischer, S. and Holland, S. (2001) 'Preference XPATH: a query language for e-commerce', *Information Age Economy*, pp.427–440, Springer Publishing Company, Incorporated.

Kopidaki, S., Papadakos, P. and Tzitzikas, Y. (2009) 'STC+ and HSTC: two novel online results clustering methods for web searching', in *WISE '09*, *Procs. of the 10th Intern. Conf. on Web Information Systems Engineering*.

Koren, J., Zhang, Y. and Liu, X. (2008) 'Personalized interactive faceted search', in *WWW '08. Procs. of the 17th International Conference on World Wide Web*, pp.477–486, ACM, New York, NY, USA.

Levandoski, J., Mokbel, M. and Khalefa, M. (2010) 'FlexPref: a framework for extensible preference evaluation in database systems', in *ICDE*, pp.828–839.

Lichtenstein, S. and Slovic, P. (2006) *The Construction of Preference*, 13th ed., Cambridge University Press, New York.

Linden, G., Hanks, S. and Lesh, N. (1997) 'Interactive assessment of user preference models: the automated travel assistant', in Jameson, C.P.A. and Tasso, C. (Eds.): *User Modeling. Procs. of the Sixth International Conference*, pp.67–78, Springer, Wien.

Oren, E., Delbru, R. and Decker, S. (2006) 'Extending faceted navigation for RDF data', in *Procs. of ISWC '06*, pp.559–572, Athens, GA, USA.

Papadakos, P. and Tzitzikas, Y. (2014) 'Hippalus: preference-enriched faceted exploration', in *Proceedings of the Workshops of the EDBT/ICDT 2014 Joint Conference* (*EDBT/ICDT 2014*), Athens, Greece, March 28, pp.167–172.

Papadakos, P., Armenatzoglou, N., Kopidaki, S. and Tzitzikas, Y. (2012) 'On exploiting static and dynamically mined metadata for exploratory web searching', *Knowledge and Information Systems*, Vol. 30, No. 3, pp.493–525.

Papadakos, P., Kopidaki, S., Armenatzoglou, N. and Tzitzikas, Y. (2009) 'exploratory web searching with dynamic taxonomies and results clustering', in *Procs. of the 13th European Conference on Digital Libraries* (*ECDL '09*).

Peintner, B., Viappiani, P. and Yorke-Smith, N. (2008) 'Preferences in interactive systems: technical challenges and case studies', *AI Magazine*, Vol. 29, No. 4, pp.13–24.

Pound, J., Paparizos, S. and Tsaparas, P. (2011) 'Facet discovery for structured web search: a query-log mining approach', in *Procs. of the International Conference on Management of Data*, *SIGMOD '11*, pp.169–180, ACM, New York, NY, USA.

Pu, P. and Chen, L. (2008) 'User-involved preference elicitation for product search and recommender systems', *AI Magazine*, Vol. 29, No. 4, pp.93–103.

Roy, S. B., Wang, H., Das, G., Nambiar, U. and Mohania, M. (2008) 'Minimum-effort driven dynamic faceted search in structured databases', in *CIKM '08*, *Procs. of the 17th ACM Conference on Information and Knowledge Management*, pp.13–22, New York, NY, USA.

Ruotsalo, T., Athukorala, K., Glowacka, D., Konyushkova, K., Oulasvirta, A., Kaipiainen, S., Kaski, S. and Jacucci, G. (2013). Supporting exploratory search tasks with interactive user modelling. In *Proceedings of ASIST*, *the 76th ASIS&T Annual Meeting*.

Sacco, G.M. and Tzitzikas, Y. (Eds.) (2009) *Dynamic Taxonomies and Faceted Search*. *Theory*, *Practise and Experience*, Springer Publishing Company, Incorporated.

Sakthivel, G., Ilangkumaran, M., Nagarajan, G., Raja, A., Ragunadhan, P. M. and Prakash, J. (2013) 'A hybrid MCDM approach for evaluating an automobile purchase model', *IJIDS*, Vol. 5, No. 1, pp.50–85.

Scherer, K.R. (2005) 'What are emotions? And how can they be measured?', *Social Science Information*, Vol. 44, No. 4, pp.695–729.

Stefanidis, K., Koutrika, G. and Pitoura, E. (2011) 'A survey on representation, composition and application of preferences in database systems', *ACM Transactions on Database Systems*, Vol. 36, No. 3, pp.19:1–19:45.

Tzitzikas, Y. and Papadakos, P. (2013) 'Interactive exploration of multi-dimensional and hierarchical information spaces with real-time preference elicitation', *Fundamenta Informaticae*, Vol. 122, No. 4, pp.357–399.

Utkin, L.V. (2014) 'Reducing the Pareto optimal set in MCDM using imprecise probabilities', *IJOR*, Vol. 19, No. 1, pp.40–59.

Wagner, A., Ladwig, G. and Tran, T. (2011) 'Browsing-oriented semantic faceted search', in *DEXA* (*1*), pp.303–319.

Wasserman, L. (2004) *All of Statistics*. *A Concise Course in Statistical Inference*, Springer Publishing Company, Incorporated.

Wellman, M. and Doyle, J. (1991) 'Preferential semantics for goals', in *AAAI*, pp.698–703.

Appendix

Evaluation questionnaires

The questionnaires used for the evaluation of the IPD hypothesis are shown in Figure A1, Figure A2 and finally Figure A3.

Figure A1 Evaluation Step A: users express their preferences for a car (see online version for colours)

Preferences Evaluation – Step 1

ID:_____
Age: _____
Gender: _____
Education: _____
Country: _____
Time to Complete: _____

Suppose that you have (it is obligatory) to change your car. You have to select and buy a new one, which you will use for the next 5 years, and of course **you will have to pay it**. Please express your preferences on paper. This paper will be handed to a different person who has at his disposal a limited collection of available cars. This person will select one for you based on the available cars and the preferences that you expressed.

You have 30 minutes at most to express your preferences. You are free to express them in any form you like, e.g. in natural language text (e.g. I prefer a car with an engine volume between 1200 and 1400 cc), by providing an ordering of the firms according to your preference (e.g. Japanese, European or BMW, Audi), by specifying the preferred (ideal) price, etc. Other characteristics could include year, body type, engine volume, power, max speed, acceleration, fuel consumption, weight, fuel type, price, trunk, etc.

Please measure how much time you spent on this exercise and give us the paper.

Figure A2 Evaluation Step B: users select a car from the list (1st page) (see online version for colours)

Preferences Evaluation - Step 2

User ID:		Car ID:	Time:

Suppose that you have (it is obligatory) to change your car and buy a new one. You are obliged to use the new car for the next 5 years, and of course you will have to pay it. Please select the ideal car for you from the list below.

Id	Brand	Model	Modification	Year	Body type	Number of doors	Engine volume cc.	Power HP	Torque Nm	Max. speed km/h	Accel. (0-100km/h) sec.	Fuel Cons. (city) l /100 km	Fuel Cons. (highway) l / 100 km	Netto weight kg.	Weight kg	Trunk l.	Fuel tank l.	Fuel type	Price EUR
1	**Alfa Romeo**	**Brera**	Brera 2.2 JTS	2005	coupe	2	2198	185	230/4500	222	8.6	13.1	7.3			236/546	70	gasoline	35810
2	**Alfa Romeo**	**MiTo**	Mi.to 1.4 MPI	2008	hatchback	3	1368	78	120/4750	165	12.3	7.7	4.8	1080			45	gasoline	15300
3	**Alfa Romeo**	**8C**	8C Spader 4.7	2009	cabriolet	2	4691	450	480/4750	290	4.2							gasoline	159300
4	**Audi**	**A3**	A3 1.4 TFSI	2008	hatchback	3	1390	125	200/1500	203	9.4	7.8	4.8	1245	1820	350	55	gasoline	22500
5	**Audi**	**S8**	S8 5.2 FSI quattro	2007	sedan	4	5204	450	540/3500	250	5.1	19.5	9.5	1940	2540	500	90	gasoline	97600
6	**Audi**	**TT**	Roadster 2.0 TFSI quattro	2010	roadster	2	1984	211/5100	350/1800	240	5.8	10.2	5.7	1405	-	260	55	gasoline	40400
7	**BMW**	**1**	118i 16V Steptronic	2008	cabriolet	2	1895	143/6000	200/3750	208	10.1	8.7	5.4	1500	1870	280	53	gasoline	28079
8	**BMW**	**3**	320i 2.0i 16V 6AT	2010	coupe	2	1995	156/6200	200/3600	218	9.1	12.1	6.2	1360	-	440	63	gasoline	33274
9	**BMW**	**7**	760i	2008	sedan	4	5972	544/5250	750/1500	250	4.6	-	-	2105	2695	500	82	gasoline	137425
10	**Citroen**	**C1**	C1 1.0	2008	hatchback	3	998	68	93/3600	157	13.7	5.5	3.9	790	1160	139/712	35	gasoline	9450
11	**Citroen**	**C3**	1.1i	2010	hatchback	5	1124	61/5500	94/3200	155	16.50	8.10	4.90	1010	-	300	45	gasoline	10400
12	**Fiat**	**Bravo**	Bravo 1.4	2007	hatchback	5	1368	90	128/4500	179	12.5	8.7	5.6	1205	1705	400/1175	58	gasoline	15470
13	**Fiat**	**Punto**	Grande Punto 1.4	2005	hatchback	3	1368	77	115/3000	165	13.2	7.7	5.2	1025		275/1030	45	gasoline	12090
14	**Ford**	**Fiesta**	Fiesta 1.25	2008	hatchback	3	1242	82	114/4200	168	13.3	7.5	4.6	1098	1490		45	gasoline	11683
15	**Ford**	**Ka**	Ka 1.3	2009	hatchback	3	1299	60	99/2500	155	15.4	8.1	4.6	962			35	gasoline	8733
16	**Hyundai**	**i10**	1.1 CRDi	2008	hatchback	5	1120	75	152/1900	167	-	-	-	-	-	-	53	diesel	9620
17	**Hyundai**	**i30**	1.5 CRDi	2008	minivan	5	1493	110	235/1900	164	14.3	6.8	4.6	1240	1840	350	55	gasoline	14000
18	**Kia**	**Ceed**	1.6 4AT	2007	touring	5	1591	122	154/5200	187	11.4	8.9	5.8	1263	-	-	53	gasoline	17900
19	**Lancia**	**Delta**	Lancia 1.4	2008	hatchback	5	1400	120	206/1750							380/1190		gasoline	18810
20	**Mazda**	**3**	1,6 Touring	2009	hatchback	5		105	145	181	12/1	8.5	5.3		1770	340	55	gasoline	18435
21	**Mazda**	**RX-8**	1,3 Energy	2008	sedan	4	1308	192	220	223	7.2	14.1	8.1	1373	1815	290	61	gasoline	29995
22	**Mercedes-Benz**	**A**	A 150	2008	hatchback	5	1498	95	140/3500	175	12.6	7.4	5.1	1245	1740	435/1370	54	gasoline	20827
23	**Mercedes-Benz**	**C**	200 CDI 5AT	2010	sedan	4	2148	136/3800	340/1800	213	10.2	9.1	5.1	1575	2060	475	66	diesel	36265
24	**Mercedes-Benz**	**SL**	SL 65 AMG V12	2008	cabriolet	2	6208	612/4750	1000/2000	250	4.2	15.1	23.4	2120	2385	235	80	gasoline	236334

Figure A3 Evaluation Step B: users select a car from the list (2nd page) (see online version for colours)

Preferences Evaluation - Step 2

Id	Brand	Model	Modification	Year	Body type	Number of doors	Engine volume cc.	Power HP	Torque Nm	Max. speed km/h	Accel. (0-100km/h) sec.	Fuel Cons. (city) l /100 km	Fuel Cons. (highway) l / 100 km	Netto weight kg.	Weight kg	Trunk l.	Fuel tank l.	Fuel type	Price EUR
25	**Mercedes-Benz**	**C**	180 Kompressor 5AT	2011	sedan	4	1796	156/5200	230/2600	220	9.9	10.7	5.8	1500	1985	475	66	gasoline	32066
26	**Mitsubishi**	**Colt**	1,3 AT Invite	2008	hatchback	3	1332	95	125 / 4000	180	11.8	7.3	4.6	945	1435		47	gasoline	12175
27	**Mitsubishi**	**X-Trail**	2,5 SE (ABA-B)	2008	crossover	5	2488	169	233	185	10.3	12.1	7.1	1623	2050	479	65	gasoline	30630
28	**Nissan**	**Micra**	1,2 Luxury	2009	hatchback	5	1240	80	110	145	17.9	8.1	5.1	1068	1490	251	46	gasoline	10720
29	**Nissan**	**Navara**	2,5 SE (8CCH-)	2008	pickup	4	2488	174	403	170	11.5	11.1	7.1	2128	2805		80	diesel	22190
30	**Opel**	**Astra**	1.3 TD	2010	hatchback	5	1248	95/4000	190/1750	170	14.7	5.1	3.6	1373	1870	370	56	diesel	14700
31	**Opel**	**Corsa**	1.4 16V 4AT 3/5	2010	hatchback	3	1364	90/5600	125/4000	166	14.8	8.7	5.6	1147	1560	285	45	gasoline	12982
32	**Peugeot**	**107**	107 1.0	2009	hatchback	5	998	68	93/3600	157	13.7	5.5	3.9	875	1190	131/751		gasoline	9950
33	**Peugeot**	**207**	1.6HDi	2009	cabriolet	2	1560	110/4000	240/1750	193	10.90	6.20	4.30	1423	1785	145	50	diesel	18750
34	**Renault**	**Clio**	Clio III 1.2	2009	hatchback	3	1149	75	105	167	13.4	7.6	4.9			288/1028	55	gasoline	11636
35	**Renault**	**Twingo**	Twingo II 1.5 dCi	2008	hatchback	3	1461	85		164		4.8	3.5			230/959	40	diesel	16591
36	**Saab**	**9/3**	9-3 2.8	2008	cabriolet	2	2792	280	370/1800	250	7.2			1580	1680		62	gasoline	45650
37	**Seat**	**Altea**	Altea 1.4	2009	minivan	5	1390	85	132/3600	169	14.8	9.7	5.8				55	gasoline	16282
38	**Seat**	**Leon**	Leon 1.4	2009	hatchback	5	1390	85	130/3800	172	13.7	9.5	5.6			341	55	gasoline	15538
39	**Skoda**	**Octavia**	OctaviaII 2.0 RS TDI	2009	hatchback	5	1984	170								560/1420	55	diesel	24466
40	**Skoda**	**Yeti**	Yeti 1.2 TSI	2009	crossover	5	1197	105	175/3500							1760		gasoline	18090
41	**Suzuki**	**Grand Vitara 3D New**	1,6 JLX-E	2009	crossover	3	1586	107	145	160	14.4	10	7	1445	1800	210	55	gasoline	18700
42	**Suzuki**	**Jimny**	1,3 JLX	2008	touring	3	1328	93	110	140		9	6	1045	1420	286	40	gasoline	16260
43	**Suzuki**	**Swift**	1,3 GLX MTA	2008	hatchback	5	1328	93	116	175		8	5	1060	1485	213	45	gasoline	11560
44	**Toyota**	**Prius**	1.8	2009	hatchback		1798	99	142	180	10.4	3.9	3.7	1495	1805	445	45	gasoline	29550
45	**Toyota**	**RAV4**	2,0 Komfort	2009	crossover		1998	152	194	175	12.0	11.6	7.4	1610	2110	410	60	gasoline	24120
46	**Toyota**	**Yaris**	1,33 Prestige	2009	hatchback		1329	101	132	175	13.4	6.2	4.4	1115	1480	272	42	gasoline	12990
47	**Volkswagen**	**Golf**	Golf VI GTI 2.0	2009	hatchback	3	1984	210	280/5200	238	6.9	10.1	5.8	1318	1817	350/1305	55	gasoline	29700
48	**Volkswagen**	**Scirocco**	Scirocco 1.4 TSI	2008	hatchback	3	1390	122	200/4000	200	9.7	8.4	5.3	1244	1690	312/1006	50	gasoline	20970
49	**Volkswagen**	**Tiguan**	Tiguan 1.4	2008	crossover	5	1390	150	240/1750	195	9.3	9.4	6.2	1451	2080	470/1510	64	gasoline	26750
50	**Volvo**	**C30**	C30 1.6	2006	hatchback	3	1596	100	150/4000	185	11.8	9.3	5.7	1210	1700	364/1010	55	petrol (gasoline)	26500