



University of
Waterloo

Technical Report
CS-2012-03

**A Study on Justifications for Choices:
Explanation Patterns and Guidelines**

**Ingrid Oliveira de Nunes
Simon Miles
Michael Luck
Carlos José Pereira de Lucena**

Faculty of Mathematics

**DAVID R. CHERITON SCHOOL OF COMPUTER SCIENCE
UNIVERSITY OF WATERLOO
WATERLOO, ONTARIO, CANADA N2L 3G1**

A Study on Justifications for Choices: Explanation Patterns and Guidelines ¹

Ingrid Oliveira de Nunes^{1,2}, Simon Miles², Michael Luck²,
Carlos José Pereira de Lucena¹

¹ Pontifical Catholic University of Rio de Janeiro - Rio de Janeiro - Brazil

² King's College London, Strand, London, WC2R 2LS, United Kingdom

{ionunes,lucena}@inf.puc-rio.br
{simon.miles,michael.luck}@kcl.ac.uk

Abstract. Many different forms of explanation have been proposed for justifying decisions made by automated systems. However, there is no consensus on what constitutes a *good* explanation, or what information these explanations should include. In this paper, we present the results of a study into how people justify their decisions. Analysis of our results allowed us to extract the forms of explanation adopted by users to justify choices, and the situations in which these forms are used. The analysis led to the development of guidelines and patterns for explanations to be generated by automated decision systems. This paper presents the study, its results, and the guidelines and patterns we derived.

Keywords: User Explanation, Guidelines, Patterns, Recommender Systems, Decision Support Systems.

¹Trabalho patrocinado pelo Ministério de Ciência e Tecnologia da Presidência da República Federativa do Brasil e FINEP.

In charge of publications:

Helen Jardine

David R. Cheriton School of Computer Science

University of Waterloo

Waterloo, Ontario, Canada N2L 3G1

Tel: +1 519 888-4567 x33293 Fax: +1 519 885-1208

E-mail: hjardine@uwaterloo.ca

Web site: <http://www.cs.uwaterloo.ca/research/tr/>

Contents

1	Introduction	1
2	Experiment Description	1
2.1	Research Questions	2
2.2	Procedure	3
2.3	Participants	5
3	Results and Analysis	6
4	Interpretation	13
4.1	Explanation for Acceptance	13
4.2	Explanation for Rejection	15
5	Guidelines and Patterns	16
5.1	Guidelines	16
5.2	Patterns	17
5.2.1	Pattern 1: Critical Attribute	18
5.2.2	Pattern 2: Cut-off	19
5.2.3	Pattern 3: Domination	19
5.2.4	Pattern 4: Minimum Requirements ⁻	20
5.2.5	Pattern 5: Minimum Requirements ⁺	21
5.2.6	Pattern 6: Decisive Criteria	21
5.2.7	Pattern 7: Trade-off Resolution	22
6	Related Work	23
7	Final Considerations	23
	References	24
A	Questionnaire	25
A.1	Introduction: Survey on Reasons for Choice	25
A.2	Part I: User Data	25
A.3	Part II: Option Selection	25
A.4	Part III: Reasons for your Choice	25

1 Introduction

The popularity of recommender systems has increased significantly in the last decade, with many commercial applications already adopting them. For many years, the main goal of research into such systems has been to improve their *accuracy*, associating this measure with the quality of the recommendation. However, as argued by McNee et. al. (McNee, Riedl & Konstan 2006), the most accurate systems (based on standard metrics) may not be those that provide the most useful recommendations to users. Other aspects, such as trust and transparency, have also been considered, and many of these can be improved by providing users with *explanations* (Tintarev & Masthoff 2007). Such explanations justify the choice of a particular recommendation to users, and their applicability extends to decision support systems (Labreuche 2011) and over-constrained problem solvers (Junker 2004).

There are different existing approaches to generating explanations, from exposing the rationale of the underlying recommendation technique to selecting the essential attributes on which the decision is based. However, there is no consensus on what constitutes a *good* explanation, and what kinds of information must be presented to users in such explanations. Even though existing work (Tintarev & Masthoff 2007) provides qualitative arguments that characterise good explanations, there is no extensive research into the kinds of explanation that users expect and need to understand and accept recommendations or decisions made on their behalf and, where work does exist, it is particular to a specific system.

In response, this paper presents a study whose main objective is to give guidance for explanation generation. The study performed consisted of a survey, from whose results we extract types of explanation that people use to justify a choice from a set of available options. As, based on the design of the study, we can assume that the explanations provided by study participants are those that the users would expect to receive, we derive a set of *guidelines* and *patterns*, which are a basis for generating explanations for users as to why particular options are chosen by a recommender system or decision support systems. Therefore, this paper presents three contributions: (i) the design and results of a study into what explanations users expect when justifying choices made; (ii) guidelines for the qualities and forms of explanation needed to best meet user expectations; and (iii) patterns for explanations to be given under different circumstances.

The remainder of this paper is organised as follows. We first detail the performed study in Section 2, then present and discuss its results in Section 3 and our interpretation in Section 4. We propose guidelines and patterns derived from this study in Section 5. Related work is detailed in Section 6 and, finally, we conclude in Section 7.

2 Experiment Description

The framework that provided guidance for this study elaboration was proposed by Basili et al. (Basili, Selby & Hutchens 1986), which includes the goal-question-metric (GQM) template (Basili & Rombach 1988), used to define the goal of the study, and later to define research questions and select metrics for answering those questions. The goal of the present study, following the GQM template, is presented in Table 1. We highlight that the work of Basili et al. focus on experimentation in the context of Software Engineering (SE), but it is sufficiently generic to be applied to other areas, and the reason for using their guidelines

and template is due to the experience of the researchers with SE studies.

Definition element	Our experiment goal
Motivation	To identify the kinds of explanations users expect to receive from software systems,
Purpose	characterise and evaluate
Object	explanations to justify a choice
Perspective	from a perspective of the researcher
Domain:people	as they are provided by people
Scope	of the context of the social network of the researcher.

Table 1: Goal Definition (GQM template).

In order to achieve this goal, the study we designed is based on a questionnaire made available online, in which participants (individuals that are part of the social network of the researchers) have to make a choice from a set of available options and later justify their decision. The arguments given by participants are carefully analysed to understand their common characteristics and also the dependency between options and the arguments given. The options given in our study consist of hotels located in New York city, USA, and this decision was made due to three reasons.

- (i) *New York is a widely known touristic city*, therefore participants are more likely to be aware of close to where they would like to stay, prices they are accepting to pay, and so on.
- (ii) *Researchers knowledge about the city*, thus we are able to select appropriate options for being part of the study.
- (iii) *Massive amount of available hotels*, which is important as our study is based on *real* hotel data so participants take it more seriously.

Next sections provide further details about our study. We start by presenting our research questions of the study in Section 2.1, and then detail the study procedure in Section 2.2. Then, we describe the participants of our study in Section 2.3, to later proceed to the analysis and interpretation of our obtained results.

2.1 Research Questions

As shown in the previous section, our main objective while performing this study is to give guidance for explanation generation. It is accepted by the research community that explanations improve software systems that provide decision-making by turning its reasoning process more explicit, thus enhancing user acceptance and trust. Nevertheless, there is no consensus of what constitutes a *good* explanation, and what kind of information must be provided for users. Therefore, the present study aims at solving this issue in a user-centric way by identifying the kinds explanation that people give — and we assume that these are those that users expect to receive — to justify a decision, and then providing guidelines and patterns that allow defining good explanations from a user perspective. In our study we addressed four different research questions, which are presented in Table 2(a).

(a) Research Questions.	(b) Evaluation Approaches.
RQ1. Do users adopt a pattern to justify an option chosen from the set of those available?	EA1. Analysis of the arguments given to justify the chosen option and identification of commonalities among arguments given by different users.
RQ2. Is there a relationship between the type of explanation given to support the decision and the chosen option?	EA2. Comparison among the arguments given to justify each different chosen option.
RQ3. Do users use a pattern to justify the rejected (not chosen) options?	EA3. Analysis of the arguments given to reject options and identification of commonalities among arguments given by different users.
RQ4. Is there a relationship between the type of explanation given to reject options and the rejected or chosen option?	EA4. Comparison among arguments given to reject options according to each different chosen and rejected option.

Table 2: Research questions and their evaluation approach.

By answering these research questions, we are able to extract patterns for user explanations to be generated by decision-making systems (based on *RQ1* and *RQ3*), and also the context in which each pattern is adopted (based on *RQ2* and *RQ4*). These explanations are associated with both chosen and rejected options — the first two questions focus on patterns and their context for explaining the chosen option; and the last two address explaining why other options are rejected (or not chosen).

2.2 Procedure

In a nutshell, our study consists of collecting information provided by participants through a web-based questionnaire, and analysing the collected data later. Our aim was to obtain a high number of participants, and therefore anyone with Internet access could access the questionnaire (more details about the set of participants are given in next section). Our study involves decision-making and explanation about this process, and we choose hotels as the domain associated with the decision. The main reason for this design choice is that most of people are aware of the attributes that characterise hotels, and have preferences for individual attributes. Moreover, we chose to provide hotels in New York city for the reasons already presented. The applied questionnaire, which can be seen in Appendix A, consists of three parts, and each of which is explained next.

User Information Data. Our study does not assume that explanations depend on people characteristics, such as age or gender; however we collect some information about the participants in case we need these, and also to be able to provide demographic information (as we make the questionnaire available online any individual can access it). The collected participant data is: (i) age; (ii) gender; (iii) location (city and country); and (iv) working/studying field.

Choosing Product. The study participant is then requested to imagine the scenario in which she is going to spend holidays in New York, and must choose a hotel for staying there from a set of options that we make available for her. As hotel rates for double rooms are very similar to those for single rooms, and people usually spend holidays with at least with one friend, we include in this hypothetical scenario that

the participant would travel with a friend and do not mind to share a bed with him or her, in order to make our scenario more realistic. For the same reason, we have selected real existing hotels to offer for participants. We have selected *five* different hotels — *Hotel 91*, *Econo Lodge Times Square*, *The Hotel at Times Square*, *Comfort Inn Times Square*, *Renaissance New York Hotel 57* — whose complete details can be seen in Appendix A.

G-1 Dominated option. Even though a dominated option (i.e. an option that has no advantage and at least one disadvantage with respect to another) is in general not chosen, we have added an option with this characteristic to the set of those available in order to capture the arguments that participants use to reject them. In our case, “The Hotel at Times Square” dominates “Comfort Inn Times Square.” In theory, the former does not dominate the latter, as two of the attributes of the “Comfort Inn Times Square” option are better than the “The Hotel at Times Square;” however our assumption (later confirmed by our study) is that one of these attributes is usually not considered (parking price) and the other (room size) has a very small difference so, in practice, it is a case of domination from the perspective of most of the participants.

G-2 Extreme options. Extreme options compromise too much one attribute (e.g. quality) to improve another (e.g. price). People in general avoid such options, as stated by the extremeness aversion principle (Simonson & Tversky 1992), so we also select extreme options to understand how participants explain their rejection or, if they choose them, why. There are two extreme options: (i) much lower quality and much lower price (*Hotel 91*); and (ii) much higher quality and much higher price (*Renaissance New York Hotel 57*).

G-3 Options that Require Trade-off Resolution. Two options that have relative pros and cons require a trade-off to be made. As this may require a different form of explanation from either category above, we include options which clearly illustrate such a need for trade-off, *Econo Lodge Times Square* and *The Hotel at Times Square*.

Reasons for Choice. The last step of our questionnaire consists of providing reasons for choosing a particular hotel. The participant is asked to say why she chose a particular option, and why she rejected the remaining options — we are assuming that if a participant does not chose an option, she is automatically rejecting it. In order to obtain good answers, we highlight for the participant that complete answers should be provided and the arguments should be good enough to convince their friend to accept the selected hotel.

The most important information collected in the study are the provided justifications expressed in natural language. So, the analysis part of the study consists first of making a careful investigation on these justifications to identify patterns and define explanation types — there might be different types for acceptance or rejection of options. Moreover, additional characteristics might be identified, such as using arguments that involve the participant preferences. Based on this initial analysis, we can extract quantitative data from the study.

Qualitative	Quantitative
<ul style="list-style-type: none"> • Justifications for acceptance • Justifications for rejection • Explanation types • Additional characteristics of justifications 	<ul style="list-style-type: none"> • Chosen hotel • Chosen hotel vs. Explanation types for acceptance • Chosen hotel vs. Explanation types for rejection

Table 3: Data collected in our study.

Gender	Male 58 (58%)		Female 42 (42%)	
Country	Brazil 78 (78%)	United Kingdom 8 (8%)	Canada 5 (5%)	Other 9 (9%)
Age	16-25 years 4 (4%)	26-35 years 61 (61%)	36-45 years 11 (11%)	>45 years 24 (24%)

Table 4: Demographic Characteristics of Participants.

In Table 2(b), we show our evaluation approach for answering each research question based on our study, which is mainly based on a classification of explanation types. We summarise all data collected in our study, which includes both “raw” data (i.e. data directly provided by participants) and derived data (such as the explanation types), in Table 3. The presented data consists of the items already discussed, and also their association, e.g. the comparison of the relationship between the chosen hotel and the explanation type adopted to justify the decision.

2.3 Participants

The participants of our survey were selected using convenience sampling, which reached a total number of 100 participants. The sample was obtained based on the social network of the researchers involved in this study, by means of two forms of publishing the survey: (i) by e-mail, using the contact list of the researcher; and (ii) by Facebook,² which is a widely known social network. The distributed message consists of an invitation to participate of the survey and a request to forward the invitation for other people.

The survey was available for participation in October 12–24, 2011 and was initiated by 191 people, who took at least one of the steps of the survey, from which 100 (52.36%) finished all the survey steps — the remaining surveys were discarded. The demographic characteristics of the participants that completed the survey are described in Table 4. Because we adopted the social network of the lead researcher to perform the study, most of the participants are aged between 26 and 35 years (61%) and are Brazilians (78%). Non-Brazilian participants are from 8 other countries: United Kingdom, Canada, Germany, United States of America, Switzerland, China, France and Netherlands, and the last six were grouped into the “Other” category in Table 4, as there are only a few participants from these countries.

²<http://www.facebook.com>

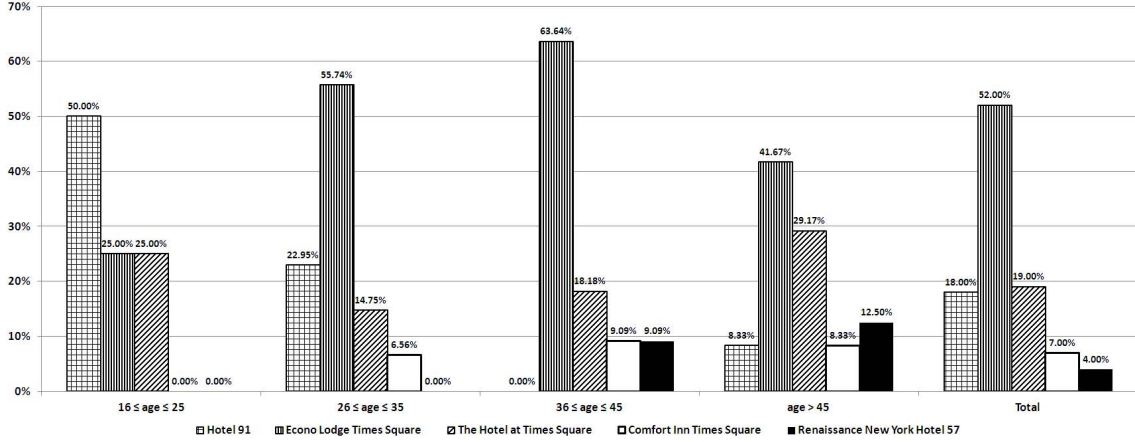


Figure 1: Hotel Choice.

3 Results and Analysis

Our collected data consists mainly of justifications expressed in natural language and, as these are qualitative data, we analyse them in a systematic way to extract quantitative information. In the section, we explain how we performed this analysis and show results obtained from our study, separating our findings according to the research questions we aim to answer. We focus only on describing the obtained data and our qualitative analysis, and we let for next section further discussions and our interpretation. Note that, at various points, we label some findings with “*Evidence X*,” in order that we can later refer to them to support our proposed guidelines.

Before proceeding to this detailing, we present the hotels chosen by our participants. This information is relevant for understanding the relationship between the chosen option and justifications, as indicated by our research questions *RQ2* and *RQ4*. Figure 1 shows how many participants selected each hotel and, as expected, the majority of participants chose a hotel from the group G-3. We also show the choice distribution according to age, and it can be observed that younger participants prefer cheaper options.

RQ1: Do users adopt a pattern to justify an option chosen from the set of those available? Each participant has to provide five justifications for their choice, from which one is a justification for why they choose a particular hotel. We have analysed all provided justifications and derived from them a classification, which we refer to as *explanation types*, consisting of *six* different types that are described as follows. This classification emerged from the qualitative analysis of collected data, which is supported by principles of grounded theory (Glaser 1992). We exemplify each of these explanation types for the acceptance scenario in Table 5.

Critical attribute. For a group of participants, there is an attribute that plays a crucial role in the decision-making process, being in most of the cases the attribute *price*. In these situations, the justification focuses only on this crucial attribute, and the remaining ones are omitted.

Dominance. As already introduced before, it is said than an option dominates another

when the first is better than the second for at least one attribute, and not worse for the others, and the domination relationship can be used as an argument to justify a decision. The acceptance of an option is justified using the dominance only when it dominates all other options, which is an uncommon situation mainly when choosing among products, because due to seller competition there is always a trade-off to be resolved, with options presenting both pros and cons; if it *does* happen, the decision is extremely easy to be made. However, one option may dominate another from a particular participant perspective, as she might not care about a set of attributes, and the remaining ones create this ideal scenario to make the decision.

Main reason. Some participants take into account many attributes to make a decision, but a particular option may be chosen (or rejected) when there is one attribute value that, together with its importance, is decisive for the choice.

Minimum requirements. People usually have a set of hard constraints that are used to filter available options by discarding those that do not satisfy all of them — this can be seen as the establishment of cut-off values. If only one option satisfies all requirements, the decision becomes easy as the justification for the option acceptance is that it satisfies all requirements. Furthermore, some participants provide a justification based on minimum requirements; however, as more than one option satisfy them, the participants also provide a criterion to decide among them, such as minimum price.

One-sided Reasons. Instead of only providing the main reason for the acceptance, many participants focus on exposing only positive aspects (or negative, in case of rejection) of the option, even though the chosen option has disadvantages (or advantages) with respect to other options and considering the participant preferences. This is an indication the existence of a minimal set of attributes that made the option to be chosen (or rejected).

Pros and Cons. The most complex type of explanation consists of making the option pros and cons explicit, and showing the reasoning process behind the choice. Based on these pros and cons evaluation, the participant states that the pros compensate cons — or do not compensate them, in case of rejection. In some cases, participants do not enumerate pros and cons, but only state “*this is (not) the best cost-benefit relationship.*”

This set of explanation types indicates that justifications for choosing an option provided by participants, and more generally by people, do follow patterns, and these can be used in software systems for explanation generation. The right hand side of Figure 2 (which shows the explanation types used to justify each hotel) shows the total number of the different explanation types adopted by the participants, who mostly adopt *one-sided reasons* and *pros and cons* to explain their choices; however, this part of the chart does not give any indication why a particular explanation type is adopted, and this issue is investigated in the next research question.

RQ2: Is there a relationship between the type of explanation given to support the decision and the chosen option? Given that we have identified patterns used to justify why a particular hotel is chosen, we next investigate if there is any relationship

Explanation Type	Example of Justification for Acceptance
Critical attribute	H_i is the cheapest option.
Dominance	H_i is better in all aspects.
Main reason	I chose H_i because it offers the benefit a_i .
Minimum requirements	From the hotels that satisfy my requirements, H_i is the cheapest.
One-sided Reasons	I chose H_i because it provides the benefits a_i and a_j .
Pros and Cons	Even though H_i is not the cheapest, it provides the benefits a_i and a_j .

Table 5: Example of Justification for Acceptance.

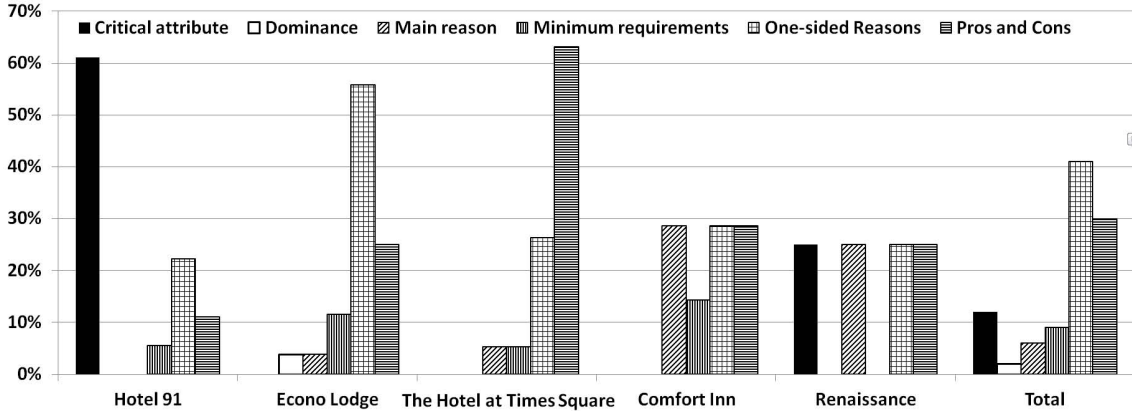


Figure 2: Explanation types used to justify each chosen hotel.

between the type of explanation given and the chosen option. Therefore, we show in Figure 2 how much each explanation type is adopted for each individual hotel.

We make three main observations by analysing the obtained the results. First, it can be seen that most of the participants that choose the “Hotel 91” (61.11%) justify their decision by giving information of a critical attribute, as price is an attribute extremely relevant for these participants, and what matters for them is basically that this hotel is the cheapest one. Some participants provide further positive information about the hotel (*one-sided reasons*, 22.22%), besides stating that it is the cheapest one — they provide other positive aspects that complements the fact that it is cheapest, i.e. they show that even though the hotel is the cheapest, the quality that they require is not compromised.

The second observation is related to the hotels of the G-3 group. As we expected, the main adopted explanation types for choosing them are *one-sided reasons* and *pros and cons*, as it can be seen in Figure 2 and is shown in more detail in Table 6. The first explanation type is used to show that a whole set of hotel characteristics is responsible for the choice made. In general, participants that choose the “Econo Lodge Times Square” exclude the cheapest hotel from the set of hotels being considered in the decision, and explain the benefits of this hotel to show that this hotel is suitable for them, i.e. there is no reason to pay more for another option if this hotel already provides what the participant wants. On the other hand, participants that choose “The Hotel at Times Square” make a detailed analysis of this hotel against the “Econo Lodge Times Square,” i.e. they discuss

Hotel	One-sided Reasons	Pros and Cons	Total
Econo Lodge Times Square	55.77%	25.00%	80.77%
The Hotel at Times Square	26.32%	63.16%	89.47%

Table 6: Main explanation types used for justifying hotels of the G-3 group.

their *pros and cons*, and show that the higher price of the former justifies the benefits it provides, when compared against the latter. With respect to these two options, we point out one last comment: there are two participants (3.85%) that used dominance to justify why they chose “Econo Lodge Times Square.” The participants ignored attributes that are not relevant for them, creating a scenario in which this hotel dominates all the others.

Finally, we discuss the results obtained for the dominated option and the most expensive option. It can be seen that there is no most adopted explanation type, and participants adopt different explanation types for justifying them. Only few participants choose these two options and, as it is not obvious why these options should be chosen, the participants give their particular explanations to justify this decision. In the first case, “Comfort Inn Times Square,” some participants are vague and say that they choose this hotel because it has the best cost-benefit relationship without giving details. The remaining ones use as arguments the two attributes that this hotel is better than “The Hotel at Times Square,” i.e. parking price and room size. The room size argument is also used with the expression of *intuition*: as the room is bigger, and the price is higher, the hotel “apparently” provides more comfort. For this same reason, some participants choose the 4-star “Renaissance New York Hotel 57,” as comfort is the most important issue for them, and they are not concerned with price, and in their justification they explain this situation, i.e. for them the price of the hotel justifies the possible comfort it offers, and this is assumed because of the hotel stars. In one case, a participant say that she prefer the most expensive (critical attribute), as she wants to maximise comfort.

RQ3: Do users use a pattern to justify the rejected (not chosen) options? Now, that we have already addressed the research questions related to choosing an option, we focus on the rejected options. By analysing justifications for rejecting options, we observe the same explanation types used for justifying the chosen option. The description given for our set of explanation types show that they can also be applied for rejecting options, e.g. if an option does not satisfy the minimum requirements, then it is rejected due to this reason. In Table 7, we show examples of how each of these explanation types is used in the context of option rejection.

As it is the case with justifications for accepting an option, we also conclude that participants do use patterns for constructing arguments to reject options, and we next analyse the relationship between the adopted explanation types and the options involved in the decision-making process.

RQ4: Is there a relationship between the type of explanation given to reject options and the rejected or chosen option? In order to understand how participants choose a particular explanation type, we analyse the relationship between the types adopted to justify rejected options from two perspectives. The first consists of analysing justifications for rejection by relating them to the hotel that is rejected, i.e. we observe

Explanation Type	Example of Justification for Rejection
Critical attribute	<i>There are other options cheaper than H_i.</i>
Dominance	<i>There is no reason for choosing H_i, as it is worse in all aspects than H_j.</i>
Main reason	<i>I did not choose H_i because it does not offer the benefit a_i.</i>
Minimum requirements	<i>H_i is too expensive.</i>
One-sided Reasons	<i>I did not choose H_i because it has the disadvantages a_i and a_j.</i>
Pros and Cons	<i>Even though H_i provides the benefits a_i and a_j, its price does not compensate it.</i>

Table 7: Example of Justification for Rejection.

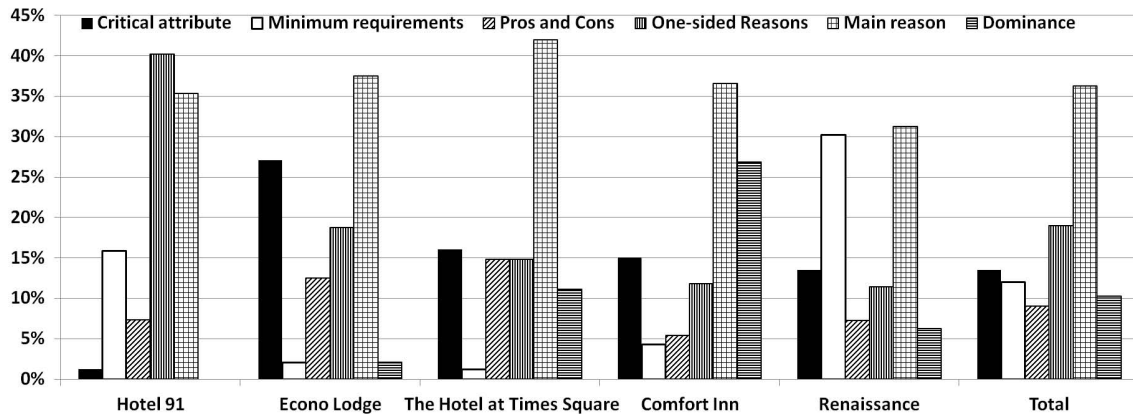
which explanation types are adopted to reject a particular hotel. The second perspective groups justifications according to the chosen hotel, i.e. we observe which explanation types are adopted to reject other options according to a particular chosen hotel. These two discussed views of justifications for rejection are presented in Figures 3(a) and 3(b).

There are many interesting aspects that can be observed in the collected data. *Critical attribute* is the type of explanation used when the decision is guided by it. For instance, if the participant wants to minimise price, the justification for the acceptance is that the chosen hotel is the cheapest, and the justification for the remaining rejected hotels is that they are more expensive (than the chosen hotel). Similarly, this situation happens with the more expensive hotel, in which the participant wants to maximise the price (as a proxy to the comfort maximisation).

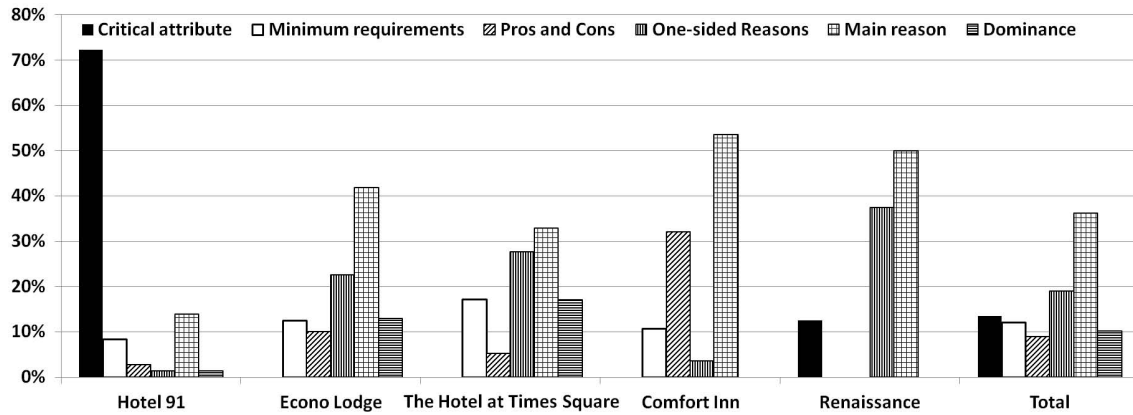
Dominance, on the other hand, is adopted when the chosen option dominates the rejected option, i.e. the comparison made in the explanation is always comparing the chosen option with the others. In many situations, preferences (hidden in justifications) of participants, who choose “Econo Lodge Times Square,” indicate that “The Hotel at Times Square” dominates “Comfort Inn Times Square,” however, this is not given as an argument to discard the last, but the participants seek for an explanation why “Econo Lodge Times Square” is better than “Comfort Inn Times Square” (*Evidence A*). Dominance is used as an argument by participants that choose “Econo Lodge Times Square” when the set of attributes that matter for the participants indicates that this option dominates both “The Hotel at Times Square” and “Comfort Inn Times Square.”

Some participants have hard constraints that they require to be satisfied by the chosen hotel, such as a maximum price that they are willing to pay, a minimum distance from the city centre or a minimum number of stars. In these situations, an option is rejected regardless the remaining options, and the justification given is that the option does not satisfy the participant *minimum requirements*.

There are three remaining types of explanation, namely *main reason*, *one-sided reasons* and *pros and cons*. The first two indicate that there is an attribute (or a set of, in case of one-sided reasons) that is really important for the participant that, even though it is not part of a hard constraint, plays a *decisive role* in the decision, i.e. because of this (these) attribute(s), the option is being rejected. This set of attributes is *kept as simple as possible* (*Evidence B*), and this can be seen in many justifications. For instance, some participants



(a) Grouped by rejected hotel.



(b) Grouped by chosen hotel.

Figure 3: Explanation types used to justify each chosen hotel.

that choose “Econo Lodge Times Square” reject “The Hotel at Times Square” and “Comfort Inn Times Square” because they do not have a refrigerator and are more expensive (than the chosen hotel). But, for justifying the “Renaissance New York Hotel 57” (which also does not have a refrigerator), they argue only that it is more expensive. It is important to note that the explanations given for *The Hotel at Times Square* and *Comfort Inn* are exactly the same, and there are many other cases in which the same explanation is given for different options rejected for the same reason (*Evidence C*). Finally, *pros and cons* are given as rejection arguments by participants in two situations: (i) it is their style of giving explanation, and they expose pros and cons for all the options; and (ii) the decision between two (or three options) is really hard, so they expose these options pros and cons to show that the chosen option has the best cost-benefit relationship for them. Thus, *pros and cons* are used only in the *absence of a decisive subset of attributes* (*Evidence D*), not considering participants individual styles.

In this way, the justification given for rejecting an option depends on both the chosen and rejected options, as the explanation given typically justifies why the rejected option is worse than the chosen one. Only in those cases in which the option is rejected due to a hard constraint (*minimum requirements*), the rejection explanation depends only on the option being rejected.

Further Observations. While analysing the collected data, we also identify other relevant characteristics present in the provided justifications. We describe each of these characteristics below, most of which can be used to suggest informal arguments for systematic approaches to decision-making.

Explicit trade-off (TO). As already mentioned in the description of the *pros and cons* explanation type, some participants state that the chosen hotel has the best cost-benefit relationship (or not the best, for rejecting a hotel), and sometimes just provide this argument without any details, e.g. “*For a trip like this, it seems the best cost-benefit among the 3-star hotels.*”

Preferences mentioned (PREF). Participants, when requested to justify their decision, provide arguments that are constructed based on their preferences (*Evidence E*), for example, a participant argue “*Absence of a fitness centre*” to justify a rejection, but this is due to the participant preference for a hotel *with* a fitness centre — and in some cases, participants make their preference explicit.

Insignificant difference (ID). The “Econo Lodge Times Square” has a US\$5.00 difference from “The Hotel at Times Square.” While some participants argue that the benefits provided by the second does not compensate the price difference, others, who have chosen the second, state that the price difference is insignificant, as it is very small, and both hotel prices can be considered the same. The same applies for room size or location, from the perspective of some participants.

Intuition (INT). One interesting characteristic of some provided justifications consists of inferring information of the hotel without any basis, i.e. some participants use their intuition to choose a hotel. For instance, one participant that chose “Econo Lodge Times Square” justify the rejection of “The Hotel at Times Square” by saying “*The name The Hotel seems to provide quality and, consequently, high price.*”

Price as a first class attribute (PRICE). The majority of participants (92%) mention the attribute “price” in their justifications, and evaluate options by comparing this attribute with all the other ones. This indicates that *cost* (which can also be time, effort, etc.) is not seen as any disadvantage that an option has when compared to another, but a fixed attribute that should be treated differently in the provided explanations (*Evidence F*).

Irrelevant attributes (IRR). When participants choose a hotel that does not offer as many benefits as the other, mainly when they choose the cheapest hotel, they use as arguments that those benefits are not important for them and, as they do not care about them, there is not reason for paying more for something that will not be used. Irrelevant attributes are mentioned in both acceptance and rejection justifications. For supporting a choice, participants state: “*Even though hotel H_i does not offer attribute a_i , this is not important to me,*” and for rejecting an option, they say “*Even though hotel H_i offers attribute a_i , this is not useful for me.*”

In Table 8, we show the percentage of participants whose justifications presented this identified characteristics. The table is split into each chosen hotel, and rows of each separate sub-table is related to the justification provided for each separate hotel. We highlight in gray the hotel that was chosen, therefore the row of a highlighted first cell is associated with justifications for acceptance.

As mentioned before, it can be seen that price should be treated as a first class attribute in explanations, as it is a crucial factor considered in the decision. In cases that a higher price is chosen, but this difference is very small, many participants acknowledge this fact. When the chosen option has a lower price, benefits provided by other options may be relevant to be mentioned, even though the decision maker does not care about it. In cases that pros and cons of a set of options make the decision hard, an explicit statement that a particular option has the best cost-benefit relationship might be helpful. Finally, participants typically do not support their arguments with their preferences.

We have now made observations based on the obtained data, and in next section we present conclusions that we extracted from it considering our data interpretation. These conclusions are used later to propose a set of guidelines and patterns to be followed by explanation generation approaches.

4 Interpretation

Our study investigates justifications given by people as explanations for their choices that can be used to convince another person to accept the choice made. The justifications give reasons to accept or to reject options. In this section, we provide an interpretation for our analysis, which explains how participants, and more generally people, choose a particular type of explanation for their decisions. We start by discussing the justifications given for choosing an option.

4.1 Explanation for Acceptance

We have provided five different options for participants, chosen in order to characterise options with certain particularities. Options that have quality lower than most of the other

Hotel 91						
Reason for	TO	PREF	ID	INT	PRICE	IRR
Hotel 91	0.00%	0.00%	0.00%	0.00%	100.00%	16.67%
Econo Lodge	5.56%	0.00%	0.00%	0.00%	77.78%	11.11%
The Hotel	0.00%	0.00%	0.00%	0.00%	83.33%	16.67%
Comfort Inn	0.00%	0.00%	0.00%	0.00%	83.33%	16.67%
Renaissance	0.00%	0.00%	0.00%	0.00%	94.44%	11.11%
Econo Lodge Times Square						
Reason for	TO	PREF	ID	INT	PRICE	IRR
Hotel 91	1.92%	3.85%	0.00%	3.85%	15.38%	0.00%
Econo Lodge	19.23%	1.92%	5.77%	1.92%	76.92%	1.92%
The Hotel	3.85%	3.85%	7.69%	1.92%	82.69%	13.46%
Comfort Inn	3.85%	1.92%	1.92%	0.00%	88.46%	7.69%
Renaissance	0.00%	0.00%	0.00%	0.00%	96.15%	13.46%
The Hotel at Times Square						
Reason for	TO	PREF	ID	INT	PRICE	IRR
Hotel 91	5.26%	0.00%	0.00%	0.00%	5.26%	0.00%
Econo Lodge	0.00%	5.26%	47.37%	5.26%	42.11%	0.00%
The Hotel	36.84%	10.53%	5.26%	5.26%	63.16%	0.00%
Comfort Inn	0.00%	5.26%	5.26%	0.00%	68.42%	5.26%
Renaissance	5.26%	0.00%	0.00%	0.00%	73.68%	0.00%
Comfort Inn Times Square						
Reason for	TO	PREF	ID	INT	PRICE	IRR
Hotel 91	28.57%	14.29%	0.00%	14.29%	28.57%	0.00%
Econo Lodge	28.57%	14.29%	0.00%	0.00%	14.29%	0.00%
The Hotel	28.57%	0.00%	0.00%	0.00%	28.57%	0.00%
Comfort Inn	28.57%	0.00%	0.00%	0.00%	42.86%	14.29%
Renaissance	14.29%	0.00%	0.00%	0.00%	85.71%	0.00%
Renaissance New York Hotel 57						
Reason for	TO	PREF	ID	INT	PRICE	IRR
Hotel 91	0.00%	25.00%	0.00%	25.00%	25.00%	0.00%
Econo Lodge	0.00%	25.00%	0.00%	0.00%	25.00%	0.00%
The Hotel	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Comfort Inn	0.00%	0.00%	0.00%	0.00%	25.00%	0.00%
Renaissance	0.00%	0.00%	0.00%	25.00%	0.00%	0.00%

Table 8: Results for additional characteristics observed in justifications.

options available, but also lower price, are justified by the “*critical attribute*,” which in this case is price. Therefore, other option details are not relevant, and do not need to be part of the explanations. In some cases, as the cheapest option being offered in our study already provides some comfort (as the hotel has 2 stars, shower, etc.), some participants mentioned that they choose the cheapest hotel, as it satisfies all their minimum requirements. The attribute *price* should always be mentioned in explanations to support the choice for this kind of options. Moreover, irrelevant attributes can be mentioned as part of explanations for this kind of option, in order to make the argumentation stronger. Explanations for the most expensive hotel (and with highest quality), although also characterised as an extreme option, does not follow this same reasoning. This kind of option is justified with all positive aspects it provides, or a main one — that is generally only offered by expensive options. In this scenario, price is typically not mentioned, as it is not a concern.

The majority of the participants have reduced their choice to options that require trade-off, by first discarding some options due to a minimum requirement, such as maximum price. There are mainly two kinds of justifications given for supporting a choice from this set of options. When the chosen option is the cheapest in this set, people use only the main reason or the benefits provided by the chosen option. On the other hand, when the option is not the cheapest, a more detailed explanation is needed, therefore the pros and cons related to the chosen option should be exposed, with the aim of showing that pros justify cons. For making this argument stronger, explicitly mentioning that the option provides the best cost-benefit relationship is helpful. In addition, when balancing pros and cons, price should be mentioned as a first class attribute.

The last kind of option, namely the dominated option, is never chosen. In our study, the option that represents this group of options (G-1) has few advantages with respect to the others, which are the room size and parking price. As these attributes are extremely relevant for some of the participants, they supported their choice evaluating pros and cons or by stating these essential aspects, in summary, this option was considered as an option of the G-3 group by those participants.

4.2 Explanation for Rejection

Having described how explanations are constructed to justify a chosen option, we now discuss how explanations for rejecting an option are built. A very important aspect of the explanation for rejecting an option, is that the *explanation is driven by the chosen option* in many cases.

As in explanations for acceptance, when an extreme option is chosen, explanations for rejecting options have a different behaviour for the two different extreme directions. When the chosen option is the cheapest and with lowest quality, the explanation given for all the remaining options are the same, which says that the other options are more expensive than the chosen one. So the typical explanations for this scenario are “*because of the price*” or “*because this option is more expensive [than the chosen option]*.” When the chosen option is the most expensive option and with the highest quality, the arguments used to reject the remaining options are the most decisive aspects that are not offered by these options. Assuming that for one individual “fitness centre” is extremely important, and a “bar” is also really important, but not as much as “fitness centre.” Therefore, in this scenario, “Econo Lodge Times Square” should be rejected because it does not have “fitness centre,” even though it also does not have a “bar.” Both aspects should be mentioned in the explanation

if they, individually, would not change the choice made. For this kind of extreme option, dominance is never used because, as the chosen option is the most expensive, it does not dominate any other option.

Next, we discuss explanation for rejection when the chosen option is part of the G-3 group, which has a different explanation for each kind of rejected option. When there is a dominated option, this is the argument to be given, however, *only if the chosen option dominates this option*. In addition, people that choose an option from the G-3 group, typically discard some options due to a cut-off value, i.e. options that do not satisfy minimum requirements, which are often part of the G-2 group. In these situations, the reason for cutting this option off should be given as the explanation, such as “*too expensive*” or “*too far away*.” Finally, to reject other acceptable options, the reasoning is similar to the one described for the most expensive option, i.e. the decisive aspects of the option must be exposed. If pros and cons of the rejected option have a similar balance to the chosen option, then this should be discussed in detail in the explanation, in order to show that pros do not compensate cons.

5 Guidelines and Patterns

This study provides us with means for understanding how users construct arguments to justify a choice made, by explaining why an option is chosen and why the remaining ones are rejected. In addition, based on the conclusion derived from this study, our ultimate goal is to provide guidance that serve as a basis for the development of explanation approaches and, with this purpose, we introduce guidelines and patterns derived from our study in this section. For each guideline, we indicate the evidence that supports it.

5.1 Guidelines

1. Provide chosen-option-oriented explanations. (*Evidence A*) The explanation generation process must be guided by a previously chosen option. The goal of the explanation is not to expose all the reasoning process used to make the decision, but to provide the main arguments that justify a chosen option and reject the remaining ones. After the choice is made, the explanations given should answer two main questions: (i) what makes the chosen option better than the others? (ii) what makes other options worse than the chosen option.?

An example of the application of this guidelines is the case of domination. Consider the scenario in which an option A is chosen, it is considered better than B and C , and B dominates C , but A does not dominate C . An easy way of explaining why C should be rejected is stating that B dominates C ; however, the explanation that should be given is to justify why A is better than C . This example can be applied for the hotels “Econo Lodge Times Square” (A), “The Hotel at Times Square” (B), and “Comfort Inn Times Square” (C).

2. Keep it simple. (*Evidence B*) The explanation given for the user should be as simple as possible, being as simple as justifying the decision with a single sentence, e.g. *A is the cheapest option*. Therefore, the less complex the explanation can be, the better. The next

three guidelines are associated with this guideline, as they show concrete ways of keeping the explanation simple.

3. Focus on the most relevant criteria. (*Evidence D*) In the given explanation, only the *decisive criteria* should be mentioned, i.e. the minimum set of attributes that makes an option selected or rejected. These decisive criteria should be derived from the comparison of the chosen option against the others.

For example, “Econo Lodge Times Square” is chosen because it has a refrigerator and its price. “The Hotel at Times Square” and “Comfort Inn Times Square” are rejected also due to these two reasons; however, even though “Renaissance New York Hotel 57” also does not have a refrigerator, the explanation associated with it mentions only its price, as this hotel is rejected independently of the refrigerator attribute.

4. Group similar options. (*Evidence C*) In many situations, an explanation to reject an option can also be given to reject other options. So, rejected options should be grouped when they are rejected for the same reason, and presented as a group and not individually. For instance, in the example above, “The Hotel at Times Square” and “Comfort Inn Times Square” should be put into a single group and receive the explanations that they are more expensive than the chosen options and do not have a refrigerator.

5. Back up explanations with user preferences, but provide them only if asked. (*Evidence E*) Explanations expose option characteristics that are relevant for the decision made. But, why are these characteristics relevant? Because of the preferences being considered in the decision-making process. If an explanation provided for a choice is “because the chosen option is the cheapest from the set of those available,” one might challenge why this is important, and the answer for it is that, for example, the user most important preference is to minimise costs. People usually do not explicitly state their preferences to justify their decisions; but, if a decision is made on a person’s behalf, it is fundamental to back an explanation up with her preferences. However, this information is not always needed, and the simplest the explanation is, the better, therefore preferences must be provided as part of explanations only upon request.

6. Use *cost* as a first class attribute. (*Evidence F*) An option is chosen by an individual when she believes that the cost being paid for that option compensates the benefits it provides. Benefits is a subset of all possible positive characteristics that an option can have, for example, a hotel that provides breakfast, “big” room, “good” location, etc; nevertheless benefits always come with a cost, which in the hotel case is its price, but in other scenarios it can be time or effort. This trade-off between benefits and costs is the key issue in the process of decision-making, therefore the option attributes that define the option costs should be made explicit and used as a first class attribute in the explanation provided to justify a decision.

5.2 Patterns

Based on our study, we derived patterns of explanations, which can be used for supporting a decision made by a software system. Moreover, we identified the components these

patterns must have, which comprise a template for an explanation pattern catalog. These components are: (i) a *classification*; (ii) a *context* in which the pattern should be applied; (iii) a *template* for the explanation; (iv) the pattern *description*; (v) an *example*; (vi) *preferences* that back up the explanation; and (vii) optionally, *extensions* to the pattern. Patterns are classified (item (i)) according to three attributes, explained below.

- *Explanation goal: accept/reject/both.* An explanation can have three different goals, as it can be given to justify a chosen option (*accept*), justify why an option is rejected or not chosen (*reject*); or give an explanation that covers both goals (*both*), typically by exposing the overall reasoning behind the decision.
- *Target: decision/option.* An explanation pattern can provide guidance to generate an argument that justifies the decision as a whole (*decision*), or the generation of an argument that supports the acceptance or rejection of a single option (*option*).
- *Position: absolute/relative.* When a pattern target is *option*, the explanation given can be based solely in the target option, or make a statement that explicitly compares the option to another one. In the first case, the explanation position is *absolute*, and in the second it is said *relative*.

Patterns are presented ordered according to their complexity, i.e. the simpler the explanation associated with a pattern is, the earlier it is presented. According to our second guideline, the explanation should be as simple as possible, therefore, if two patterns can be used in a particular situation, the simplest one must be applied.

5.2.1 Pattern 1: Critical Attribute

Classification:

- *Explanation goal:* both
- *Target:* decision

Context: this pattern is applied in two situations: (i) there is an attribute that is extremely important for the user and this is the only one to be taken into consideration; (ii) all available options satisfy all constraints and there is one criterion to choose the best.

Template:

Option chosen option was chosen because it has the best value for critical attribute.

Description: some users have a single criterion to choose an option, and may additionally have a set of constraints that is satisfied by all options. In these situations, the decision becomes trivial, as well as its associated explanation, which consists of stating that the chosen option was selected according to this single criterion.

Example: the user wants to buy the cheapest flight from Rio de Janeiro to London, regardless number of stops, flying time, airline company and so on.

Back up preference: preference that establishes criterion used to make the choice, and possibly other preferences satisfied by all options.

Extensions: rejected options might have characteristics not present in the chosen option, which are typically considered as benefits; however, for the decision maker, these characteristics are irrelevant. The explanation in this case can be extended by stating that the chosen option does not have such characteristics, but these are not relevant given the provided preferences.

5.2.2 Pattern 2: Cut-off

Classification:

- *Explanation goal:* reject
- *Target:* option
- *Position:* absolute

Context: an option does not satisfy a user requirement (hard constraint), or does not satisfy a constraint that is very relevant for the user, but there are other options that satisfy it.

Template:

<i>Option <u>rejected option</u> was rejected because it does not satisfy constraints associated with <u>attribute</u>.</i>

Description: in many situations, users have a set of requirements that **must** be satisfied (or hard constraints), and therefore, an option that does not satisfy at least one of these requirements cannot be chosen. As not satisfying at least one of these constraints is enough for rejecting an option, only the constraint associated with the most important attribute for the user is part of the explanation. In addition, there are cases in which some requirements are not hard constraints, because the user might accept options that do not satisfy it when there is no other available options. When there are options that satisfy these “almost-hard” constraints, options that do not satisfy it can be rejected with an explanation of a not satisfied hard constraint.

Example: the maximum amount of money that a user is spending for a hotel is US\$300.00 for two nights. A hotel that costs US\$350.00 for two nights is rejected because it is too expensive, i.e. due to a cut-off value.

Back up preference: not satisfied hard-constraints (or “almost-hard”).

5.2.3 Pattern 3: Domination

Classification:

- *Explanation goal:* reject
- *Target:* option
- *Position:* relative

Context: the chosen option dominates a rejected option.

Template:

There is no reason to choose option rejected option, as option chosen option is better than it in all aspects, including cost.

Description: when an option dominates another, there is no argument that justifies considering the latter better than the former. Therefore, exposing this fact is enough to explain the rejection of the dominated option. Nevertheless, the domination argument must be used only if the option that dominates the rejected option is the chosen one.

Example: there are two hotel rooms available for the user: standard queen room and superior queen room. The difference between them is that the superior queen room is bigger, it has a sitting area, it has a bath besides the shower, and, because of these extra features, it is more expensive. All user constraints are satisfied by both options, and she does not care about these three extra features, but cares about the price. Therefore, according to the preferences of this user, the superior queen room is dominated by the standard queen room.

Back up preference: preferences that establish that individual attributes of the chosen option are considered better than the rejected one.

Extensions: in this pattern, attributes that users do not care about might be mentioned to support the domination relationship. See more details in the extensions of Pattern 1: Critical Attribute.

5.2.4 Pattern 4: Minimum Requirements⁻

Classification:

- *Explanation goal:* reject
- *Target:* option
- *Position:* relative

Context: user established a set of minimum requirements for options, and a way of choosing from the ones that satisfy it. According to those requirements, some options were discarded. Other option attributes might have been used for making the decision, but only one of them makes the difference.

Template:

Even though option rejected option satisfies all your requirements, it has a worse value for attribute than option chosen option.

Description: this pattern addresses justifying rejecting options that differ only by a single attribute (that matters for the user) from the chosen option, and these rejected options, and also the chosen one, satisfy a set of user requirements. This scenario might happen when provided preferences consist only of these minimum requirements plus a preference to choose among them or when, even though when the decision involves a careful evaluation of pros and cons of each individual option, after choosing a particular option, one or more options are distinguished from it by a single attribute, relevant for the decision. Options that follow in this category, can be rejected using a simple explanation — rather than a more complex one, possibly used in the decision-making process — consisting only of the requirements satisfaction and a single attribute that makes the difference.

Example: a user wants to stay in a 2-star hotel, whose price is up to US\$150.00 per night, and is within the city centre, with breakfast. Given these requirements, she wants the cheapest one. The chosen hotel costs US\$120.00 per night, there is one that costs US\$130.00; and all the remaining attributes have the same values for both options. The rejection of the second hotel is explained using this pattern.

Back up preference: user requirements and preference used to choose from the options.

5.2.5 Pattern 5: Minimum Requirements⁺

Classification:

- *Explanation goal:* accept
- *Target:* option
- *Position:* absolute

Context: user established a set of minimum requirements for options, and a way of choosing from the ones that satisfy it. According to those requirements, some options were discarded. Other option attributes might have been used for making the decision, but only one of them that made the difference.

Template:

<i>Besides satisfying all your requirements, option <u>chosen option</u> has the best value for <u>attribute</u>.</i>

Description: users, in certain decisions, establish a set of minimum requirements that reduces the set of available options to a subset, in which options differ only by a single criterion from the chosen option. This is a situation in which the decision becomes easy, and also the explanation, which consists of acknowledging users that, from those options that satisfy their requirements, the chosen option is the preferred one according to a particular criterion.

Example: a user wants to stay in a 2-star hotel, whose price is up to US\$150.00 per night, and is within the city centre, with breakfast. Given these requirements, she wants the cheapest one. The chosen hotel costs US\$100.00 per night, and other available hotels provide the same features and are more expensive. Therefore, the explanation for the chosen hotel is that it is the cheapest from those that satisfy requirements.

Back up preference: user requirements and preference used to choose from the hotels.

Extensions: in this pattern, attributes that users do not care about might be mentioned to support the decision. See more details in the extensions of Pattern 1: Critical Attribute.

5.2.6 Pattern 6: Decisive Criteria

Classification:

- *Explanation goal:* both
- *Target:* option
- *Position:* absolute

Context: even though there are other attributes that contribute for the option acceptance (or rejection), there is a subset of them that would confirm this decision regardless of the values of the other attributes.

Template:

Option option was [chosen | rejected] because of its set of decisive attributes .

Description: options, when compared, might have different pros and cons. However, some attributes are the most decisive in the decision (according to their value and importance), while others — which can make a difference in particular cases — do not impact on the decision between two options. Therefore, the only attributes that must be part of the explanation are those that impact on the decision, leaving aside remainder.

Example: three hotel options are given for a user: (i) hotel *A* is a 3-star hotel, cheaper than the other two options and has a refrigerator in the room; (ii) hotel *B* is also a 3-star hotel, more expensive than the former, with a better location; and (iii) hotel *C* is a luxury 4-star hotel, much more expensive than the others and, like hotel *B*, does not have a refrigerator. While the rejection of *B* is justified by the absence of the refrigerator *and* its price; the rejection of *C* is justified only because of its price, as this is the decisive criterion for not choosing it.

Back up preference: preferences over the set of decisive attributes.

5.2.7 Pattern 7: Trade-off Resolution

Classification:

- *Explanation goal:* both
- *Target:* option
- *Position:* absolute

Context: there is no set of attributes that is decisive.

Template:

Template for rejected options:

Even though option rejected option provides better pros than the chosen option, it has worse cons .

Template for the chosen option:

Even though option chosen option does not have the best value for cons , its values for pros compensate its cons.

Description: a set of decisive attributes does not exist in all situations. Options might provide different pros and cons in a way that all attributes are important for making the decision, therefore, all option attributes that differ for these options have to be evaluated, and their evaluation has to be informed to the user.

Example: a user is provided with two hotel options. Both of them are 3-star hotels, the first is cheaper (US\$115.00 per night) and its room has a refrigerator, and the second is

more expensive (US\$130.00 per night), better located (two blocks closer to the city centre), and has breakfast included. As, according to the user preference in our example, location has a higher priority than price, and their difference is strong enough when considering the included breakfast and the refrigerator, the second hotel is chosen.

Back up preference: all user preferences used to evaluate pros and cons of options that require trade-off analysis.

Extensions: in situations that pros and cons of the chosen option creates a balance that is very similar to the one of another option, it might be not obvious for the user why pros compensate cons. So, additionally, it can be explicitly said for the user, i.e. inform the user that the chosen option provides the best cost-benefit relationship.

6 Related Work

Recommender systems have different aims that should be achieved beyond accuracy, such as trust, effectiveness and satisfaction, and these have been addressed through explanations that expose the rationale behind the adopted recommendation approach (content-based, collaborative, or hybrid). For example, if a collaborative approach is adopted, the user may receive as an explanation a histogram of ratings of the product given by similar users. McSherry (McSherry 2005) focused on case-based reasoning approaches, in which products are seen as cases from which one should be selected when it is similar to the case provided by the user, and the explanation is based on selected similar cases. Another direction is explanation interfaces (Pu & Chen 2007), which organise recommended products in a way that causes trade-off situations to be resolved explicitly for users, thus facilitating the decision making process.

Even though explanations improve recommender systems, they currently focus on explaining the *means* used to obtain recommendations, but stating that “someone like you chose this product” or “you like similar products” is not sufficient for justifying a recommendation and for users to accept it. Characteristics of *good* explanations have been defined based on the analysis of existing approaches (Tintarev & Masthoff 2007), and these can be used as metrics to *evaluate existing approaches*. Our work, on the other hand, identifies *good* explanations that should be given to users, and can be used as guidance for elaborating new explanation approaches. The challenge of obtaining these explanations from existing recommendation approaches, however, still remains.

7 Final Considerations

In this paper, we have presented a study performed to understand how people justify their decisions, by giving explanations why they choose a particular option from the set of those available, and why remaining options are rejected. The study consisted of providing participants (100 people) with a set of carefully chosen hotel options, and requesting them to give reasons for the choice. Based on collected data, we have identified explanation types that are patterns of justifications given by people, and how they are selected to be given as explanation — for both chosen and rejected options. Assuming that explanations given by people are the explanations that users expect to receive as reasons for a choice, our study allowed us to propose a set of guidelines and patterns for the development of explanation approaches. Future work involves producing explanations for choices made

by our decision-making technique (Nunes, Miles, Luck & Lucena 2012), which takes into consideration this guidance derived from our study.

References

- Basili, V. R. & Rombach, H. D. (1988), ‘The tame project: towards improvement-oriented software environments’, *Software Engineering, IEEE Transactions on* **14**(6), 758–773.
- Basili, V. R., Selby, R. W. & Hutchens, D. H. (1986), ‘Experimentation in software engineering’, *IEEE Trans. Softw. Eng.* **12**(7), 733–743.
- Glaser, B. G. (1992), *Emergence vs. Forcing: Basics of Grounded Theory Analysis*, Sociology Press, Mill Valley, USA.
- Junker, U. (2004), Quickxplain: Preferred explanations and relaxations for over-constrained problems, *in* ‘AAAI’04’, AAAI Press, USA, pp. 167–172.
- Labreuche, C. (2011), ‘A general framework for explaining the results of a multi-attribute preference model’, *Artif. Intell.* **175**, 1410–1448.
- McNee, S. M., Riedl, J. & Konstan, J. A. (2006), Being accurate is not enough: how accuracy metrics have hurt recommender systems, *in* ‘CHI ’06’, ACM, pp. 1097–1101.
- McSherry, D. (2005), ‘Explanation in recommender systems’, *Artif. Intell. Rev.* **24**, 179–197.
- Nunes, I., Miles, S., Luck, M. & Lucena, C. (2012), User-centric preference-based decision making, *in* ‘AAMAS’12’. (to appear).
- Pu, P. & Chen, L. (2007), ‘Trust-inspiring explanation interfaces for recommender systems’, *Knowledge-Based Systems* **20**, 542–556.
- Simonson, I. & Tversky, A. (1992), ‘Choice in context: Tradeoff contrast and extremeness aversion’, *Journal of Marketing Research* **29**(3), 281–295.
- Tintarev, N. & Masthoff, J. (2007), A survey of explanations in recommender systems, *in* ‘Proceedings of the 2007 IEEE 23rd International Conference on Data Engineering Workshop’, IEEE Computer Society, Washington, DC, USA, pp. 801–810.

A Questionnaire

This appendix presents the questionnaire used in our study. Participants had the option of doing it either in English or in Portuguese. If participants tried to go back in the questionnaire, they were notified that changes would not be stored.

A.1 Introduction: Survey on Reasons for Choice

The purpose of this survey is to collect data that helps on understanding the **reasons for a choice**. The survey is completely anonymous and all information collected will be used solely for analysis within the context of this study. The survey has three steps and the estimated time for implementation of the survey is around 10 minutes.

Please, click on the image below to start the survey in English:

A.2 Part I: User Data

- Age: a positive integer;
- Gender: a value from {Male, Female};
- Country: a value from a provided list of countries;
- City: a string;
- Working/Studying Field: a string.

*All fields are mandatory.

A.3 Part II: Option Selection

Assume you are going to spend holidays in New York with a close friend (you do not mind sharing a bed with him/her). You were given the following hotel options, from which you have to choose one. Which would you choose?

NB1. Prices are in American dollars.

NB2. Provided options are based in real data, but details were changed for the purposes of this study.

- *Chosen Option*: {Hotel 91, Econo Lodge Times Square, The Hotel at Times Square, Comfort Inn Times Square, Renaissance New York Hotel 57};

Available options were given as shown in Figure 4.

A.4 Part III: Reasons for your Choice

Could you please tell us why did you choose “*chosen hotel*”, and why did you reject the other options? Assume that this **justification must be good enough so that is can be used to convince your friend that your choice is the best one**. Please, use full sentences.

- Why did you [*accept/reject*] “Hotel 91”?
[text area in which participants write their explanation]

	Hotel 91	Erono Lodge Times Square	The Hotel at Times Square	Comfort Inn Times Square	Renaissance New York Hotel 57
	2 Stars	3 Stars	3 Stars	3 Stars	4 Stars
	91 East Broadway, Lower East Side, NY 10002 New York	302 West 47th Street, Midtown, NY 10036 New York	59 West 46th Street, Midtown, NY 10036 New York	129 West 46th St, Midtown, NY 10036 New York	130 East 57th Street, Midtown, NY 10022 New York
Distance from Times Square	Map	Map	Map	Map	Map
Price (per night)	\$97.3	\$129.99	\$134.99	\$144.99	\$219.0
Check-in	From 15:00	15:00 - 23:30	15:00 - 00:00 hours	15:00 - 00:00 hours	From 15:00 hours
Check-out	Until 11:00	Until 12:00	Until 12:00 hours	Until 12:00 hours	Until 12:00 hours
Internet	Wired internet is available in the entire hotel and is free of charge.	Wi-fi is available in the entire hotel and is free of charge. Wired internet is available in the hotel rooms and is free of charge.	Wi-fi is available in the entire hotel and is free of charge.	Wi-fi is available in the entire hotel and is free of charge.	Wired internet is available in public areas and is free of charge. Wired internet is available in the hotel rooms and costs USD 16.95 per 24 hours
Parking	Public parking is possible at a location nearby (reservation is not needed) and costs USD 32 per day.	Public parking is possible at a location nearby (reservation is not needed) and costs USD 30 per day.	Public parking is possible at a location nearby (reservation is not needed) and costs USD 35 per day.	Public parking is possible at a location nearby (reservation is not needed) and costs USD 23 per day.	Private parking is possible on site (reservation is not needed) and costs USD 55 per day. Public parking is possible at a location nearby (reservation is not needed) and costs USD 55 per day.
24-hour Front Desk	X	X	X	X	X
Express Check-in/Check-out		X	X		X
Luggage Storage	X	X	X	X	X
Elevator	X	X	X	X	X
Bar			X		X
Restaurant					X
Laundry	X		X	X	X
Fitness Center			X	X	X
Business Center		X	X	X	X
Room Service					X
Breakfast in the Room					X
Room Type	Queen Room	Superior Queen Room	Queen Room	Queen Room	Queen Room
Breakfast Included		X	X	X	X
Safety Deposit Box		X	X	X	X
Air Conditioning	X	X	X	X	X
Heating	X	X	X	X	X
Iron		X	X	X	X
Hairdryer	X	X	X	X	X
Bath or Shower	X	X	X	X	X
Toilet	X	X	X	X	X
Alarm Clock or Wake up Service	X	X	X	X	X
Telephone	X	X	X	X	X
TV	X	X	X	X	X
Cable TV	X	X	X	X	X
Radio	X	X	X	X	X
Tea/Coffee Maker		X	X	X	X
Seating Area			X		X
Work Desk	X	X	X	X	X
Refrigerator		X			
Room size	19.0	21.0	21.0	22.0	24.0

Figure 4: Hotel Options.

- Why did you [*accept/reject*] “Econo Lodge Times Square”?
[text area in which participants write their explanation]
- Why did you [*accept/reject*] “The Hotel at Times Square”?
[text area in which participants write their explanation]
- Why did you [*accept/reject*] “Comfort Inn Times Square”?
[text area in which participants write their explanation]
- Why did you [*accept/reject*] “Renaissance New York Hotel 57”?
[text area in which participants write their explanation]

Available options are presented again (Figure 4).