11
Conclusion

Making decisions over a possibly huge set of options is part of many of the tasks that humans face in their everyday lives. Such decisions are not only time-consuming, but also require cognitive effort demanded by humans, as choosing an option from an available set of options often requires resolution of trade-offs. Moreover, as nowadays the number of options available to users is massive, analysing all available options goes beyond their cognitive limitations, making them often unsatisfied with their choices.

Approaches to representing and reasoning about preferences, as well as explanation approaches to justify decisions made by computer systems, have been proposed in order to support and automate decision making, and this is the context of this thesis. We proposed a new preference metamodel, which allows the representation of high-level preferences. This metamodel was developed based on a study involving almost 200 participants, whose goal was to understand how humans express preferences and the expressions they use. With the aim of making decisions in a similar way to humans, and based on preferences they explicitly provide, we also presented a decision making technique, which receives as input preferences in a language that is based on our metamodel and chooses an option from a set available, taking into account user-centric principles.

In order for users to understand why an option was chosen and to trust the decision, we proposed an explanation generation technique that uses models built by our decision making technique to justify choices. Our explanation approach is based on guidelines and patterns that we derived from a study to justify choices, which involved 100 participants. Finally, we presented the results of a user study, performed to evaluate the different aspects of our approach. This evaluation shows that (i) our preference language is adequate for users to express their preferences; (ii) our decision making technique makes choices that users consider as having good quality; and (iii) the provided explanations allows users to understand why the choice was made and improves decision confidence. By making a side-by-side comparison of our explanation technique with the two similar existing approaches, we were able to determine that ours is significantly better than one of them with respect to transparency and trust in choice. Although our explanation technique

has no significant difference with respect to the second compared approach, the comparison showed the particular scenarios in which our explanation technique needs to be improved. Therefore, limitations of our approach were also identified with this user study, not only associated with explanations. For example, we identified cases where users provide dyadic preferences, but also indicate that values referred to by those preferences are preferred to all other possible values for that attribute, and that explanations must mention important attributes associated with unsatisfied preferences.

Next, we detail the contributions of this thesis and discuss future work.

11.1 Contributions

As the result of the work presented in this thesis, many contributions can be enumerated, which are detailed next. Some of them serve as a basis or to evaluate our three main contributions: the preference metamodel, the decision making technique, and the explanation technique.

Study of How Humans Express Preferences. Our first study, presented in Chapter 2, provided a deeper understanding of how humans express their preferences about a domain. We analysed and discussed different aspects of preferences specified by the study participants, such as how useful the provided preferences are for making a choice on their behalf, and how preferences change after the participants face a concrete decision making situation. We also derived from this study common expressions that humans use to state their preferences.

High-level Preference Metamodel. Considering the results of our study of how humans express preferences, we proposed a preference metamodel (Chapter 3) — which also includes an ontology metamodel to represent application areas and a metamodel to represent propositional formulae — that allows representation of different types of preferences, such as constraints, goals and qualifying statements, which use expressive speech acts to indicate preference. Initial versions of this metamodel were published elsewhere (Nunes et al. 2010a, Nunes et al. 2010b). The metamodel is represented in UML, but is also formally specified using the Z notation.

Systematic Review of Reasoning about Preferences. As many different areas of computer science investigate preferences, we provided a systematic review of reasoning about preference approaches in Chapter 5, including work in the context of decision theory, artificial intelligence, constraint programming, databases

and semantic web. Each piece of work was presented following an evaluation framework, which facilitates their comparison, and we further analysed them, discussing their positive and negative aspects.

User-centric Preference-based Decision Making Technique. We proposed an automated decision making technique (Nunes et al. 2012a, Nunes et al. 2012b), presented in Chapter 6, that uses preferences expressed by users in a high-level language, to resolve trade-offs based on priorities provided by users combined with user-centric principles. Our technique provides the novelty of exploiting different natural language expressions and user-centric principles in automated decision making. These two particularities of our technique consist of two ways of significantly improving research in this area: while expressive speech acts and other expressions give valuable information that can be used to generate low-level preference representations, such as utility functions; our (and possibly others) user-centric principles can be used to reduce the amount of preferences obtained from users, as they can predict how users would resolve trade-offs. Moreover, these principles of human decision making explain situations in which a decision made by a human is "irrational" according to classical decision theory, and by taking these principles into account, automated systems can make decisions that are more acceptable to users.

Explanation Guidelines and Patterns. We performed a study (Nunes et al. 2012c) that allowed us to investigate how humans justify their decisions (Chapter 8), by arguing why they choose a particular option from the set of those available, and why the remaining options are rejected. With the data collected from this study, we have derived *guidelines* and *patterns* of explanations to be given to users to justify decisions made by the system.

Explanation Generation Technique. We presented a means of generating explanations for users to justify choices made by our decision making technique (Nunes et al. 2013) (Chapter 9). This is based on proposed guidelines and patterns, and provides a means of identifying parameters of explanation templates, which are part of the patterns. The technique not only identifies these parameters, but also provides an algorithm to choose which explanation should be used in different cases.

Evaluation. In order to evaluate different aspects of our approach, we performed a *user study*, presented in Chapter 10, in which participants had to specify their preferences, receive a choice (or recommendation), receive an explanation for that choice, and finally receive alternative explanations for it. With this study,

we evaluated our preference language, our decision making technique and our explanation generation technique, compared to existing approaches. The results of this study showed that our approach performs well in these three different aspects, but also identified its limitations.

11.2 Future Work

The contributions of this thesis advance research work on preferences, with the proposal of a preference metamodel, a novel decision making technique and an explanation generation approach. However, our work has limitations, leading to ongoing and future work, some aspects of which are discussed as follows.

Replication Studies. Each of our studies was performed in the context of one domain only: laptops (study of how humans express preferences); hotels (study of explanations to justify choice); and mobile phones (user study performed to evaluate our approach). In order to confirm the results obtained from these studies, it is important to replicate them in other domains and with other subjects. Moreover, the recommendation of other domain specialists could be taken into account in our first study.

Preference Consistency. One of the assumptions of our decision making technique is that the provided preferences are *consistent*. This is unlikely to happen, as confirmed by our study of how humans express preferences in which none of the preferences provided was inconsistent, and there was only one case in which preferences were inconsistent in our user study. However, when inconsistency *does* arise, wrong decisions can be made, with possibly inadequate explanations, which compromise user acceptance. Therefore, it is important to elaborate an approach that is able to check whether a set of provided preferences is consistent.

Preference Elicitation. Although our preference metamodel allows the derivation of a language in which users can express preferences in a way close to natural language, the activity of providing preferences requires a significant effort, as our user study revealed. Therefore, it is important to consider an approach that is able to implicitly capture an initial set of preferences, so that users can refine this set later. The advantage of using a high-level language is that users are able to understand the elicited preferences, and possibly make changes. Furthermore, as identified in our study of how humans express preferences, other kinds of support could be provided, e.g. reminding users of (generally important) attributes that were not mentioned.

Preferences not Covered by the Decision Making Technique. The language in which preferences that are the input of our decision making technique are expressed corresponds to a restricted version of our preference metamodel. As our metamodel is associated with expressions and terms that humans use to state preferences, we still constrain users while providing their preferences. Therefore, the investigation of how these limitations can be addressed is important. While some of these expressions that were not addressed should be handled during the decision making process, others may be associated with the translation or interpretation of terms used as proxy for others. For example, if the user wants to *maximise mobility* (of a laptop), this can be translated into the minimisation of laptop dimensions and weight. Therefore, an user interface language can be used as an abstraction layer on top of our decision making technique, in these situations.

Decision Making Technique Variability. The empirical evaluation of our technique showed that it is able to make a choice on behalf of the users as good as that made by a human domain expert. However, the conducted user study indicated that, even though our technique indicates good recommendations and helps users to make choices, it is not *always* able to make the right choice on their behalf, preventing them from delegating tasks to a system, which is our ultimate goal. As our decision making technique has variable parts (e.g. modifier scale, functions and weights), which were instantiated in this thesis after running the technique with different alternatives, it is future work to improve results by exploring this variability — either experimentally or by proposing individual-specific approaches — and investigating other user-centric principles to be adopted. Machine learning techniques can be adopted to instantiate these variables, as well as using fuzzy logics in the modifier scale to improve the interpretation of expressive speech acts and rates.

Inter-agent Decision Making. In this thesis, we have considered decision making in the context of a single agent. However, decisions can also be made in situations in which: (i) a single decision maker takes into account preferences of other agents, e.g. if a decision maker wants to decide at which hotel to stay, and one of her preferences is to stay at the same hotel as her colleague, then the preferences of her colleague should be also taken into account; and (ii) multiple decision makers must make a joint decision. Our goal is to extend our approach in order to address these situations, and this involves explicitly representing third-party preferences, reasoning about them and explaining decisions made in these scenarios.

Mixed Initiative Decision Making. Our approach receives as input a set of preferences and a set of available options and give as output a choice with an associated explanation. The automation of decision making is helpful as this task requires humans to demand high amounts of cognitive effort. However, as our evaluation showed, our technique is not always able to make an adequate choice. Besides improving our technique, better results can be achieved by adopting a mixed initiative approach: an initial set of preferences is given, which can be later refined according to the choice presented and explanations given. Moreover, this initial set of preferences can be obtained by mining preferences from historical data.

In summary, the work presented in this thesis advances work on automated decision making in three main directions: preference representation, preference reasoning and explanation generation. Clearly there is still much to do in order to make concrete our vision of having agents able to make decisions on behalf of users in a multi-agent scenario, but our work consists of a significant step towards this vision.