Eliciting GAI preference models with binary attributes aided by association rule mining

Sergio Queiroz[1]

Centro de Informática, Universidade Federal de Pernambuco, Av. Prof. Luiz Freire - s/n - Cidade Universitária - CEP 50740-540 - Recife (PE) - Brazil. E-mail: srmq@cin.ufpe.br

Abstract Generalized additive independent (GAI) preference models are, in many situations, more interesting than additive models, as GAI models allow interdependencies between attributes. However, they are more difficult to construct (elicit), not only because the number of questions needed to be asked increases, but also because we must know what the interdependencies are, i.e. the structure of the model. We introduce the use of association rules to select attributes and detect simple interactions between them during an interactive preference elicitation process that intends to build a GAI utility function for the preferences of the user of a recommender system. Using this strategy, we have built a recommender system prototype that suggests touristic sites in a city. We show that, after the elicitation, the recommendation problem can be solved as an instance of the non-linear generalized additive knapsack problem.

Keywords: Preference Elicitation, Multiattribute Recommender Systems, Generalized Additive Preferences, Non-Linear GAI Knapsack Problem.

1 Introduction

The development of decision support systems and web recommender systems has stressed the need for models that can handle users preferences and perform preference-based recommendation tasks. In some applications, like many recommender systems, the decision maker (DM) express his opinions about some alternatives (in a explicit or implicit manner) and the application will try to find among the available objects in a database the ones that are more probable to be liked by him. However, these approaches are only possible when the objects are available in a standardized form, as for example CDs, books, and DVDs [6]. When the objects are composed by configurable attributes, the enumeration of all possible configurations may be prohibitive. This condition characterizes the problems where the space of alternatives has a combinatorial structure, i.e. it is a Cartesian product of variables. In this case, an efficient form of representing and reasoning with preferences over combinatorial domains is needed.

In this way, several current works in preference modeling and decision theory aim at developing compact preference models that achieve a good trade-off between two conflicting goals: i) the power to describe sophisticated decision behaviors; and ii) the practical necessity of keeping the elicitation effort at an admissible level as well as the need for fast procedures to solve preference-based optimization problems.

A good compromise between model simplicity and representational power is reached by GAI (*generalized additive independence*) models [4, 2]. Having the DM preferences modeled as GAI functions brings many benefits. The descriptive power of such models allows interactions between attributes while preserving decomposability (provided that the preferences exhibit some structure). In this way, preferences may be compactly stored. Moreover, given an utility function, we can easily compute the utility of an item, allowing us to quickly calculate its value. But still, there is no easy way to elicit GAI functions.

In this work we use association rule mining [1] in order to select the attributes to expose to the DM at each moment in the elicitation process. Also, association rules are used as a heuristic to detect dependencies between attributes. Using this strategy, we have built a recommender system prototype that suggests touristic sites to visit in a city, under the constraint of the time that the DM will stay at the city. We show that, after the elicitation process, the best recommendations can be obtained as a solution to a non-linear GAI-decomposable knapsack problem. Formally, this correspond to solve this optimization problem:

$$\begin{aligned} \text{maximize} \quad & u(x) \\ \text{under the constraint that} \quad & \sum_{j=1}^{n} w_j x_j \leq c\,, \\ & x = (x_1, x_2, \ldots, x_n) \in \{0,1\}^n. \end{aligned} \tag{1}$$

We suppose that all coefficients w_j, as well as c are non-negative real numbers and that $u(x)$ is a GAI-decomposable utility function over the combinatorial domain of alternatives, i.e. $u : \{0,1\}^n \mapsto \mathbb{R}$. As the utility function $u(x)$ is not additively decomposable, the resulting knapsack problem is much more difficult to solve than for the additive case, given that we cannot exploit the structural independencies to develop more efficient algorithms [3]. We use GAI-Networks [5], a graphical model for the representation of GAI models, in order to efficiently treat the optimization problem.

References

[1] Rakesh Agrawal, Tomasz Imielinski, and Arun N. Swami. Mining association rules between sets of items in large databases. In *ACM SIGMOD*, pages 207–216, Washington, D.C., 26–28 May 1993.

[2] Fahiem Bacchus and Adam J. Grove. Graphical models for preference and utility. In *UAI*, pages 3–10, 1995.

[3] Kurt M. Bretthauer and Bala Shetty. The nonlinear knapsack problem – algorithms and applications. *European Journal of Operational Research*, 138(3):459–472, 2002.

[4] P. C. Fishburn. Interdependence and additivity in multivariate, unidimensional expected utility theory. *International Economic Review*, 8:335–342, 1967.

[5] Christophe Gonzales, Patrice Perny, and Sergio Queiroz. GAI-networks: Optimization, ranking and collective choice in combinatorial domains. *Foundations of computing and decision sciences*, 32(4):3–24, 2008.

[6] Francesco Ricci, Adriano Venturini, Dario Cavada, Nader Mirzadeh, Dennis Blaas, and Marisa Nones. Product recommendation with interactive query management and twofold similarity. In *ICCBR 2003*, volume 2689 of *Lecture Notes in Computer Science*, pages 479–493. Springer, 2003.