Case-Based Ranking for Decision Support Systems

Paolo Avesani, Sara Ferrari, and Angelo Susi

ITC-IRST,
Via Sommarive 18 - Loc. Pantè, I-38050 Povo, Trento, Italy
{*avesani,susi,sferrari*}*@itc.it*

Abstract. Very often a planning problem can be formulated as a ranking problem: i.e. to find an order relation over a set of alternatives. The ranking of a finite set of alternatives can be designed as a preference elicitation problem. While the case-based preference elicitation approach is more effective with respect to the first principle methods, still the scaling problem remains an open issue because the elicitation effort has a quadratic relation with the number of alternative cases. In this paper we propose a solution based on the machine learning techniques. We illustrate how a boosting algorithm can effectively estimate pairwise preferences and reduce the effort of the elicitation process. Experimental results, both on artificial data and a real world problem in the domain of civil defence, showed that a good trade-off can be achieved between the accuracy of the estimated preferences, and the elicitation effort of the end user.

1 Introduction

Very often a planning problem can be formulated as a ranking problem: i.e. to find an order relation over a set of alternatives. This order relation is usually interpreted as a sort of priority index that supports a selection policy.

Although the problem of finding an order over a set of items is ubiquitous, we have conducted our research in the domain of environmental decision support systems. A typical problem in civil defence is arranging emergency plans for a set of areas with high environmental hazard. Since preparing a single emergency plan requires a great deal of effort, a new issue arises: to establish an order over the candidate areas that indicates how to proceed. A similar problem arises when the training of civile population has to be planned: the simulation of a single emergency plan is expensive and has a great impact on everyday life.

Both scheduling the preparation of the emergency plans and scheduling the simulation of these plans require to take into account many factors, because the final order of a set of alternatives has to combine multiple criterias. For example, the relevance of an emergency plan requires combining the assessment of elements such as the environmenatal hazard and the systemic hazard. Usually it is hard to understand in advance the mutual dependecies of these criterias; in many cases the solution is proceeding with an elicitation process of all the relevant factors. The ranking problem can be designed as a preference elicitation problem.

The preference elicitation can be designed as an ex-ante process or as an ex-post process. In the former case, the preferences are formulated before we know the alternatives, while in the latter case, the preferences are formulated after establishing a finite set of alternatives. It is worthwhile to note that, in this context, the goal is not to find the top rank but to produce the total order.

The focus of this work is on the ex-post approach, where the real cases (i.e. the alternatives) play a key role in the preference elicitation process. In particular we are interested in developing a methodology where such a process is guided by the cases, not by general principles. We believe that by reasoning on cases, we may support the emergence of relevance factors in determining a rank over a finite set.

The main contribution of this paper is the design of a case-based elicitation process to support the planning problem in terms of rank assessment. More specifically, we adopt a pairwise approach in which the preference elicitation strategy is based on the comparison of two alternatives. In this paper, we propose a simple iterative model based on two automated steps: the development of a case selection policy, and the approximation of case preferences. The former aims at supporting a strategy for pairwise comparison, and the latter aims at reducing the elicitation effort.

A well known drawback of the case-based elicitation approach is the restriction on the case base size: this constraint results from the interactive process of acquiring preferences from the end users. As the size of the case base increases, the number of preferences that must be elicited grows quadratically. Usually the threshold is around ten cases [17].

The objective of this work is to illustrate how the use of machine learning techniques, like boosting [11], enables us to increase the threshold on the size of the case base while reducing the cognitive load of the end user. The preliminary results, both on artificial data and data obtained from real world experiments, showed that our solution is effective in finding a good balance between the size of the case base and the end users effort.

The structure of the paper is as follows. In the next section, we describe our motivation for this work. In the third section we discuss related work. In the fourth section we introduce the case-based model for the pairwise elicitation process. The fifth section explains how the machine learning techniques have been exploited in the case-based model, and the sixth and the seventh sections will present and discuss the experimental results on artificial and real world data respectively.

2 Motivation

The preference elicitation problem is well known in many scientific communities and it has been extensively investigated [3, 12, 14, 18]. The ex-ante approach is based on the notion of utility functions, which are designed to combine the various factors involved. This way of proceeding doesn't take into account the real set of cases where the utility function has to be applied. The basic assumption is that it is always possible to encode the first principle rationality in terms of an utility function.

A typical example is represented by the overlay model [2], widely used in the environmental sciences. In this domain, the following equation can be used to estimate the natural risk: $R = HVv$ (risk, hazard, vulnerability and value respectively). The values of these three features are acquired separately for each case and the risk value for each of them is derived by the equation. The total rank of the cases follows from the order relation of numbers.

Although the development of Geographical Information Systems (GIS) has made this approach very popular, there are many drawbacks [6]. First of all the unrealistic assumption that the acquisition of the feature values is context independent; a further related assumption that simplifies the complexity of the problem is concerned with the homogeneity of the case description. Moreover this method sometime produces only partial orders. When the main purpose of an order relation is supporting a decision-making step a partial rank could be not satisfactory.. In general, with the ex-ante approaches it is not possible to know in advance whether a predefined criteria will produce a total or a partial order because the final outcome is strongly related to the given set of cases.

To overcome this drawback an alternative approach has been proposed by Saaty [17] where the order relation is derived from the analysis of the cases. The pairwise comparison allows the simplification of the elicitation process that iteratively acquires the preference relation with respect to two alternative cases. Through this method the rationality of ranking can emerge from the comparison of the real cases and not from a generalized first principle. The elicitation process requires filling an half matrix where both rows and columns represent the same set of cases, and each entry a pair of cases (we assume that the preference structure is invariant with respect to the order of pair presentation).

The average effort to accomplish this task is related to the size of the case base: $|C|(|C|-1)/2$. It can become very hard for the end user to afford this approach when the number of cases increases. For this reason the literature has identified a threshold of 9 cases [17]. Given this constraint only an approach based on preference approximation will be sustainable once the size of the case base exceed the threshold above. A new challenge arises to design innovative strategies that do not require an exhaustive acquisition of all the pairwise preferences. Our objective is to find a trade-off between the elicitation effort and the accuracy of the estimated preferences.

In the fourth section we will introduce a model for a case-based preference elicitation process that supports a mixed-initiative strategy to achieve a total order of a finite set of cases. A step of pairwise analysis is in charge of the end user, and a step of preference approximation is in charge of the computer. The role of the machine learning techniques is twofold: estimating the rank, and shortening the iterative process of preference elicitation.

These two roles are closely related and together aim at making the case-based approach of preference elicitation competitive against the first principle one.

3 Related Work

The preference elicitation problem has received a growing attention in recent years with the advance of the Internet. The growing importance of e-commerce has promoted the development of personalization services that automatically elicit preferences from users.

However, preference elicitation in the domain of e-commerce is slightly different than in the domain of environmental resource management. Although existing work attempts to exploit the notion of case-based comparison [13] or to support the retrieval through an order relation [4], their foci are the selection of the top of a set of alternatives. In the domain of e-commerce, the objective is similar to that of information filtering, while in environmental planning the ultimate goal is a total ranking because it is more suitable to support a policy of case selection. Let us remind that in a decision support system the interest is on the rank of the whole set of items, while in a recommendation system the interest is on a single item, i.e. the top rank.

In this perspective, the work most similar to ours is the Analytical Hierarchy Process (AHP) and its variations [17, 15]. This method has been developed in the last decade in the domain of environmental resource management. The basic idea is to exploit a case-based pairwise comparison to elicit the preference structure and consequently to derive a rank according to a predefined goal. For all possible pairs from a finite set of cases, each expert is asked which of the two cases he prefers. Once the matrix of the pairwise preferences is filled, the corresponding rank is synthesized for the set of cases. It is not worthwhile to go in detail to explain the role of the hierarchy in AHP since it is not meaningful for the purpose of this paper, nevertheless it has to be stressed that this method is a successful alternative to the first principle approaches and has been applied in many domains [16].

Although it has been widely used, AHP suffers of two main drawbacks that prevents the method from scaling up on the number of cases, as we have already mentioned before. First, the cardinality of the set of cases has to be less than ten; second, the constraint on the preference range has to be elicited from a scale of nine values (to state how much a case has to be preferred with respect to the other). A non trivial boolean preference increases the cognitive overload of the user, making the method prone to the error because it assumes the capability to vote independently from the context. In a real world application, if the user is not familiar with the scale of values or the related semantic is not well defined in advance, the acquired information will be noisy. An additional objective of our work is to relax this constraint and to enable a boolean pairwise preference.

A recent work in the area of preference elicitation has developed a system that acquires the preference structure of the end user interactively: VEIL [3]. This system supports an incremental utility elicitation process on large datasets, to allow the building of a model of the user's preferences. In particular the system addresses the problem of visualizing many alternatives emphasizing some of their features considered of interest for the users. When the users are interacting with the system the whole set of cases are simultaneously presented to them. The user's operations are interpreted to implicitely derive from them the preference structure. This approach doesn't allow a rich visualization of the case descriptions. Moreover, the users are not supported in browsing the space of alternatives, and they are in charge of focusing their attention on relevant cases. In the next section we will show how our model takes this issue into account, and how the machine learning techniques can be effective in supporting the pair selection.

The works mentioned above do not fit the typical case-based reasoning loop [1], although we have stressed the role of cases in the elicitation process. The classical perspective of case-based reasoning has been adopted by work developed in the decision theory community. In [12] it is assumed that a set of preference structures, complete or incomplete, has been elicited from a set of users; the goal is the elicitation of the preferences of a new user incrementally, using the closest existing preference structures as potential defaults. This work explore the use of three different distance measures to compute the similarity between the preference structures. These similarity metrics among order relations are crucial to support the retrieval step of a CBR loop. In the case-based approach to decision support systems, we don't have an explicit step of retrieval; rather, an approximation step. The similarity metrics on preference structures are of potential use to design nearest neighbour classifiers supporting the preference approximation. In the following section, we explain the use of a classifier to approximate a partial unknown preference structure.

Our approach to the approximation of the preference elicitation is based on the work described in [10] where the RankBoost methodology is described and used in the context of information retrieval. In particular, the goal of this work is discovering a total order relation for a given set of instances described with some features, given a set of ordering functions for this features and a target user ordering function represented as a binary preference relation between couples of instances. Although these techniques have not been developed to support the approximation of preference structures, we will show in the section 5 how we have applied them to support the problem of preference elicitation.

To summarize, we are interested in exploiting the machine learning techniques mentioned above to combine a solution for the twofold problem: supporting the pairwise elicitation selection and the pairwise preferences approximation. In the next section, we explain the process that integrates the two steps in such a way that each one can benefit from the other.

4 A Model for Case-Based Preference Elicitation

The objectives of this work can be summarized as follows:

- **Elicitation advisory process.**
  Designing a case-based decision support system to support the interactive elicitation of a preference structure over a finite set of cases.
- **Pairwise preference elicitation.**
  Supporting the elicitation of a boolean preference structure through a pairwise comparison of cases.
- **Preference acquisition policy.**
  Developing an acquisition policy to support the expert in the process of preference elicitation: i.e. recommendation of the pairs of cases tha have to be analyzed first.
- **Preference learning step.**
  Exploiting learning techniques to approximate the unknown pair preferences taking advantage of the previously elicited preferences.
- **Error and effort minimization.**
  Achieving a good trade-off between the accuracy of approximated rank and the effort required to collect the explicit definition of the preference structure.

– **Mutual recursive benefit.**
Improving the accuracy of rank approximation taking advantage of an effective acquisition policy, and at the same time improving the acquisition policy taking advantage of the by product result of the learning step (i.e. the pairs where the learner fails).

Before introducing the interactive case-based preference elicitation model we briefly sketch the general architecture that underlies the generic process. Figure 1 illustrates the schema of the process; the abstract view highlights the interactive and the static components.

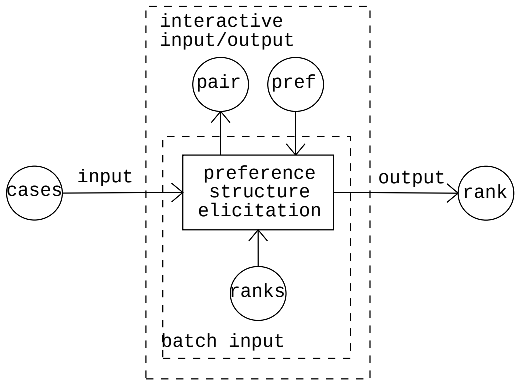


Fig. 1. Generic architecture for the preference elicitation process.

The types of data involved in the elicitation process can be detailed as follows:

– data - The input data represent the finite collection of cases that has to be ranked. The data are the initial input of a case-based preference elicitation process.
– pair - A pair is a couple of alternative cases where the end user preference is unknown. Given a pair the end user is interactively asked for preference elicitation. The pairs are generated run time by an automated process.
– pref - The pref is the order relation between two alternative cases elicited by the end user. The preference is formulated as a boolean choice from a pair, i.e. two alternative cases.
– ranks - The ranks are a collection of order relations defined on the finite collection of cases. When the ranks are derived by the feature-based description of the cases they are referred as *r*anking features.
– rank - The rank represents the target preference structure and it is the main output of the whole process. The rank produced in output could be an approximation of the exact rank.

Starting from this generic architecture, we can design at least four different models of the iterative process that enable the ranking of a set of cases through the interactive acquisition of pairwise preference: the basic single-user loop model, the iterated single-user loop model, the basic multi-user loop model and the self multi-user loop model.

The basic single-user loop model is depicted in Figure 2 and it is based on an iteration of three steps in the following sequence:

1. Pair sampling
   An automated procedure selects from the case base a pair of cases whose relative preference is unknown according to a predefined acquisition policy.
2. Preference acquisition
   Given a pair of alternative cases the end user chooses which one is to be preferred with respect to the ultimate goal.
3. Rank learning
   Given a partial elicitation of the user preferences a learning algorithm produces an approximation of the unknown preferences and then a correspondent ranking of the whole set of cases is derived.

If the result of the learning step is considered enough accurate or the end user has been overloaded the iteration halts and the latest approximated rank is given as output; otherwise an other cycle of the loop is carried on. It is to be noticed that the first and the third steps are automated while the second step is in charge to the end user.

The basic single-user loop model is characterized by two elements: first only one user takes part to the elicitation process, second it is assumed that the preference elicitation is monotonic (i.e. the user doesn't see a pair twice).

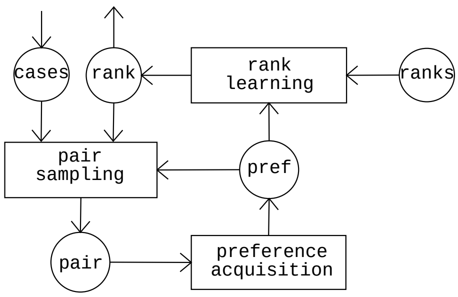


Fig. 2. Basic single-user loop: monotonic elicitation process.

In the iterated single-user loop model there is the assumption that the preference acquisition is context sensitive or better is history dependent. The choice between two alternative cases depends on what has been previously elicited. Because the opinion of the users may change and they should have the opportunity to revise their preferences presenting to them the same pair of alternative cases.

The basic multi-user loop and the self multi-user loop models refer the scenario where many users take part to the elicitation process in order to take into account multiple points of view. More specifically the fourth model introduces a further restrictive condition: the learning step can't take advantage of the predefined order relations over the set of cases.

In the following we will restrict our attention to the basic single-user loop model.

5 The Learning Step

In this section we will focus our attention on the approximation of a preference structure. To explain how a boosting approach can be exploited for this purpose we arrange an abstraction of the problem definition.

The basic elements we have seen until now are the case base, the ranking features, the target rank and the elicited pairwise preferences. The case base is defined as a finite set C of cases. The ranking features $F = (f_1, \ldots, f_N)$ are defined as a finite set of N features that describe the case, where $f_i : C \to \overline{\mathbb{R}}$ ($\overline{\mathbb{R}} = \mathbb{R} \cup \{\perp\}$) and the interpretation of the inequality $f_i(c_0) > f_i(c_1)$ means that c_0 is ranked above c_1 by f_i and $f_i(c) = \perp$ if c is unranked by f_i. The target rank is defined as a feedback function $\Phi : C \times C \to \mathbb{R}$ whose interpretation is that $\Phi(c_0, c_1)$ represents how important it is that c_1 be ranked above c_0; positive values means that c_1 should be ranked above c_0; $\Phi(c_0, c_1) = 0$ indicates that there is no preference between c_0 and c_1 (we assume $\Phi(c, c) = 0$ and $\Phi(c_0, c_1) = -\Phi(c_1, c_0)$ for all $c_0, c_1 \in C$). Finally the set of pairwise elicited preferences is defined as $C_\Phi = \{c \in C | \exists c' \in C : \Phi(c, c') \neq 0\}$ that represents the feedback of the user or the finite support of Φ.

The goal of the learning step is to produce a ranking of all cases in C, including those not ranked by the f_i, represented in the form of a function $H : C \to \mathbb{R}$ with a similar interpretation to that of the ranking features (c_1 is ranked higher than c_0 by H if $H(c_1) > H(c_0)$).

The methodology we have choosen is boosting, a method to produce highly accurate prediction rules by combining many weak rules which may be moderately accurate; here we refer to the boosting algorithm introduced in [10]. In the current setting the objective is a learning algorithm that will produce a function $H : C \to \mathbb{R}$ whose induced ordering of C will approximate the relative orderings encoded by the feedback function Φ using the information from the set of features F.

Before introducing the boosting algorithm we require to define an additional notion of *crucial pair*, a density function $D : C \times C \to \mathbb{R}$ such that

$$D(c_0, c_1) = max(\{0, \Phi(c_0, c_1)\})\gamma$$

setting to 0 all negative entries of Φ; γ is a positive constant chosen so that

$$\sum_{c_0, c_1} D(c_0, c_1) = 1$$

A pair c_0, c_1 is said to be *crucial* if $\Phi(c_0, c_1) > 0$, so that the pair receives non-zero weight under D. The algorithm we used is designed to find an order function H with a small weighted number of crucial-pair misorderings, in other words with a small ranking loss $rloss_D(H)$ defined as:

$$rloss_D(H) = \sum_{c_0, c_1} D(c_0, c_1)[\![H(c_1) \leq H(c_0)]\!] = Pr_{(c_0, c_1) \sim D}[H(c_1) \leq H(c_0)]$$

Algorithm **RankBoost**
Input: initial distribution D over $C \times C$
Output: the final Hypothesis $H(c)$

begin
  Initialize $D_1 = D$;
  For $t = 1, \ldots, T$:
    $h_t = WeakLearner(.)$, where $h_t : C \rightarrow \mathbb{R}$;
    $\alpha_t = ChooseAlpha(.)$, where $\alpha_t \in \mathbb{R}$;
    $D_{t+1}(c_0, c_1) = \frac{D_t(c_0,c_1)}{Z_t} e^{\alpha_t (h_t(c_0) - h_t(c_1))}$;
  return $H(c) = \sum_{t=1}^{T} \alpha_t h_t(c)$;
end.

Fig. 3. RankBoost algorithm.

where $[\![H(c_1) \leq H(c_0)]\!] = 1$ if $H(c_1) \leq H(c_0)$ is true, 0 otherwise.

The algorithm *RankBoost* in Figure 3 operates in rounds. The basic round is based on two procedures called *WeakLearner* and *ChooseAlpha* that are invoked to produce respectively a *weak hypothesis* h and the value for the parameter α. *RankBoost* maintains a distribution D_t over $C \times C$ that is passed on round t to the weak learner. The algorithm chooses D_t to emphasize different parts of the training data. A high weight assigned to a pair of cases indicates a great importance and the weak learner has to order that pair correctly.

Weak hypotheses have the form $h_t : C \rightarrow \mathbb{R}$. The boosting algorithm uses the weak hypothesis h to update the distribution as shown in the algorithm. Suppose that c_0, c_1 is a crucial pair so that we want c_1 to be ranked higher than c_0 (in all other cases D_t is 0). Assuming that the parameter $\alpha_t > 0$ this rule decreases the weight $D_t(c_0, c_1)$ if h_t gives a correct ranking ($h_t(c_1) > h_t(c_0)$) and increases the weight otherwise. Thus, D_t tends to concentrate on the pairs whose relative ranking is hardest to determine.
The final hypothesis H is the α_t weighted sum of the weak hypothesis h_t.

During the development of the system we have implemented the weak learner proposed in [10] that takes as input the ranking features $f_1, \ldots, f_N$ and the distribution D, and as output the learner h. Other possible weak learner schemas can be designed, for example the nearest neighbour learners based on the similarity metrics proposed in [12]; nevertheless combining nearest neighbour learners with a boosting schema produces classifiers more sensitive to the variance of the data.

6 The Pair Sampling Step

If the learning step illustrated in the previous section plays a key role in producing an accurate approximation of the target rank, an other task remains to be tackled: to reduce the elicitation effort of the end users. The purpose of the acquisition policy is to sample the pairs whose relative preferences is unknown. Of course if the number of elicited preferences increases, the approximation error tends to decrease. When the acquisition process is exhaustive the approximation error is zero because all the information is elicited. The challenge is to find a kind of minimum set that maximize the accuracy of the estimated rank. Although the model enables a one-shot pairs sample, the basic intuition is that an incremental strategy could be much more effective because after each learning step a new awareness on what is difficult to approximate is acquired. Therefore an acquisition policy can take advantage focusing the sampling on the pairs where the estimated preferences are not enough accurate [5].

As we have seen in the previous section the boosting technique produces, as a by-product of the learning algorithm, a distribution of weights that emphasizes the pair of cases for which it was harder to approximate the right preference. The rationality of acquisition policy combines this information with an other source of knowledge: the notion of problem complexity [9]. The complexity of the approximation problem can be characterized by the relationship between the target rank and the order relation used as the base of the inductive process. Four types of categories can be defined with increasing degree of complexity: *isotonic*, *antitonic*, *non-monotonic* and *non-projective*.

Although many acquisition policies can be designed all of them share the same structure based on two components: a bootstrap component and a feedback component. The former samples a set of pairs without any hints because at the beginning no information are avaliable (remind that the sampling step is the first of the process), the latter implements the rationality that we sketched above.

In the experiments illustrated in the following section we have adopted a "bipartite" strategy for bootstrapping the pair sampling. The hypothesis is to maximize the minimum coverage of the all cases: the pairs are generated without selecting the same case twice.

The next two sections are devoted to the experimentations of the model looking first at the evaluation setup and then at the empirical results.

7 Evaluation Setup

The evaluation of the case-based preference elicitation model presented until now has been organized distinguishing between on-line and off-line experiments. The off-line experiments have been designed to assess the properties of the process and of the proposed solutions. The on-line experiments have been performed as a case study on a real world problem and with real users in the domain of the civil defence (see Section 9 for the details). As depicted in the Figure 4 all the experiments have tested the same three steps loop using the same acquisition policy and the same learning algorithm. Only the preference acquisition was different: in the on-line experiments the real users have been directly involved while in the off-line experiments they have been simulated.

For the off-line experiments many artificial datasets have been synthesized producing case bases with a number of instances ranging from ten to hundred. For each case base an arbitrary preference function Φ has been associated as the target rank and in a similar way the ranking features f_i have been derived from the case description. The Φ function has been used both to simulate the real user in the preference acquisition step and to assess the accuracy of the approximated rank function H.

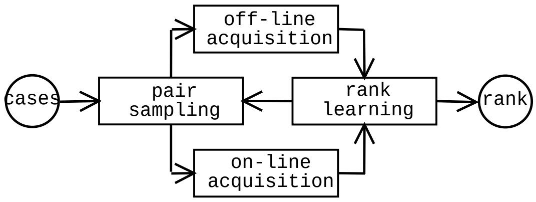


Fig. 4. Experimental Loops: the on-line and off-line elicitation.

The assessment of the accuracy has been performed using the notion of disagreement [7] between the two preference functions Φ and H: i.e. the percentage of approximated pair relations not aligned with the correspondent pair relations defined by the target rank. Of course such a measure depends on the total amount of elicited pairwise preferences. It is formulated as percentage of the total amount of pairs without taking into account the symmetrical pair relations.

To simulate more difficult problems different target functions Φ have been configured on the same case base changing their relationships with the ranking features f_i, whose definition is strongly related to the specific case base. Problems with different complexity degree have been generated using the four categories mentioned above.

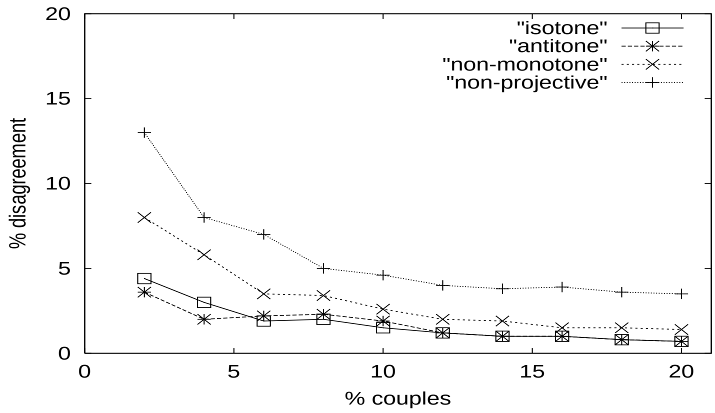
8 Experimental results

We performed many trials to assess the properties of the proposed model and overall to provide empirical evidence that it could be effective.

The aim of the first set of experiments was to assess the claim that a target order relation can approximated taking advantage of a set of orders defined on the same collection of cases, e.g. the ranking features. The results are illustrated on the left side of the Figure 5 where on the y axis is illustrated the disagreement measure and on the x axis the elicitation effort. The elicitation effort is formulated as the percentage of the $|C|(|C|-1)/2$ total amount of relations; the figure doesn't provide the detail of the incremental dynamics of the acquisition policy. The plots refer to the behaviour of the model on a case base of 50 instances and a case description of 20 features. The different curves correspond to different types of problem complexity without changing the case base size. The values represent the average performance over 50 trials. It is worthwhile to notice that, although it is not reported on the plots, the performance of the model is invariant respects with the dimension of the case base and the number of ranking features.

There is a lower bound for the elicitation effort because the end users are constrained to see all the cases at least once, i.e. $|C|/2$ pairs of cases. Such a threshold is around the 2%: the leftmost point of the plots. Even though the performance is already good with the 5% of pairs, the performance of the method becomes stable around 10%, i.e. $2|C|$ pairs of cases. We can conclude that elicitating a number of preferences double respects with the number of cases it possible to lower the approximation error less than 5%. We claim that improving the acquisition policy could lower the elicitation effort achieving a number of pairs equal to the size of the case base.

The second plot of Figure 5 illustrates the results of the experiments related to the comparison between the boolean and the rated pairwise preferences. In these experiments we modified the configuration of the preference function Φ extending the values range from $\{-1, 1\}$ to $[-9, 9]$. The two curves show how the two different ways to elicit preferences don't differ with respect to the disagreement measure; this similarity increases with the cardinality of the case base, so we can conclude that the cognitive overload required by a user to provide a rated preference relation doesn't reduce the approximation error.



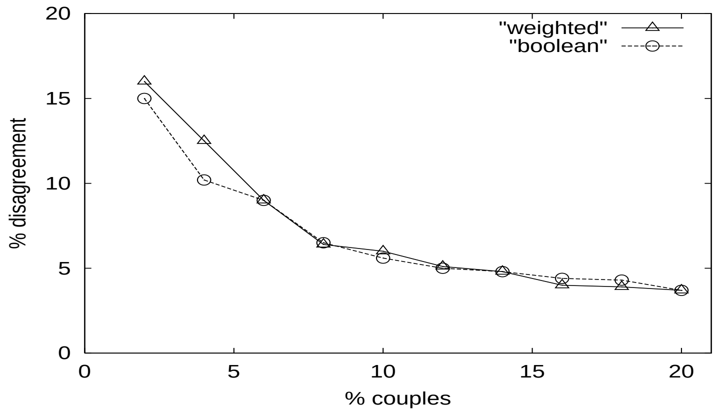


Fig. 5. Evaluation tests: on the y axis is shown the percentage of the wrong pairwise order relations with respect to the target ones: i.e. the disagreement between the approximated ranks and the exact rank; on the x axis is shown the amount of relations given to the learning algorithm: i.e. the percentage of pairs presented to the users. The first plot shows the performance of the method on few case bases of 50 instances with incresing difficulties; the second plot shows the comparison between a boolean choice of alternatives and a fined grained elicitation of preferences.

9 The case study of civile defence

After the good results achieved in the laboratory, there was the challenge of confirming this success also with a test on the field. We chose to arrange a real world experimentation in collaboration with the Geological Department of Trentino, an Italian northern region.

The expert selected the following problem: ranking a set of areas that a preliminary and rough analysis has recognized as risky from the hydrogeological point of view. The ultimate goal was to implicitely elicit a rationality for planning further studies of deeper analysis and the related emergency plans. In particular this process of assessment was required to take into account the systemic vulnerability, i.e. the capability of the social organizations to deal with unexpected events.

The starting point was a collection of 30 areas, mainly landslides, divided into two categories of risk (the source database is “Aree Vulnerate Italiane” developed by a project led by the National Research Council of Italy). From this perspective the goal was to confirm this partition of the areas and at the same time to achieve a total order of the areas of the two sets. An area was represented by a detailed map and by a record with a quite rich description including the geographical positioning, the administrative information, the risk analysis (with data like the landslide volume, land chemical composition, the site slope, the involved infrastructures, the threatened people, ...). Part of these features have been used to induce the order relation functions.

Experimental sessions have been performed with three experts of the Geological Department acting as a single decision maker. The typical session was characterized by the direct interaction of the users (not mediated by a technician) with the web-based interface developed to deploy the implementation of the model. Although the experimentation has not been modified with respects to the definition given in the Section 4, the two steps of sampling and approximation were transparent to the users.

At the beginning of the experimental session a set of pairs of alternative cases was presented to the users that were free to choose which couple of cases to inspect. At this stage the team of expert negotiated the order relation between the two alternative cases, eventually annotating the single case: sometimes extending the case description with new features, sometimes modifying the original values of the case description.

Once the user has elicited the preferences for all the proposed pairs of cases there are two choices: to ask for a new set of pairs or to ask for a synthesis of the approximated rank. Our experts carried on three cycles of the loop with a set of 15 pairs for each step.

At the end of the elicitation phase the final rank was synthesized and the coverage [8] of the induced order relation has been tested with the team of expert. The disagreement on the coverage of the order relation was 4%, a result obtained presenting only the 10% of pairs to the experts. These achievements, accomplished on the field in a real world setting, agree with the promising results of the previous off-line experiments.

10 Conclusions and Future work

In this paper we presented a case-based model for the mixed initiative elicitation process that combines a step of a pairwise case selection and a step of a pairwise preference approximation. We illustrated how a case-based approach can be sustainable also with a large number of cases whether learning phase is combined with the end user elicitation. Moreover we have provided the experimental evidence that a rankboost algorithm is effective to achieve a good trade-off between the accuracy of the final rank and the elicitation effort. A meaningful result is concerned with the boolean preference: we have showed that when the number of cases increases the rated preference and the related cognitive overload become useless.

For lack of the space, we have not included a further interesting result on the characterization of the problem complexity. Four categories of increasing complexity have been introduced discriminating with respect to the relationship between the target rank and the order relations derived from the supporting funtions. A paper devoted to this aspect is in preparation.

The paper omitted illustrating the design detail of the acquisition policy. Up to now we refer to a baseline that has allowed achieving the results above. We are currently working on the enhancement of this baseline to obtain a much more effective acquisition policy that would enable a better exploitation of the available information: the by product of the boosting step and the understanding of the problem complexity.

Last but not least, after these preliminary promising results, a new challenge is open: to extend this approach dealing with the multiple viewpoints or the multiagents architecture, according to a decision theory or an artificial intellingence perspective respectively.

Acknowledgments

We would like to thank the Geological Department of Provincia Autonoma di Trento and in particular Silvana Arnoldi and Roberto Cavazzana for their invaluable support to accomplish the case study.

References

1. A. Aamodt and E. Plaza. Case-based reasoning: Foundational issues, methodological variations, adn system approaches. *AI Communications*, 7(1):39–59, 1994.
2. E. Beinat and P. Nijkamp. *Multicriteria Evaluation in Land-Use Management: Methodologies and Case Studies*. Kluwer, 1998.
3. Jim Blythe. Visual Exploration and Incremental Utitlity Elicitation. In *Proceedings of AAAI/IAAI 2002*, pages 526–532, Edmonton, Alberta, CA, July 2002.
4. K. Bradley and B. Smyth. Personalized information ordering: A case study in online recruitment. In *Proceedings of the Twenty-second SGAI International Conference on Knowledge Based Systems and Applied Artificial Intelligence*, Cambridge, UK, 2002.
5. B. Caprile, C. Furlanello, and S. Merler. Highlighting hard patterns via Adaboost weights evolution. In J. Kittler and F. Roli, editors, *Multiple Classifier Systems, Lecture Notes in Computer Science 2364*, pages 72–80. Springer, 2002.
6. A. Carrara. Potentials and pitfalls of gis technology in assessing natural hazards. *Earth Surface Processes and Landforms*, 16:427–445, 1991.
7. William W. Cohen, Robert E. Schapire, and Yoram Singer. Learning to Order Things. In Michael I. Jordan, Michael J. Kearns, and Sara A. Solla, editors, *Advances in Neural Information Processing Systems*, volume 10. The MIT Press, 1998.

8. B.A. Davey and H.A.Priestley. *Introduction to lattices and order*. Cambridge University Press, 1990.
9. G. Devetag and M. Warglien. Representing others' preferences in mixed motive games: Was schelling right? *to appear*, 2002.
10. Y. Freund, R. Iyer, R. Schapire, and Y. Singer. An Efficient Boosting Algorithm for Combining Preferences. In *Proceedings 15th International Conference on Machine Learning*, 1998.
11. Y. Freund and R. Schapire. A Short Introduction to Boosting, 1999.
12. Vu Ha and Peter Haddawy. Toward case-based preference elicitation: Similarity measures on preference structures. In Gregory F. Cooper and Serafín Moral, editors, *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence (UAI-98)*, pages 193–201, San Francisco, July 24–26 1998. Morgan Kaufmann.
13. McGinty L. and Smyth B. Comparison-based recommendation. In Springer, editor, *Proceedings of the European Conference on Case-Based Reasoning*, Aberdeen, Scotland, 2002.
14. G. Linden, S. Hanks, and N. Lesh. Interactive assessment of user preference models. In *User Modeling*, 1994.
15. Thomas L. Saaty. Fundamentals of the analytic network process. In *Proceedings of International Symposium on Analytical Hierarchy Process*, 1999.
16. Thomas L. Saaty and Luis G. Vargas. *Decision Making in Economic, Political, Social and Technological Environments With the Analytic Hierarchy Process*. RWS Publications, 1994.
17. Thomas L. Saaty and Luis G. Vargas. *Models, Methods, Concepts & Applications of the Analytic Hierarchy Process*. Kluwer Academic, 2000.
18. J. von Neumann and O. Morgestern. *Theory of Games and Economic Behaviour*. Princeton University Press, 1994.