Case-Based Preference Elicitation (Preliminary Report)

Vu Ha Tri Le Peter Haddawy
Decision Systems and Artificial Intelligence Lab
Dept. of EE&CS
University of Wisconsin-Milwaukee
Milwaukee, WI 53201
{vu,lvtri,haddawy}@cs.uwm.edu

Abstract

While decision theory provides an appealing normative framework for representing rich preference structures, eliciting utility or value functions typically incurs a large cost. For many applications involving interactive systems this overhead precludes the use of formal decision-theoretic models of preference. Instead of performing elicitation in a vacuum, it would be useful if we could augment directly elicited preferences with some appropriate default information. In this paper we propose a case-based approach to alleviating the preference elicitation bottleneck. Assuming the existence of a population of users from whom we have elicited complete or incomplete preference models, we propose eliciting the preference model of a new user interactively and incrementally, using the closest existing preference models as potential defaults. A notion of closeness demands a measure of distance among preference orders, which is formally defined and investigated.

Introduction

We are interested in the problem of building interactive systems for which task objectives will be communicated in the form of user preferences. Utility theory, the branch of decision theory that deals with representation of preferences, provides a rich normative framework that is capable of capturing such aspects of preferences as tradeoffs among objectives and attitudes toward risk. But the task of eliciting a utility function is typically time consuming and tedious. For many applications involving interactive systems such overhead can preclude the use of utility theory since the cost in time and effort may be too large relative to the value of the sought after solution or time may simply be limited. For example, we would be very unhappy if a system designed to help us select a video to watch required as much time as the entire length of the film just to elicit our preferences.

To reduce elicitation overhead, practitioners typically make use of assumptions (e.g. additive independence) that simplify the elicitation task by allowing a high-dimensional utility function to be decomposed into a simple combination of lower-dimensional sub-utility functions. But even under such assumptions, elicitation of a complete utility function can still be too time consuming and, furthermore, the assumptions preclude representation of many kinds of interesting and common preferences.

An alternative approach to alleviating the elicitation bottleneck is to work with partially elicited preferences to produce one or more candidate solutions, avoiding incuring the entire elicitation cost up front. The candidate solutions are then used to elicit further preference information from the user, in an interactive and incremental manner.

In previous work (Ha & Haddawy 1997), we investigated an approach in which we first elicit partial preference information to produce a set of candidate solutions. We then use the set of candidate solutions to identify the additional preference information to elicit that would likely be most helpful in narrowing down this set. But in order to be able to make useful inferences after eliciting only some preference information, we were forced to assume additive independence of the underlying utility function and to assume that the sub-utility functions are known.

In contrast, Linden, etal. (Linden, Hanks, & Lesh 1997) supplement elicited preferences with default information to obtain a complete utility function, which is to be continually adjusted based on the user's feedback. This default preference information represents preferences that are assumed applicable to all users, such as preferences for reduced cost. The user is presented with the optimal solution according to the constructed utility function, together with some extreme points in the solution space (e.g., cheapest and shortest-time flights). The utility function is then modified based on the user's critiques. They too assume an additive utility model in order to facilitate the incorporation of user's feedback.

In this paper we also take the approach of supplementing elicited preferences with default information. But rather than applying one default uniformly to all users, we use elicited preferences to select an appropriate default from a set of defaults. Specifically, we investigate a case-based approach to providing such default information. The idea is based on the observation that people tend to form clusters according to their preferences or tastes, an observation that has been analyzed in a large amount of literature in the area of market segmentation. We envision our system to maintain a population of users with their preferences partially or completely specified in a given domain. When encountering a new user A, the system elicits some preference information from A and then determines which user in the population has the preference structure that is closest to the preference structure of A. The preference structure of that user will be set as an initial default representation of A's preferences. In contrast with the previously discussed work of (Ha & Haddawy 1997) and (Linden, Hanks, & Lesh 1997), we do not make any restrictive assumptions concerning the form of the underlying utility functions.

Realization of this approach to elicitation minimally requires that we have a measure of the similarity between two preference structures. In this paper we report on progress toward defining a distance measure on preference structures. The distance is defined to be the probability that two randomly chosen decision alternatives are ranked differently by the two preference structures. We investigate the properties of this distance measure and show how it can be computed in various situations.

The case-based approach we are advocating was inspired by the work on collaborative filtering (Resnick *et al.* 1994; Konstan *et al.* 1997), in which the filtering system predicts how interesting a user will find items he has not seen based on the ratings that other users give to items. Each user in a population rates various alternatives, e.g. newsgroup postings or movies, according to a numeric scale. The system then correlates the ratings in order to determine which users' ratings are most similar to each other. Finally, it predicts how well users will like new articles based on ratings from similar users. The work on collaborative filtering is not cast in the framework of decision theory and no theoretical framework or justification are provided for the similarity measures used. The work that we present here can be viewed as an attempt to provide a formal basis for some of the work in this area.

A Review of Utility Theory

We start out with a brief review of utility theory for decision making. The process of making decisions is generally modeled as the identification of the optimal alternative(s) from a set M of alternatives, using a weak order $\prec$, i.e. an asymmetric ($a \prec b \Rightarrow b \not\prec a$), negatively transitive ($a \not\prec b, b \not\prec c \Rightarrow a \not\prec c$) binary relation on the set of alternatives. We will call this relation the *preference order* of the decision maker: $a \prec b$ indicates that alternative b is prefered to alternative a. We use the notation $a \sim b$ to denote indifference, i.e. $a \not\prec b$ and $b \not\prec a$, and the notation $a \preceq b$ to denote that b is prefered to a or a and b are indifferent.

An important technique that is often used in association with preference orders is the use of *consistent functions* that capture preference orders.

Definition 1 *A real-valued function $f : M \to R$ is said to be* consistent *with a preference order $\prec$ on M if for all $a, b \in M$, $a \prec b \Leftrightarrow f(a) < f(b)$ and $a \sim b \Leftrightarrow f(a) = f(b)$.*

When a decision problem involves no uncertainty about the alternatives, we will call the alternatives *outcomes*, and denote the set of outcomes by Ω. For ease of technical exposition, we assume that Ω is finite and $\Omega = \{s_1, \ldots, s_n\}$. It can be proven (Kreps 1988) that there is a function u, called *value function*, that is consistent with $\prec$. We sometimes write $\prec$ as $\prec_u$ to emphasize this relationship.

As an example, suppose the outcomes of a decision problem are characterized by two binary attributes "Health" (H or $\bar{H}$) and "Wealth" (W or $\bar{W}$): $\Omega = \{HW, H\bar{W}, \bar{H}W, \bar{H}\bar{W}\}$, and that $\prec$ is a preference order over these outcomes such that $\bar{H}\bar{W} \prec \bar{H}W \prec H\bar{W} \prec HW$. A value function that is consistent with this preference order is $v : \Omega \to R$ where $v(\bar{H}\bar{W}) = 0$, $v(\bar{H}W) = .2$, $v(H\bar{W}) = .5$, and $v(HW) = 1$.

When the alternatives are uncertain, they are usually modeled by probability distributions over outcomes and are called *prospects*. For example, an alternative can be a probability distribution p over Ω such that $p(\bar{H}\bar{W}) = .1$, $p(H\bar{W}) = .5$, and $p(HW) = .4$. We denote the set of all probability distributions over Ω by $\mathcal{S}$. The central result of utility theory is a representation theorem that identifies a set of conditions guaranteeing the existence of a function consistent with the preferences of a decision maker (von Neumann & Morgenstern 1944; Savage 1954). The theorem states that if the preference order of a decision maker satisfies a few "rational" properties, then there exists a real-valued function, called *utility function* $u : \Omega \to R$, over outcomes such that $p \prec q \Leftrightarrow E_p(u) < E_q(u)$, and $p \sim q \Leftrightarrow E_p(u) = E_q(u)$. Here $E_p(u)$ denotes the expected value of function u with respect to the distribution p [1]. It is often convenient to extend u, by means of expectation, to a function $u : \mathcal{S} \to R$ that maps a prospect $p \in \mathcal{S}$ to $E_p(u)$. This function is clearly consistent with the preference order $\prec$, and we sometimes write $\prec$ as $\prec_u$ to emphasize this relationship. Utility functions are known to be unique up to a positive linear transformation.

For the sake of simplicity, we shall use the term "utility function" to talk about both value and utility functions when no distinction between them is needed. Two utility functions that induce identical orders are said to be *strategically equivalent*. Otherwise, they are said to be *strategically different*. Note that strategic equivalence is an equivalence relation.

Distance Among Preference Orders

Suppose that there are two users with corresponding preference orders $\prec_1$ and $\prec_2$, which are weak orders on the set of alternatives. We are interested in defining the distance between these two preference orders. Intuitively, this distance should be proportional to the chance that a uniformly randomly chosen pair (a, b) of alternatives will cause a conflict between the two users, i.e, the two users will rank a and b differently. We use an indicator function, defined below, to capture conflicts; a conflict occurs when the indicator function takes on the value 1:

$$c_{\prec_1,\prec_2}(a,b) = \begin{cases} 1 & \text{if } (a \preceq_1 b \wedge a \succ_2 b) \vee \\ & \quad (a \succ_1 b \wedge a \preceq_2 b) \vee \\ & \quad (b \preceq_1 a \wedge b \succ_2 a) \vee \\ & \quad (b \succ_1 a \wedge b \preceq_2 a) \\ 0 & \text{otherwise} \end{cases}$$

The Certainty Case

In the case of certainty, the distance between two orders $(\Omega, \prec_1)$ and $(\Omega, \prec_2)$ is defined as:

$$\begin{aligned} d(\prec_1, \prec_2) &= \Pr(\prec_1 \text{ and } \prec_2 \text{ rank } s_j \text{ and } s_k \text{ differently}) \\ &= \frac{2}{n(n-1)} \sum_{1 \le j < k \le n} c_{\prec_1,\prec_2}(s_j, s_k). \end{aligned}$$

Example 1 *Suppose that there are three kinds of night entertainment in Milwaukee: going to watch a basketball game (B), a movie (M), and going to a blues pub (P). If Xaviera's preference order $\prec_X$ is such that $B \prec_X M \prec_X P$ and Yvette's preference order $\prec_Y$ is such that $M \prec_Y P \prec_Y B$, then the distance between their preferences is $d(\prec_X, \prec_Y) = 2/3$ since out of three pairs of alternatives, they agree only on the relative ranking of the pair $\{M, P\}$ and disagree on those of the other two.*

We note that it is also common to define the distance between two orders over Ω by using the *rank correlation coefficient* [2]. Specifically, suppose that $r_{ij}, i = 1, 2, j = 1, \ldots, n$ is the ranking (an integer number between 1 and n) that $\prec_i$ assigns to outcome s_j. Then the distance $\delta(\prec_1, \prec_2)$ can be defined as:

$$\delta(\prec_1, \prec_2) = \sqrt{\frac{3}{n^3} \sum_{i=1}^{n} (r_{1j} - r_{2j})^2}.$$

Function δ is also a metric with range $[0, 1]$. In general, δ agrees with d in the following sense: $d(\prec_1, \prec_2) < d(\prec_1, \prec_3)$ implies $\delta(\prec_1, \prec_2) < \delta(\prec_1, \prec_3)$ and vice versa. However there are abnormal cases where they disagree.

The Uncertainty Case

In the case of uncertainty, the distance between two orders $(\mathcal{S}, \prec_1)$, $(\mathcal{S}, \prec_2)$ is defined as:

$$\begin{aligned} d(\prec_1, \prec_2) &= \Pr(\prec_1 \text{ and } \prec_2 \text{ rank } p \text{ and } q \text{ differently}) \\ &= \int_{\mathcal{S}} \int_{\mathcal{S}} c_{\prec_1,\prec_2}(p, q) dp dq \end{aligned}$$

When the preferences $\prec_i$ are captured by utility functions u_i, we sometimes write $d(u_1, u_2)$ instead of $d(\prec_1, \prec_2)$. It is not hard to see that two utility functions that are linear transformations of each other (i.e. there

[1] In the case when the set of outcomes is finite, as assumed in this paper, $E_p(u)$ can be viewed as the inner product $E_p(u) = p \circ u$ of the probability vector p and the utility vector u.

[2] For example, the rank correlation coefficient is used in (Ha & Haddawy 1997) to measure the disagreement of two attribute rankings.

are $\alpha, \beta \in R, \alpha \neq 0$ such that $u_2 = \alpha u_1 + \beta$, a condition that is always satisfied when there are exactly two outcomes) have distance of 0 or 1 depending on whether α is positive or negative.

Example 2 *Suppose that Xaviera and Yvette are involved in a decision problem involving lotteries over three possible outcomes a, b, and c. Furthermore suppose that Xaviera's preferences and attitude toward risks are captured by a utility function u_X such that $u_X(a) = 0, u_X(b) = 1, u_X(c) = 2$, and Yvette's preferences and attitude toward risks are captured by a utility function u_Y such that $u_Y(a) = 0, u_Y(b) = 2, u_Y(c) = 3$, then the distance between their preferences and attitudes toward risks is:*

$$\begin{aligned} & d(\prec_1, \prec_2) \\ & = \int_S \int_S c_{\prec_{u_X}, \prec_{u_Y}}(p,q) dp dq \\ & = \int_S \int_S Neg[(u_X \circ (p-q))(u_Y \circ (p-q)] dp dq \\ & = 1/9. \end{aligned}$$

Here p and q run over the probability simplex $\mathcal{S}$, which is the equilateral triangle in R^3 with vertices $(0,0,1), (0,1,0), (1,0,0)$; $u_X = (0,1,2)$, $u_Y = (0,2,3)$; "$\circ$" denotes the inner product of two vectors, and Neg is the function that returns 1 if its argument is negative and 0 otherwise.

Note that Xaviera and Yvette agree on the ranking of certain outcomes, but have different attitudes toward risky decisions. Now suppose that Zelda's utility function is $u_Z = (0,2,1)$, which implies that, unlike Xaviera and Yvette, she prefers c with certainty to b with certainty. Calculations show that $d(u_X, u_Z) = 1/3 > 1/9$, which is what we would expect: Yvette's preferences are more similar to Xaviera's than Zelda's are.

Some Properties of the Distance Function

Proposition 1 *The distance function (for both certainty and uncertainty cases) is a metric on the space of all preference orders of alternatives, i.e., it satisfies the following three metric properties for all preference orders $\prec_i, i = 1,2,3$:*

(i) **Reflexivity.** $d(\prec_1, \prec_2) \geq 0$,

"=" iff $\prec_1$ and $\prec_2$ are identical.

(ii) **Symmetry.** $d(\prec_1, \prec_2) = d(\prec_2, \prec_1)$.

(iii) **Triangle Inequality.** $d(\prec_1, \prec_3) \leq d(\prec_1, \prec_2) + d(\prec_2, \prec_3)$.

Proof: It is obvious that the distance function only takes nonnegative values, and the distance between two identical orders is zero. In the case of certainty, it can also be proven that zero distance implies two identical orders. For the case of uncertainty, let u_1 and u_2 be two utility functions that capture $\prec_1$ and $\prec_2$, respectively. Note that the subset of $\mathcal{S} \times \mathcal{S}$ that consists of pairs of prospects that cause conflict between $\prec_1$ and $\prec_2$ is a union of two convex cones, and it can be shown that this union has empty volume if and only if the two utility functions u_1 and u_2 are positive linear transformations of each other. The first property is then established. The symmetry of the distance function trivially follows from the symmetry of the conflict function. Finally, to prove the triangle inequality, we note that for all preference orders $\prec_i, i = 1,2,3$ and alternatives a, b, $c_{\prec_1,\prec_3}(a,b) = 1$ implies either $c_{\prec_1,\prec_2}(a,b) = 1$ or $c_{\prec_2,\prec_3}(a,b) = 1$, and that for all events X, Y, $\Pr(X \vee Y) \leq \Pr(X) + \Pr(Y)$. □

Closely related to the distance concept is the *similarity* concept in fuzzy-set theory (Zadeh 1971; Ovchinnikov 1991). A similarity relation s is a binary fuzzy relation defined on a set U that satisfies the following three properties for all $u, v, w \in U$:

(i) **Reflexivity.** $s(u,u) = 1$.

(ii) **Symmetry.** $s(u,v) = s(v,u)$.

(iii) **$*$-Transitivity.** $s(u,v) * s(v,w) \leq s(u,w)$,

where $*$ is a t-norm operation, that is, a commutative, associative, non-decreasing operation on $[0,1]$, with 1 being the neutral element ($1 * x = x * 1 = x, \forall x \in [0,1]$) and 0 being the absorbent element ($0 * x = x * 0 = 0, \forall x \in [0,1]$). Noticeable t-norms are min, product, and Lukasiewicz operation ($l(x,y) = \max\{0, x+y-1\}$).

It is not hard to see that the complement of our distance measure, defined as $s(\prec_1, \prec_2) = 1 - d(\prec_1, \prec_2)$ — the probability that two users with preference orders $\prec_1$ and $\prec_2$ will have the same preference over a randomly chosen pair (a,b) of alternatives — is a fuzzy similarity relation with respect to Lukasievicz t-norm.

Distance Among Partially Specified Preference Orders

While the distance function proposed above provides a similarity measure that could be used to cluster preferences of multiple users, it is not much good for preference elicitation. For the purpose of elicitation we need to be able to compute the distance when at least one of the preference orders is only *partially* specified.

We first clarify what we mean by "partially specified" preference order. Recall that a preference order $\prec$ on the set M of alternatives is a weak order, i.e., an asymmetric, negatively transitive binary relation on M. For a decision maker whose preferences we have completely elicited, given any two alternatives a and b, we know that there are only three possibilities: either she prefers b to a ($a \prec b$), a to b ($b \prec a$), or indifferent between a and b ($a \sim b$). However, when we have little information about the preferences of the decision maker, it might be the case that we can not say anything about her preferences over some two alternatives a and b, meaning that none of the above three possibilities applies. In such cases, we say that the preference between a and b is not specified and denote it by $a \bowtie b$. Thus, a partially specified order can be viewed as a *partially specified*, or *generalized weak order* where for any pair of alternatives (a, b), exactly one of the four relations $\prec$, $\succ$, $\sim$, and $\bowtie$ holds, corresponding to whether b is prefered to a, a is prefered to b, a and b are indifferent, or we do not have information about the decision maker's preferences over a and b.

For simplicity, we will call a partially specified preference order, which is a 3-tuple $(\prec, \sim, \bowtie)$, a *partial preference order* and denote it by just $\prec$. We need a slightly different definition of consistent functions – functions that are consistent with partial preference orders.

Definition 2 *A real-valued function $f : M \to R$ is said to be* consistent *with a partial preference order $\prec$ on M if for all $a, b \in M$, $a \prec b \Rightarrow f(a) < f(b)$ and $a \sim b \Rightarrow f(a) = f(b)$.*

In other words, consistent functions capture all information contained in $\prec$, and they might contain more than that.

We now turn our attention to defining the distance among partially specified preference orders. We consider both the certainty case and the uncertainty case. Suppose the preferences of two users are partially specified by two partial preference orders $\prec_1$ and $\prec_2$ on M. For each partial preference order $\prec_i$, let U_i^* be the set of all utility functions that are consistent with it. Note that since strategic equivalence is an equivalence relation, we can define the set of equivalence classes with respect to this relation. We denote this set by U_i. Furthermore, we can view U_i as a subset of U_i^* that contains strategically different functions, and any member function of U_i^* is strategically equivalent to some member of U_i.

Given this notation, we can define the distance $d(\prec_1, \prec_2)$ in two ways. First, we can define the distance between two partial preference orders to be the probability that a uniformly randomly chosen pair (a, b) of alternatives will cause a conflict between two utility functions u_1, u_2 that are chosen independently and uniformly randomly from U_1 and U_2. In other words, $d(\prec_1, \prec_2)$ is defined as:

$$\Pr(c_{\prec_{u_1}, \prec_{u_2}}(s_j, s_k) = 1 | 1 \le j < k \le n, u_i \in U_i, i = 1, 2)$$

in the case of certainty and

$$\Pr(c_{\prec_{u_1}, \prec_{u_2}}(p, q) = 1 | p, q \in \mathcal{S}, u_i \in U_i, i = 1, 2)$$

in the case of uncertainty.

An alternative approach is to define the distance between $\prec_1$ and $\prec_2$ to be an interval whose endpoints are the minimum and the maximum of $d(u_1, u_2)$, where $u_i \in U_i, i = 1, 2$:

$$\mathbf{d}(\prec_1, \prec_2) = [\min_{u_i \in U_i} d(u_1, u_2) \ \max_{u_i \in U_i} d(u_1, u_2)].$$

An advantage of the first approach is that the distance among partial preference orders remains both as a probability number and a metric on the space of all partial preference orders, a property that makes it seem to be a theoretically appealing extension of the basic distance among complete preference orders.

In contrast, the second approach gives us more flexibility in defining the concept of closeness among partial preference orders. For example, given three partial preference orders $\prec_i, i = 1, 2, 3$, we can take the conservative approach and say that $\prec_1$ is closer to $\prec_2$ than to $\prec_3$ if the upper bound of $\mathbf{d}(\prec_1, \prec_2)$ is less than the lower bound of $\mathbf{d}(\prec_1, \prec_3)$. We can also take other approaches such as the optimistic approach (*minimin*: closer when the lower bound is smaller) and pessimistic approach (*minimax*: closer when the upper bound is smaller).

Computing the Distance Among Partial Preference Orders

In this section, we investigate the problem of computing the distance among partial preference orders in the case of certainty. We consider the first definition of the distance function proposed in the previous section.

We will be talking about users with partial preference orders $\prec_i$ on the set of outcomes $\Omega = \{s_1, \ldots, s_n\}$ and corresponding sets of consistent, strategically different value functions U_i. We assume that each member value function $u_i \in U_i$ is injective, i.e., $u_i(s_j) \neq u_i(s_k), \forall 1 \leq j < k \leq n$. This assumption has a non-sufficient prerequisite that the partial preference orders $\prec_i$ are *strict partial orders*, i.e., there are no alternatives s_j and s_k such that $s_j \sim s_k$. We make this assumption in order to simplify our distance-computing algorithm; removing this assumption will result in slightly different algorithms that will be reported in the longer version of this paper.

Observe that since the outcomes are chosen uniformly randomly from Ω and independently from u_i, we have that the distance $d(\prec_1, \prec_2)$ is equal to

$$\begin{aligned} &= \Pr(c_{\prec_{u_1}, \prec_{u_2}}(s_j, s_k) = 1 | 1 \leq j < k \leq n, u_i \in U_i) \\ &= \frac{2}{n(n-1)} \sum_{1 \leq j < k \leq n} \Pr(c_{\prec_{u_1}, \prec_{u_2}}(s_j, s_k) = 1 | u_i \in U_i). \end{aligned}$$

We thus need to compute $\Pr(c_{\prec_{u_1}, \prec_{u_2}}(s_j, s_k) | u_i \in U_i, i = 1, 2)$ for each $1 \leq j < k \leq n$. When both $\prec_1$ and $\prec_2$ have rankings between s_j and s_k, this probability is either 1 or 0 depending on whether they agree or conflict. So we only need to consider the case when at least one of $\prec_1$ and $\prec_2$ does not have a ranking between s_j and s_k. To this end we represent each partial order $\prec_i$ with a directed acyclic graph (DAG) G_i that has vertices $s_1, \ldots, s_n$ and edges $s_p \leftarrow s_q$ exactly where $s_p \prec_i s_q$. Then each complete order $\prec_{u_i}$ induced by a value function $u_i \in U_i$ can be represented as a *complete directed acyclic graph* that is a complete extension of G_i.

Denote the number of complete extensions of a DAG G by $\mathcal{E}(G)$. In the case when G is not a DAG, i.e., when G has a directed cycle, let $\mathcal{E}(G) = 0$. The total number of pairs of complete extensions of a pair of DAGs (G_1, G_2) is then just $\mathcal{E}(G_1)\mathcal{E}(G_2)$. Since u_1 and u_2 are chosen independently from U_1 and U_2, respectively, the probability of conflict caused by s_j and s_k can be computed as follows. If neither $\prec_1$ nor $\prec_2$ have rankings between s_j and s_k, then

$$\begin{aligned} &\Pr(c_{\prec_{u_1}, \prec_{u_2}}(s_j, s_k) = 1 | u_i \in U_i, i = 1, 2) \\ &= \frac{1}{\mathcal{E}(G_1)\mathcal{E}(G_2)} \times \\ &\quad [\mathcal{E}(G_1 + \{s_j \leftarrow s_k\})\mathcal{E}(G_2 + \{s_j \rightarrow s_k\}) \\ &\quad + \mathcal{E}(G_1 + \{s_j \rightarrow s_k\})\mathcal{E}(G_2 + \{s_j \leftarrow s_k\})] \end{aligned}$$

Otherwise, suppose that, say $s_j \prec_1 s_k$ and $\prec_2$ does not have a ranking between s_j and s_k ($xs_j \bowtie_2 s_k$), then

$$\Pr(c_{\prec_{u_1}, \prec_{u_2}}(s_j, s_k) | u_i \in U_i, i = 1, 2)$$

$$= \frac{\mathcal{E}(G_1)\mathcal{E}(G_2 + \{s_j \rightarrow s_k\})}{\mathcal{E}(G_1)\mathcal{E}(G_2)}.$$

Conclusions and Future Research

In this paper we have only provided the initial theoretical basis for a cased-based approach to preference elicitation. The next step is to implement the approach to see how well it works in practice. In order to do this several additional problems need to be addressed. We need a good way of determining how to sample preferences from a population of users in order to populate our database of cases. If the initial goal of elicitation is to find the case (or cluster) that most closely matches the preferences of our user, then one would like to have an elicitation strategy that would be efficient in discriminating among the available cases. Finally, the preference structure that is retrieved as being the closest match will typically have some conflicts with the directed elicited user preferences. It would be reasonable to modify that retrieved structure so that those conflicts are eliminated.

References

Ha, V., and Haddawy, P. 1997. Problem-focused incremental elicitation of multi-attribute utility models. In *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence*, 215–222.

Konstan, J. A.; Miller, B. N.; Maltz, D.; Herlocker, J. L.; Gordon, L. R.; and Reidl, J. 1997. Applying collaborative filtering to usenet news. *Communications of the ACM* 40(4):77–87.

Kreps, D. 1988. *Notes on the Theory of Choice.* Underground Classics in Economics. Boulder: Westview Press.

Linden, G.; Hanks, S.; and Lesh, N. 1997. Interactive assessment of user preference models: The automated travel assistant. *User Modeling.*

Ovchinnikov, S. V. 1991. Similarity relations, fuzzy partitions, and fuzzy orderings. *Fuzzy Sets and Systems* 40:107–126.

Resnick, P.; Iacovou, N.; Suchak, M.; Bergstrom, P.; and Reidl, J. 1994. Grouplens: An open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 Computer Supported Cooperative Work Conference*, 175–186.

Savage, L. 1954. *The Foundations of Statistics.* New York: John Wiley & Sons. (Second revised edition published 1972).

von Neumann, J., and Morgenstern, O. 1944. *Theory of Games and Economic Behavior.* Princeton Univesity Press.

Zadeh, L. A. 1971. Similarity relations and fuzzy orderings. *Information Sciences* 177–200.