Published in Proceedings of the 2008 AAAI Workshop on Advances in Preference Handling, J. Chomicki, V. Conitzer, U. Junker, and P. Perny, eds., AAAI Technical Report WS-08-09, Menlo Park, California, USA, pp. 128-133

Comparing Preferences Expressed by CP-networks (Extended Abstract)

Andrew W. Wicker and **Jon Doyle**
Department of Computer Science
North Carolina State University
Raleigh, NC 27695-8206
{AWWicker, Jon_Doyle}@ncsu.edu

Abstract

Comparisons of similarity or dissimilarity between systems of preferences over multiple attributes play important roles in interest matching, social networking, collaborative filtering, and personalization. We develop metrics over preferences represented compactly by conditional preference networks (CP-networks) and their variants. Our metrics exhibit intuitive properties and support efficient (polynomial-time) algorithms for computing similarities.

Introduction

Decisions to interact with or accept recommendations from others in social networking environments rely in part on assessments of the degree to which one's interests and tastes are shared with or are different from those of others. We focus on comparing decision-theoretic preferences of agents on the assumption that such preferences can serve as a reasonable proxy for or representation of many pertinent aspects of interests and tastes. This assumption seems reasonable in that the notion of preference formalized in economic decision theory provides quite a bit of expressive power, both in traditional multiattribute utility function approaches and in theories of preference *ceteris paribus* at the focus of much recent research (Wellman & Doyle 1991; Doyle, Shoham, & Wellman 1991).

In previous work (Wicker 2006; Wicker & Doyle 2007), we studied the use of preference-similarity measures developed by Kemeny and Snell 1962 and Bogart 1973 in comparing the interests of agents. That work used these existing metrics to compare preference orders expressed by conditional preference networks, or CP-networks (Boutilier, et al 1999; 2004), which provide a simple graphical representation for one common form of preference expression (as depicted in Figure 2). These metrics, however, have bad computational properties.

In this paper, we extend the results of (Wicker 2006) by constructing metrics on CP-networks that can be computed efficiently and that reflect comparative judgments related to those of the Kemeny, Snell, and Bogart metrics. We continue a focus on CP-networks rather than more general logical languages for specifying preferences, for example (Doyle, Shoham, & Wellman 1991; Kiessling 2004; Chomicki 2003). We reformulate the usual development of CP-networks in functional terms so as to better facilitate the statement and analysis of our metrics and other functions of preference orders. For lack of space here we state our theorems without proofs, which can be found in (Wicker & Doyle 2008).

Preferences

Attributes

We consider preference representations that employ a finite set $\mathcal{A}$ of one or more *attributes* of possibly different types. In all the following, we presume an enumeration of the attributes as $\langle \mathcal{A} \rangle = \langle a_0, a_1, \ldots, a_n \rangle$. We write individual attributes as a or a_i, with subscripts on attributes normally referring to the enumeration ordering. Each attribute a has a set or domain of attribute *values* V_a that it can take, which we here assume is finite. We often write V_i as shorthand for V_{a_i}, and sometimes write a as shorthand for V_a. Distinct attributes $a \neq a'$ can have the same set of values, that is, $V_a = V_{a'}$.

Outcomes

An *outcome* is a partial function $\omega : \mathcal{A} \to \bigcup_{a \in \mathcal{A}} V_a$ such that $\omega(a) \in V_a$ for each attribute a for which $\omega(a)$ is defined. If ω is defined exactly for the attributes in a set A, we say that A is the *domain (of definition)* of ω and write $\mathrm{dom}(\omega) = A$. We write Ω to denote the set of all outcomes, and write Ω_A to denote the set of outcomes with domain of definition A.

Each outcome $\omega \in \Omega_A$ thus corresponds to a total function from A to $\bigcup_{a \in A} V_a$. The set Ω_A is thus isomorphic to $\prod_{a\in A} V_a$, with the factors of the product ordered by ascending attribute enumeration order, and with each outcome corresponding to a tuple of attribute values for attributes in the domain of definition. We sometimes write ω_i to refer to the value $\omega(a_i)$ of a_i in ω. We write Ω_a to abbreviate $\Omega_{\{a\}}$, which is isomorphic to the set of values V_a of a. With these definitions, we have $\Omega_\emptyset = \{\ ()\ \}$ being the set consisting of the unique 0-tuple, $\Omega_{\mathcal{A}}$ being the set of total outcomes defined on all attributes, and $\Omega = \bigcup_{A\subseteq\mathcal{A}} \Omega_A$.

We write $\rho_A(\omega)$ or $\rho(\omega, A)$ to denote the *retraction* of an outcome ω to an outcome over a set of attributes contained in A, with $\rho_A(\omega)$ obtained by discarding the values in ω corresponding to attributes not in A. Formally, we define $\rho_A(\omega)$ to be the outcome such that $\rho_A(\omega)(a)$ is $\omega(a)$ if $a \in A \cap \mathrm{dom}(\omega)$, and is undefined for all other attributes. If the domain of definition of ω is A' and $A \subseteq A'$, then the domain of definition of $\rho_A(\omega)$ will also be A. If $A \not\subseteq A'$, then the domain of definition of $\rho_A(\omega)$ will be a proper subset of A. In particular, if A and A' are disjoint, we define $\rho_A(\omega) = ()$.

We write $\eta_A(\omega)$ or $\eta(\omega, A)$ to denote the *expansion* of an outcome ω to outcomes over a set of attributes containing A. Where the value of $\rho_A(\omega)$ is a single outcome, the value of $\eta_A(\omega)$ is a set of outcomes, namely, all those outcomes with domains of definition including A that would return to ω under retraction. Formally, we define

$$\eta_A(\omega) \stackrel{\text{def}}{=} \{\omega' \in \Omega_{A\cup\mathrm{dom}(\omega)} \mid \rho_{\mathrm{dom}(\omega)}(\omega') = \omega\}.$$

If every outcome in a set S of outcomes has the same domain of definition, we say that S is *homogeneous* and extend the notion of expansion to S in the natural way by defining $\eta_A(S) = \bigcup_{\omega\in S} \eta_A(\omega)$.

To simplify some definitions and computations, we assume that the values of each attribute a are totally ordered by a *reference order* $\sqsubset_a$. We combine the enumeration order of attributes with the reference orders of attribute values to obtain a reference order on outcomes. Specifically, we define the reference order on each set Ω_A to be that obtained as the lexicographic ordering with respect to the attribute enumeration and reference order of the attribute values. A finite set of outcomes consists of a finite number of attributes, each of which has finitely many values. In this case, a reference ordering provides a *reference enumeration* of outcomes. Moreover, for finite sets of binary attributes, we can obtain a simple enumeration of outcomes by interpreting each outcome as the binary representation of an integer, and using the natural ordering of the resulting integers.

Preference orders

Economics formalizes preferences in terms of notions of weak preference, strict preference, and indifference. *Weak preference* refers to a partial preordering of a set of alternatives, that is, a reflexive and transitive binary relation $\succsim$. *Strict preference* consists of the strict partial order obtained as the strict part of a weak preference ordering, that is, an irreflexive, antisymmetric, and transitive binary relation $\succ$ defined so that $x \succ y$ iff $x \succsim y$ and $y \not\succsim x$. *Indifference* consists of the equivalence relation formed by the symmetric portion of a weak preference ordering, defined so that $x \sim y$ iff $x \succsim y$ and $y \succsim x$. Our focus in this paper is on strict preference, so when we speak of orders, we normally will mean strict partial orders.

We write $\mathcal{O}$ to denote the set of all preference orders over homogeneous outcomes, defined as follows. Specifically, we write O_A to mean the set of all strict partial orders over Ω_A, write O_a to mean the set of orderings over the set $\Omega_{\{a\}}$ of values of attribute a, and write $\mathrm{dom}(o)$ to denote the common domain of definition $\mathrm{dom}(\omega)$ of outcomes ordered by o. We then obtain the full set of orders $\mathcal{O} = \bigcup_{A\subseteq\mathcal{A}} O_A$ by combining all the limited sets of orders. We write $\perp_A$ to mean the *empty order* over Ω_A in which no outcome is strictly preferred to any other.

Preference *ceteris paribus*

In the present treatment, we interpret preferences over the values of one attribute as preferences *ceteris paribus*, that is, as expressions of preference for one value over another other things being equal. Formally, we interpret a preference $v \succ v'$ for one value $v \in V_a$ over some other value $v' \in V_a$ as expressing a preference order $[\![v \succ v']\!]$ over full outcomes in $\Omega_{\mathcal{A}}$ so that $\omega \succ \omega'$ in $[\![v \succ v']\!]$ whenever $\omega_i = v$, $\omega'_i = v'$, and $\omega_j = \omega'_j$ for each $j \neq i$.

More generally, for any set of attributes $A \subseteq \mathcal{A}$, we define the ceteris paribus *expansion* $\eta_A(v \succ v')$ of the comparison $v \succ v'$ from the domain V_a to domain Ω_A by

$$\begin{aligned}\eta_A(v \succ v') \stackrel{\text{def}}{=} \{(\omega, \omega') \in \Omega_A \times \Omega_A \mid\ & \\ \rho_{\{a\}}(\omega) = v \wedge\ & \\ \rho_{\{a\}}(\omega') = v' \wedge\ & \\ \rho_{A\setminus\{a\}}(\omega) = \rho_{A\setminus\{a\}}(\omega')\}.&\end{aligned}$$

We clearly have $[\![v \succ v']\!] = \eta_{\mathcal{A}}(v \succ v')$. We extend the order-interpretation notation to write $[\![v \succ v']\!]_A$ to denote the expansion $\eta_A(v \succ v')$, that is, the preference order over Ω_A entailed by the condition $v \succ v'$.

Note that if $a \notin A$, then the expansion $\eta_A(v \succ v') = [\![v \succ v']\!]_A$ consists of the empty order over Ω_A that leaves all outcomes incomparable.

If $\omega \in \Omega$ and v and v' denote values of a, we define the *restricted preference condition* $[\![\omega \Rightarrow v \succ v']\!]_A$ to mean the restriction of the order $[\![v \succ v']\!]_A$ to outcomes subsuming ω, that is,

$$\begin{aligned}[\![\omega \Rightarrow v \succ v']\!]_A \stackrel{\text{def}}{=} [\![\{(\omega', \omega'') \in [\![v \succ v']\!]_A \mid\ & \\ \rho_{\mathrm{dom}(\omega)}(\omega') = \omega \wedge\ & \\ \rho_{\mathrm{dom}(\omega)}(\omega'') = \omega\}]\!]_A.&\end{aligned}$$

We have $[\![\omega \Rightarrow v \succ v']\!]_A = \emptyset$ if $a \in \mathrm{dom}(\omega)$, for then the only possible restricted comparisons are ones that have the same value for a, which cannot be preferred to itself. We also have $[\![\omega \Rightarrow v \succ v']\!]_A = \emptyset$ if A and $\mathrm{dom}(\omega)$ are disjoint.

If S is a set of restricted or unrestricted conditions on preferences, we write $[\![S]\!]_A$ to denote the preference order entailed by the transitive union of the conditions in S, that is,

$$[\![S]\!]_A = \left[\bigcup_{s\in S}[\![s]\!]_A\right]^*,$$

where the notation R^* means the transitive closure of the binary relation R. In particular, if $o \in O_a$ and $a \in A$, we write $[\![o]\!]_A$ to denote the expansion of the order over Ω_a to an order over Ω_A formed as the transitive union

$$[\![o]\!]_A = \left[\bigcup_{(v,v')\in o}[\![v \succ v']\!]_A\right]^*$$

of the set of expansions of all comparisons $v \succ_o v'$.

Reusing the same notation, we generalize the notion of *ceteris paribus* expansion to change of basis of an order to an order over any other set of outcomes. We define the expansion or change of $\omega \succ \omega'$, where $\omega, \omega' \in \Omega_{A'}$, to a comparison between outcomes over a domain A by

$$\begin{aligned}\eta_A(\omega \succ \omega') \stackrel{\text{def}}{=} \{(\omega'', \omega''') \in \Omega_A \times \Omega_A \mid \\ \rho_{A'}(\omega'') = \rho_{A\cap A'}(\omega) \wedge \\ \rho_{A'}(\omega''') = \rho_{A\cap A'}(\omega') \wedge \\ \rho_{A\setminus A'}(\omega'') = \rho_{A\setminus A'}(\omega''')\}.\end{aligned}$$

Clearly, if $A' \subseteq A$, this expands the original comparison to ones over the larger domain of attributes in a way consistent with the preceding definition of *ceteris paribus* expansion.

If $A \subseteq A'$, a change of basis from A' to A has the effect of reducing the comparison in $\Omega_{A'} \times \Omega_{A'}$ to a comparison in $\Omega_A \times \Omega_A$. We say that $o \in O_A$ is the *compaction* or *minimal-basis representation* of $o' \in O_B$ just in case $o' = \eta_B(o)$ and $A \subset A'$ whenever $o'' \in O_{A'}$ and $o' = \eta_{A'}(o'')$. If o is the compaction of o', then we write $o = \kappa(o')$. We observe that no order can be reduced to a smaller basis than that of its compaction in a consistent way, so $[\![o]\!]_A = \emptyset$ if $A \subset \mathrm{dom}(\kappa(o))$.

We define general order expansions of an order $o \in O_{A'}$ to a set of attributes A in terms of the transitive union of their individual comparison expansions by

$$[\![o]\!]_A = \left[\bigcup_{(\omega,\omega')\in o}[\![\omega \succ \omega']\!]_A\right]^*.$$

KSB metric extension

Kemeny and Snell 1962 developed a metric $d_A : O_A \times O_A \to \mathbb{R}$ on finite orders (see (Wicker 2006)) by advancing axioms such metrics must satisfy beyond the ordinary axioms for metrics, and Bogart 1973 provided an enlarged set of axioms extending this metric to strict preference orders. Although defined in terms of a natural matrix representation of orders, their metric has a simpler restatement in terms of the set-theoretic representation of the preference orders. If one regards strict orders o and o' as sets of ordered pairs and writes the symmetric difference of these sets as $o \bigtriangleup o'$, then we also have

$$d_A(o, o') = |o \bigtriangleup o'|. \tag{1}$$

We now extend the KSB metric d_A on the several order sets O_A to a single metric $d : \mathcal{O} \times \mathcal{O} \to \mathbb{R}$ over the full set of orders $\mathcal{O}$. We do this by recasting both orders as orders over their minimal common domain and taking the KSB distance of those minimal representations.

If $\kappa(o) \in O_A$ and $\kappa(o') \in O_{A'}$, then we say that $[\![\kappa(o)]\!]_{A\cup A'}$ and $[\![\kappa(o')]\!]_{A\cup A'}$ are the *minimal common expansions* of the the orders o and o', respectively. That is, we expand the compaction of each order to the minimal attribute set such that the orders are each over the same outcomes. We say that $A \cup A'$ is their *minimal common domain* and denote this by $\mu(o, o')$.

Theorem 1. *For each $o \in \mathcal{O}$ and $A \subseteq \mathcal{A}$, we have $[\![o]\!]_A = [\![\kappa(o)]\!]_A$*

We define an extended KSB metric $d : \mathcal{O} \times \mathcal{O} \to \mathbb{R}$ over the full set of orders $\mathcal{O}$ by finding the KSB distance of orders when translated to their minimal common domain. Formally, if $o, o' \in \mathcal{O}$, then we define the distance $d(o, o')$ by

$$d(o, o') \stackrel{\text{def}}{=} d_{\mu(o,o')}([\![o]\!]_{\mu(o,o')}, [\![o']\!]_{\mu(o,o')}). \tag{2}$$

Clearly, if two orders are both over their minimal common domain, then d agrees with the KSB distance.

Another way to obtain this distance measure is to expand each of the orders under comparison to the full set of attributes, find the KSB distance of the expanded orders, and normalize by the number of complete outcomes over the attribute set by which their minimal common domain has been expanded. That is, (2) is equivalent to

$$d(o, o') = \frac{d_{\mathcal{A}}([\![o]\!]_{\mathcal{A}}, [\![o']\!]_{\mathcal{A}})}{|\Omega_{\mathcal{A}\setminus\mu(o,o')}|}. \tag{3}$$

Conditional preference (CP) networks

Conditional preference networks (CP-networks) were developed to provide a natural and compact representation of simple *ceteris paribus* preferences (Boutilier, et al 1999; 2004), namely preferences over the possible values of individual attributes when these preferences depend on the values taken by other attributes.

Network structure

An *attribute graph* over a set of attributes A is a pair (A, E) such that $E \subseteq A \times A$ represents a set of directed edges from parent to child attribute nodes. It is common to restrict attention to acyclic graphs, but we do not assume that here. We write G_A to denote the set of all attribute graphs over A, and $\mathcal{G} \stackrel{\text{def}}{=} \bigcup_{A \subseteq \mathcal{A}} G_A$ to mean the set of all attribute graphs. We represent an individual graph $g \in \mathcal{G}$ as $(A(g), E(g))$ or just (A, E).

The key aspects of attribute graphs in analyzing CP-networks are the set $Par(a, g) = \{a' \in A(g) \mid (a', a) \in E(g)\}$ of *parents* and the set $Chd(a, g) = \{a' \in A(g) \mid (a, a') \in E(g)\}$ of *children* of a node a in the attribute graph g. When the graph in question is clear from the context, we sometimes write $Par(a)$ instead of $Par(a, g)$. Naturally, if A does not appear in a graph g, we have $Par(a, g) = Chd(a, g) = \emptyset$.

CP-networks represent preferences by means of *CP-tables* associated with each attribute node. In each row of each such table, the last column states an order over the values of the attribute node, and the other columns, if any, state an assignment of values to the parent attributes. Each CP-table has one row for every combination of values for the parent attributes, making the size of the table exponential in the number of parents.

We represent CP-tables as functions from outcomes to orders. Formally, for each $A \subseteq \mathcal{A}$, the set of CP-tables for (parents) A and (child) a is the set of functions $T(A, a) \stackrel{\text{def}}{=} (\Omega_A \to O_a)$. We obtain the set of CP-tables for attribute a by combining these functions over different possible sets of parents into the set of functions $T(a) \stackrel{\text{def}}{=} \bigcup_{A \subseteq \mathcal{A}} T(A, a)$. We define the set of all CP-tables by $\mathcal{T} \stackrel{\text{def}}{=} \bigcup_{a \in \mathcal{A}} T(a)$.

We thus represent *CP-networks* by combining attribute graphs with appropriate CP-tables. Formally, a CP-network $N = (g, t)$ or $N = (A, E, t)$ consists of an attribute graph $g = (A, E)$ together with a function $t : A \to \mathcal{T}$ such that $t(a) \in T(Par(a), a)$. We write $\mathcal{N}_A$ to denote the set of all networks over G_A, and $\mathcal{N}$ to denote the set of all networks over $\mathcal{G}$.

Network semantics

If $N = (A, E, t)$ is a CP-network, $[\![N]\!]_A$ denotes the overt meaning of the network as the induced order over Ω_A. Specifically, if $a \in A$, $\omega \in \Omega_{Par(a,N)}$, and $o = t(a, \omega)$, we interpret the ω row of the CP-table for a in N as making the restricted preference statement $\omega \Rightarrow o$. We then obtain the induced order $[\![N]\!]_A$ by transitive closure of the conjunction

$$\begin{array}{|c|}\hline x : y \succ \bar{y} \\ \hline \bar{x} : \bar{y} \succ y \\ \hline \end{array} \quad \xrightarrow{expansion} \quad \begin{array}{|c|}\hline xz : y \succ \bar{y} \\ \hline x\bar{z} : y \succ \bar{y} \\ \hline \bar{x}z : \bar{y} \succ y \\ \hline \bar{x}\bar{z} : \bar{y} \succ y \\ \hline \end{array}$$

$$t \in T(\{X\}, Y) \qquad\qquad \eta(t, \{X, Z\})$$

Figure 1: CP-table expansion. The table on the left involves only one parent, attribute X. On addition of a parent attribute Z, each row in the original table splits into rows for each value of the new attribute, with each of the split rows indicating the same value ordering as in the unsplit row of the original table.

of these statements, that is,

$$[\![N]\!]_A = \left[\!\!\left[\{ \omega \Rightarrow t(a, \omega) \mid a \in A,\ \omega \in \Omega_{Par(a,N)} \} \right]\!\!\right]_A . \tag{4}$$

Some presentations of CP-networks express this same induced order in algorithmic terms, saying that $\omega \succ \omega'$ just in case there exists a “value-worsening” sequence of outcomes $\omega = \omega^1, \ldots, \omega^k = \omega'$ such that each pair of successive outcomes ω^j and ω^{j+1} differ in exactly one attribute, for which the value in ω^{j+1} is less preferred than the value in ω^j according to the CP-network preference tables.

The preference statements contributing to CP-table meanings in (4) can vary in size, depending on the graph structure of the network. In fact, we can express the network semantics as a set of statements of uniform size and structure simply by considering meanings in terms of the full set of attributes $\mathcal{A}$. That is, we consider the meaning to be the *ceteris paribus* expansion of the overt meanings to the full set of attributes, given by

$$\begin{aligned}[\![\omega \Rightarrow t(a, \omega)]\!] = \big[\!\!\big[\{ \omega' \Rightarrow t(a, \omega) \mid \\ \omega' \in \Omega \wedge \\ \omega = \rho_{Par(a,N)}(\omega') \} \big]\!\!\big] .\end{aligned}$$

In this view, each table over $Par(a, N)$ expands into a table over $\mathcal{A}$ in which each row over $\Omega_{Par(a,N)}$ subdivides into rows over $\Omega_{\mathcal{A}}$ that agree on the attributes in $Par(a, N)$. The order specified in each of the subdivision rows is the same as the *ceteris paribus* order specified in the subdivided row when expanded from $\Omega_{Par(a,N)}$ to $\Omega_{\mathcal{A}}$. More precisely, each of the $|\Omega_{Par(a,N)}|$ rows in the table for attribute a expands into

$$|\Omega_{\mathcal{A} \setminus Par(a,N)}| = \frac{|\Omega_{\mathcal{A}}|}{|\Omega_{Par(a,N)}|}$$

rows in the table over all attributes in $\mathcal{A}$.

To facilitate the discussion, we define the *expansion* $\eta(t, A')$ of a CP-table $t \in T(A, a)$ to a table $t' \in T(A', a)$ over an extended set of attributes $A' \supseteq A$ by the requirement that $t'(\omega) = t(\rho_A(\omega))$ for each $\omega \in \Omega_{A'}$. Figure 1 depicts a simple example of a CP-table expansion from $\Omega_{\{X\}}$ to $\Omega_{\{X,Z\}}$.

Observe that orders stated in the expanded CP-table do not depend on the values of the expansion attribute, and are identical with the values assigned for the original attributes in the original table.

Metrics on CP-networks

Referential distance

To obtain a distance metric on CP-networks, we only need to apply the distance metric on orders to the orders denoted by the CP-networks. Formally, we define the *referential* metric $d^{\mathrm{r}} : \mathcal{N} \times \mathcal{N} \to \mathbb{R}$ by

$$d^{\mathrm{r}}(N, N') \stackrel{\mathrm{def}}{=} d([\![N]\!], [\![N']\!]). \tag{5}$$

Although the referential metric provides a reasonable and precise comparison between CP-networks, it has terrible computational properties. First of all, computing the distance between two networks over the same n values requires examination of on the order of $(2^n)^2$ comparisons between outcomes for binary attributes, and examination of even more for attributes with more than two values. Such computations are infeasible except for the smallest networks. Second, the indirect connection of network distance with network structure makes it difficult to predict the magnitude of distances between networks from network differences themselves. Put together, these problems impel one to seek metrics on networks defined in terms of the structural properties alone, without the referential detour through the orders over outcomes indicated by the network semantics.

For the purpose of measuring similarity between preference orders, we need not demand perfect agreement between d^{r} and some new distance measure d', only strategic equivalence in the sense that the two measures agree on relative distance comparisons. Formally, we seek an efficiently-computable d' such that for every $N, N', N'' \in \mathcal{N}$, we have

$$d^{\mathrm{r}}(N, N') < d^{\mathrm{r}}(N, N'') \text{ iff } d'(N, N') < d'(N, N''). \tag{6}$$

Simple structural distance

To avoid the high cost of computing the referential distance measure, we look to identify distance measures defined directly over the CP-network representation rather than indirectly through the meanings of these networks. The simplest candidate along these lines is an edit distance measure taking into account both graph and order elements.

In (Wicker & Doyle 2008), we analyze five types of CP-network edit operations: addition of a new attribute, removal of an attribute, addition of an edge, removal of an edge, and change of preference order.

We clearly can transform any network into the empty network by successively changing all CP-table entries to the empty order, then removing all edges, and finally removing all nodes. We can thus transform any network into any other by transforming the first to the empty network and then inverting the sequence of operations needed to reduce the other network to the empty network. Indeed, this sort of transformation through "zero" forms the basis for the KSB axioms on order distances.

The main drawback of this simple structural metric is that it yields distances and relative comparisons at odds with those obtained using the referential semantics.

Table expansion metric

To find a metric on networks that has the low-cost computability of the simple edit distance but exhibits order-compatibility with the underlying reference metric, we introduce the *table-expansion metric* $d^{\mathrm{e}} : \mathcal{N} \times \mathcal{N} \to \mathbb{R}$, defined for networks $N = (A, E, t)$ and $N' = (A', E', t')$ by

$$d^{\mathrm{e}}(N, N') \stackrel{\mathrm{def}}{=} \sum_{a \in \mathcal{A}} \sum_{\omega \in \Omega_{\mathcal{A} \setminus \{a\}}} d(\eta(t(a), \mathcal{A} \setminus \{a\})(\omega), \eta(t'(a), \mathcal{A} \setminus \{a\})(\omega)) \tag{7}$$

This measure compares orders specified by tables under all conditions. To do this, it interprets each row of each table as specifying one or more entries in the full condition table, and adds up the KSB distances between the orders indicated by each of these maximally-specific rows.

The table expansion metric works no matter what size the original networks are, and works even if the orders specified in table entries are not total orderings of the attribute values. The table expansion metric also works with tables that lack some rows, if one regards the missing rows as having empty orders.

Computing the table expansion metric directly is not feasible because there are exponentially many rows in the full condition table. However, one can compute the metric in time proportional to the size of the table by simply finding the least common refinements of comparable nodes in the networks under comparison, and then weighting each table entry of the nodes under comparison by the number of conditions in the full expansion.

Theorem 2. *If* $N = (A, E, t)$, $N' = (A', E', t')$, $A^+ = A \cup A'$, *and* $Par^+(a) = Par(a) \cup Par'(a)$, *then*

$$d^{\mathrm{e}}(N, N') = (|\Omega_{A^+}|)^{-1} \cdot \sum_{a \in A^+} |\Omega_{(A^+ \setminus Par^+(a))}| \cdot \sum_{\omega \in \Omega_{Par^+(a)}} d(\eta(t(a), Par^+(a))(\omega), \eta(t'(a), Par^+(a))(\omega))$$

Consider the CP-networks N_1 and N_2 depicted in Figure

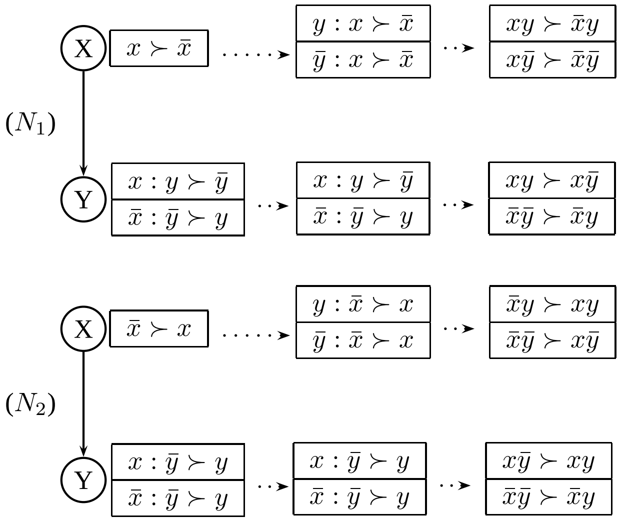

Figure 2: Two CP-network expansions of tables. Each CP-network (left column) is shown with the expanded CP-tables (middle column) and the the orders specified by each row (right column).

2. We can see the expansion of each of the CP-table rows into the maximally specific CP-table. We count the KSB distances between each of the corresponding rows in these maximally specific tables and get $d^{\mathrm{e}}(N_1, N_2) = 3$. For comparison, we also get $d^{\mathrm{r}}(N_1, N_2) = 9$.

Although a CP-network might have nodes that are children of all the nodes, in practical applications one expects to see bounded branching in the networks. In this case, distances can be computed efficiently.

Theorem 3. *For each integer k, the expansion distance between two networks in which the number of parents of nodes is bounded by k can be computed in time polynomial in the sizes of the two networks.*

It is not hard to show that expansion distance d^{e} provides a lower bound on referential distance d^{r}.

Theorem 4. $d^{\mathrm{e}}(N, N') \leq d^{\mathrm{r}}(N, N')$ *for each* $N, N' \in \mathcal{N}$.

It is also not hard to show half of the desired strategic equivalence (6).

Theorem 5. *If* $d^{\mathrm{r}}(N, N') \leq d^{\mathrm{r}}(N, N'')$, *then* $d^{\mathrm{e}}(N, N') \leq d^{\mathrm{e}}(N, N'')$.

We currently lack proof or disproof of the other half of the desired equivalence.

Conclusions and future work

We have described a table expansion metric on CP-networks that uses expansions of CP-tables to determine similarity. This approach builds upon a previous approach, which utilized the KSB distance metric to determine similarity between preference orderings represented by CP-networks. The table expansion metric, however, provides a much more efficient computation than the KSB distance on the induced orderings.

References

Bogart, K. P. 1973. Preference structures I: Distances between transitive preference relations. *Journal of Mathematical Sociology, 3(1)*, 49-67.

Boutilier, C.; Brafman, R. I.; Hoos, H. H.; and Poole, D. 1999. Reasoning with conditional ceteris paribus preference statements. In *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence, Stockholm, Sweden*, 71-80.

Boutilier, C.; Brafman, R. I.; Domshlak, C.; Hoos, H. H.; and Poole, D. 2004. CP-nets: A tool for representing and reasoning with conditional ceteris paribus preference statements. *Journal of Artificial Intelligence Research, 21*, 135-191.

Chomicki, Jan. 2003. Preference Formulas in Relational Queries. *ACM Transactions on Database Systems, 28(4)*, 427-466.

Doyle, J.; Shoham, Y.; and Wellman, M. P. 1991. A logic of relative desire (preliminary report). In *Proceedings of the Sixth International Symposium on Methodologies for Intelligent Systems, Charlotte, North Carolina, USA*, 16–31.

Kemeny, J. G., and Snell, J. L. 1962. *Mathematical Models in the Social Sciences*. MIT Press.

Kiessling, Werner. 2004. Preference Constructors for Deeply Personalized Database Queries. Technical Report, Department of Computer Science, Universität Augsburg.

Wellman, M. P., and Doyle, J. 1991. Preferential semantics for goals. In *Proceedings of the Ninth National Conference on Artificial Intelligence, Anaheim, California, USA*, 698-703.

Wicker, A. W. 2006. Interest-matching comparisons using CP-nets. Master's thesis, Department of Computer Science, North Carolina State University.

Wicker, A. W., and Doyle, J. 2007. Interest-matching comparisons using CP-nets. In *Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence, Vancouver, British Columbia, Canada*, 1914-1915.

Wicker, A. W., and Doyle, J. 2008. Comparing preferences expressed by CP-networks. *In preparation*.