Voting with CP-nets using a Probabilistic Preference Structure

Cristina Cornelio, Umberto Grandi, Judy Goldsmith, Nicholas Mattei,
Francesca Rossi and K. Brent Venable

Abstract

Probabilistic conditional preference networks (PCP-nets) provide a compact representation of a probability distribution over a collection of CP-nets. In this paper we view a PCP-net as the result of aggregating a collection of CP-nets into a single structure. We use the resulting PCP-net to perform collective reasoning tasks, e.g. determining the most preferred alternative, when a group of agents expresses their preferences via CP-nets. We propose several PCP-net based methods to perform CP-net aggregation and evaluate these methods both axiomatically and experimentally.

1 Introduction

The study of preferences plays an important role in the field of artificial intelligence [19] and machine learning [7]. The ability to express preferences in a faithful way, which can be handled efficiently, is essential in many reasoning tasks. In settings such as e-commerce, on demand video, and other settings where supply outstrips an individuals ability to view all the available choices, we require an efficient formalism to model and reason with complex, interdependent preferences [8]. We may also use these preferences to make decisions about joint plans, actions, or items in multi-agent environments [18]. Agents express their preferences over a set of alternative decisions, these preferences are aggregated into one decision which satisfies as many agents as possible.

Often multi-attribute preference modeling and reasoning causes a combinatorial explosion, often leading to high computational cost [5, 6, 9]. The set of alternatives is often described as a product of multiple features, for example, a user's preferences over a set of cars, which can be described by their colors, technical specifications, cost, reliability, etc. A number of *compact representation languages* have been developed to tackle the computational challenges arising from these problems. Among others, we mention conditional preference structures (CP-nets) [3], soft constraints [2, 19], and GAI-nets [10]. In this paper we focus on CP-nets as the tool for modeling the preferences of a single agent. CP-nets are a qualitative preference modeling framework that allow for conditional preference statements.

Preferences are often uncertain: we may be unsure about our preference ordering over certain items, or there could be noise in our preference structure due to lack of precision in elicitation or sensor collection. We may also need to represent preferences that are directly conflicting, such as disagreement in multi-agent systems or voting settings [14, 15, 20]. PCP-nets model uncertain preferences natively while using the same preferential dependency structure used with CP-nets [1, 4]. However, in a PCP-net, a preference ordering over a variable's domain is replaced by a probability distribution over all possible preference orderings of the variable's domain. Thus, a PCP-net defines a probability distribution over a collection of CP-nets: all those CP-nets that can be obtained from the PCP-net by choosing a variable ordering from the distribution over all orderings. Given a PCP-net, one can define the optimal variable assignment in two natural ways: as the most probable optimal outcome, or as the optimal outcome of the most probable CP-net induced by the PCP-net. If the dependency structure of the PCP-net has a bounded size, both kinds of optimal outcomes can be found in time polynomial time in the size of the PCP-net.

In this paper we investigate the use of PCP-nets as a way to aggregate the preferences expressed by a collection of CP-nets. When a set of agents each expresses their preferences as a CP-net, we may want to find the alternative which is most preferred by the agents, or decide whether one alternative is collectively preferred to another.

Previous efforts to tackle these collective reasoning problems have primarily focused on the question of optimality. Generally, a voting method is defined in order to determine the single most preferred alternative. These methods are usually sequential and work at the feature level of the CP-nets [11, 12, 13, 20, 21]. In contrast, we aggregate the collection of CP-nets into a single structure, a PCP-net, on which we can directly perform collective reasoning tasks. By maintaining structure, a PCP-net can then be used for other reasoning tasks, such as extracting the complete ranking over alternatives, instead of only determining the winner. The PCP-net also allows us to compute the probability of dominance: given two outcomes we can compute the probability with which the first is preferred to the second. Our aggregation methods allow us to store only a single, compact, aggregated structure instead of a large collection of CP-nets. Additionally, this aggregated structure, the PCP-net, allows us to preform other collective reasoning tasks that a voting method which only returns a top element, would not allow us to compute.

We propose two methods for aggregating a collection of CP-nets into a single PCP-net: by extracting the probability distribution directly from the CP-nets (proportion method) or by minimizing the error between the aggregated structure and the original set of CP-nets (least-squares method). We combine these two methods with the two ways of extracting an optimal outcome from a PCP-net (most probable optimal outcome and the optimal outcome of the most probable CP-net), obtaining four approaches, each of which can be seen as a voting rule that take as input a profile of individual CP-nets and outputs a winning alternative. We show that the output of the sequential voting method with majority for CP-nets [11] coincides with one of the four methods we define and the four methods we propose can return a disjoint set of outcomes. We analyze the methods according to several axiomatic properties usually considered to be desirable in voting rules [16] and through an experimental comparison of the four methods. For the empirical experiments, we compute the output of each of the four methods and comparing these outputs via dominance queries on the original set of CP-nets, determining which alternative most satisfies the individual agents' preferences. The experimental results show that the proportional methods strictly dominate the least-squares methods. Moreover, the optimal outcome of the most probable induced CP-net is always better than the most probable optimal outcome.

Notice that the outcome computed by the proportional method combined with taking the optimal outcome of the most probable induced CP-net coincides with the result of the sequential aggregation procedure, proposed by Lang and Xia [11]. This is not surprising, as Lang and Xia's sequential method was designed to obtain the outcome which best satisfies the preferences of the whole collection of agents. However, we can obtain this outcome by generating a PCP-net representing the collection of given CP-nets. This allows us to do more than just find a collectively optimal outcome, since the PCP-net can also be used to answer dominance queries or to find the next best outcome in a linearization of the induced preference ordering.

2 Background

In this section we give a brief introduction to CP-nets and PCP-nets. Throughout we assume that the domain of each variable is binary and the induced-width [4] of the dependency graph is bounded from above by a constant.

2.1 CP-nets

A CP-net is a graphical model for compactly representing conditional and qualitative preference statements [3]. CP-nets exploit conditional preferential independence by decomposing an agent's preferences via the *ceteris paribus* assumption (all other things being equal).

Definition 1 *A **CP-net** is a directed graph where each node represents a variable (often called features) $F = \{X_1, \ldots, X_n\}$ each with finite domains $\mathcal{D}(X_1), \ldots, \mathcal{D}(X_n)$. For each feature X_i, there is a set of* parent *features $Pa(X_i)$ that can affect the preferences over the values of X_i. This defines a* dependency graph *in which each node X_i has edges from all features in $Pa(X_i)$. Given this structural information, the user explicitly specifies her preference over the values of X_i for* each complete assignment *on $Pa(X_i)$. This preference is a total order over $\mathcal{D}(X)$.*

An *outcome* in a CP-net is a complete assignment to all the variables. The semantics of CP-nets depends on the notion of a *worsening flip*, which is a change in the value of a variable to a value which is less preferred by the cp-statement for that variable. We say that one outcome α is *better* than another outcome β ($\alpha > \beta$) if and only if there is a chain of worsening flips from α to β. This definition induces a preorder over the outcomes.

In general, finding an optimal outcome (that is, an outcome that has no other outcome better than it) and testing for optimality in this ordering is NP-hard. However, in acyclic CP-nets, there is only one optimal outcome and this can be found in as many steps as the number of features via a *sweep forward procedure* [3] which is linear in the number of features. On the other hand, determining if one outcome is better than another according to this ordering, called a dominance query, is NP-hard even for acyclic CP-nets [5, 9].

In this paper we consider a collection of CP-net, also called a *profile* in voting theory terminology. A profile of CP-nets is a set of CP-nets on the same set of variables: $P = (C_1, \cdots, C_m)$. In this paper we are only interested in O-legal profiles of acyclic CP-nets where each variable has a binary domain.

Definition 2 *(O-legality) Given a profile of m CP-nets $C_1, \cdots, C_m$ over n variables $X_1, \cdots, X_n$, with dependency graphs $D_1, \cdots, D_m$. Let us call D the union of such D_i's. Consider also a linear order on the variables $O = X_1 \prec \cdots \prec X_n$. The profile is said to be O-legal if, for each edge (X_i, X_j) in D, we have $X_i \prec X_j$ in O.*

2.2 PCP-nets

A PCP-net (Probabilistic CP-net) is a generalization of a CP-net, where, for each feature X, instead of giving a preference ordering over the domain of X, we give a probability distribution over the set of all preference orderings for the domain of X [1, 4]. Given a feature X in a PCP-net, its *PCP-table* is a table associating each combination of the values of the parent features of X with a probability distribution over the set of total orderings over the domain of X.

Given a PCP-net $\mathcal{Q}$, a *CP-net induced by $\mathcal{Q}$* has the same variables, each with the same domain, as $\mathcal{Q}$. The dependency edges of the induced CP-net are a subset of the edges in the PCP-net. Thus, CP-nets induced by the same PCP-net may have different dependency graphs. Moreover, the CP-tables are generated accordingly for the chosen edges. For each independent feature, one ordering over its domain (i.e., a row in its PCP-table) is selected. Similarly, for dependent features, an ordering is selected for each combination of the values of parent features. Each induced CP-net has an associated probability, obtained from the

PCP-net by taking the product of the probabilities of the deterministic orderings chosen in the CP-net.

More precisely, given a PCP-net, we have a probability p for each row $u : x > \bar{x}$ of each PCP-table, while the probability of $u : \bar{x} > x$ corresponds to $1 - p$. The probability of an induced CP-net C is f_C, computed as the product of the probability p or $1 - p$ of the rows chosen for the CP-net.

Example 1 *Consider the PCP-net shown with two features, X_1 and X_2, with domains $\mathcal{D}_{X_i} = \{x_i, \bar{x}_i\}$.*

Structure:

$$X_1 \xrightarrow{p} X_2$$

Feature X_1:

X orderings	$\mathbb{P}$
$x_1 > \bar{x}_1$	r
$\bar{x}_1 > x_1$	$1 - r$

Feature X_2:

A values	*B orderings*	$\mathbb{P}$
x_1	$x_2 > \bar{x}_2$	q_1
	$\bar{x}_2 > x_2$	$1 - q_1$
$\bar{x}_1$	$x_2 > \bar{x}_2$	q_2
	$\bar{x}_2 > x_2$	$1 - q_2$

A induced CP-net C (with probability $f_C = [(1 - r) \cdot (1 - q_1) \cdot q_2]$) is shown below.

Structure:

$$X_1 \longrightarrow X_2$$

Feature X_1:

A orderings
$\bar{x}_1 > x_1$

Feature X_2:

A values	*B orderings*
x_1	$\bar{x}_2 > x_2$
$\bar{x}_1$	$x_2 > \bar{x}_2$

Since we have a probability distribution over the set of all induced CP-nets, we can give the following definitions:

Definition 3 *Given a PCP-net, its* ***most probable optimal outcome*** *is the outcome with the highest probability. We compute the probability of an outcome o being optimal by taking the sum of the probabilities of the induced CP-nets that have o as the optimal outcome.*

Definition 4 *Given a PCP-net, its* ***most probable induced CP-net*** *is the induced CP-net with the highest probability, considering the probability distribution induced by the PCP-net.*

Notice that the optimal outcome of the most probable induced CP-net may be different from the most probable optimal outcome [4]. To compute these two outcomes, it is possible to generate two Bayesian Networks and compute their maximal joint probability [17]. Computing the result for either notion of optimal outcome has polynomial computational complexity if the induced width [4] of the dependency graph of the PCP-net is bounded.

3 Building the PCP-net

We assume an O-legal profile of m CP-nets over variables $X_1, \cdots, X_n$ each with binary domain. The profile may contain several occurrences of the same CP-net, so each CP-net comes with its frequency (that is, the number of its occurrences). We can define the *normalized frequency* of a CP-net C_i as $\mathbb{P}(C) = p_i = \frac{f_i}{m}$ where f_i is the frequency of the CP-net C_i and m the number of users. Thus a profile of m CP-nets can also be written as $P = ((C_1, f_1), \cdots, (C_k, f_k))$, with $\sum_{i=1}^{k} f_i = m$.

Given a set of CP-nets defined on the same variables and with a probability distribution over them, a PCP-net that generates exactly this distribution may not exist. This can be seen in the following profile of CP-nets over two variables X_1 and X_2:

- $(C_1, 0.5)$: $(x_1 > \bar{x}_1)$, $(x_1 : x_2 > \bar{x}_2)$ and $(\bar{x}_1 : \bar{x}_2 > x_2)$
- $(C_2, 0.4)$: $(x_1 > \bar{x}_1)$, $(x_1 : \bar{x}_2 > x_2)$ and $(\bar{x}_1 : x_2 > \bar{x}_2)$
- $(C_3, 0.1)$: $(x_1 > \bar{x}_1)$ and $(x_2 > \bar{x}_2)$

The PCP-net representing such a profile must satisfy the following system of equations, where p is the probability of $x_1 > \bar{x}_1$, q is the probability of $x_1 : x_2 > \bar{x}_2$, and r is the probability of $\bar{x}_1 : x_2 > \bar{x}_2$:

$$\begin{cases} f_{C_1} : pq(1-r) = 0.5 \\ f_{C_2} : p(1-q)r = 0.4 \\ f_{C_3} : pqr = 0.1 \end{cases} \quad \Rightarrow \quad \begin{cases} p = 2.4 \\ q = 0.25 \\ r = 0.12 \end{cases}$$

This solution is unique but not feasible, as p, q and r are probabilities so they should all be at most 1.

In general, given a profile with n features $X_1, \cdots, X_n$, we have a number of probability variables equal to

$$N_p = \sum_{i=1}^{n} 2^{|Pa(X_i)|},$$

while the number of equations is

$$2^{\sum_{i=1}^{n} 2^{|Pa(X_i)|}} = 2^{N_p}.$$

The system is thus over-constrained and will rarely admit a solution. Therefore, this aggregation method is not a feasible one for aggregating even O-legal CP-net profiles. In the next section we will define other aggregation approaches.

3.1 Aggregation Methods

In this section we define two methods to represent a profile of CP-nets using a PCP-net. As noted above, we are not guaranteed to find a PCP-net representing the exact distribution of induced CP-nets in the profile. Thus we must resort to methods approximating this ideal distribution.

The first method we propose generates a PCP-net by taking the union of the dependency graphs of the given CP-nets and determining the probabilities in the PCP-tables from the frequency of the CP-nets in the profile.

Definition 5 *Given a profile of CP-nets $P = ((C_1, f_1), \cdots, (C_k, f_k))$, the* Proportion (PR) *aggregation method defines a PCP-net whose dependency graph is the union of the graphs of the individual CP-nets in the profile. Given a variable X and an assignment u to its parents, the probabilities in the PCP-tables are defined as follows:*

$$\mathbb{P}(x > \bar{x}|u) = \sum_{C_i : x > \bar{x}|u} \mathbb{P}(C_i)$$

$$\mathbb{P}(\bar{x} > x|u) = 1 - \sum_{C_i : x > \bar{x}|u} \mathbb{P}(C_i).$$

That is, the probability of the ordering $x > \bar{x}$ for variable X, given assignment u of the parents of X, is the sum of the probabilities of the CP-nets that have that particular ordering over the domain of X.

The second method minimizes the mean squared error between the probability distribution induced by the PCP-net over the CP-nets given in input and the normalised frequency observed in the input.

Definition 6 *Let $P = ((C_1, f_1), \cdots, (C_k, f_k))$ be a profile of CP-nets. The* Least Square (LS) *aggregating method defines a PCP-net where the graph is the union of the graphs of the CP-nets and the probabilities q^i_j in the PCP-tables (where q^i_j is the probability of the j-row of the PCP-table of the variable X_i in the PCP-net) solve the following:*

$$\underset{\boldsymbol{q} \in [0,1]^r}{\operatorname{argmin}} \sum_C (f_C(q) - \mathbb{P}(C))^2$$

where $\boldsymbol{q}$ is the vector of q^i_j ordered lexicographically with i as first variable, C varies over all CP-nets induced by the union graph, $f_{C_i}(q)$ are the formulas introduced in Section 2.2, and $\mathbb{P}(C) = p_i$ if $C = C_i$, $\mathbb{P}(C) = 0$ otherwise.

It is important to observe that computing the PCP-net using method PR may require exponential time, as the PCP-net resulting from a generic profile may have an exponential number of cp-statements. However, we can ensure that the union graph of an O-legal profile has bounded width – making PR a polynomial method – by requiring the following O-boundedness condition: for each feature j there are sets $PP(X_j) \subseteq \{X_1, \ldots, X_n\}$ of possible parents such that (i) $|PP(X_j)| < k$ for all j, and (ii) for all individuals i, $Pa_i(X_j) \subseteq PP(X_j)$. On the other hand, method LS requires writing an exponential number of equations, one for each induced CP-net. For this reason, in some of the experiments in Section 5 we use a modified version of LS which uses a linear number of equation but, as a downside, is a further approximation of the probability distribution over the induced CP-nets.

3.2 Voting Rules

Let $\mathcal{P}$ be the set of all CP-net profiles P of m voters over a set of alternatives X, a CP-voting rule $r : \mathcal{P} \to X$ is a function that maps each profile P into a alternative $r(P) \in X$.[1]

We define four CP-voting rules by combining the two aggregation methods PR and LS presented in Definition 5 and 6 with the two possible ways of extracting an optimal outcome from a PCP-net presented in Definitions 3 and 4:

- PR_O: PR and most probable optimal outcome;
- PR_I: PR and optimal outcome of most probable induced CP-net;
- LS_O: LS and most probable optimal outcome;
- LS_I: LS and optimal outcome of most probable induced CP-net.

Computing the optimal outcome for PR_O and PR_I is polynomial if the graph of the resulting PCP-net has bounded width. This is the O-boundedness condition introduced at the end of Section 3.1.

We first observe that PR_I returns the same result as the *sequential voting rule with majority* [11], that consists of applying the majority rule "locally" on each issue in the order given by O.

[1] In what follows we assume CP-voting rules are resolute and return a unique winner.

Theorem 1 *Given any profile of CP-nets, PR_I produces the same result as sequential voting with majority.*

Proof. Consider a variable X_i with domain $\{x_i, \bar{x}_i\}$, and an assignment u for the parents of X_i. With sequential voting we choose the value of the domain that corresponds to the first value of the ordering that maximizes the following:

$$\max_{j \in \{1,\cdots,m\}} \{[\sum_{C_j : x_i > \bar{x}_i | u} \mathbb{P}(C_j)], [1 - \sum_{C_j : x_i > \bar{x}_i | u} \mathbb{P}(C_j)]\}.$$

With PR_I, we create a PCP-net that has, for the row in the PCP-table of X_i corresponding to assignment u for its parents, the probability $\sum_{C_j : x_i > \bar{x}_i | u} \mathbb{P}(C_j)$ for $x_i > \bar{x}_i$ and $1 - \sum_{C_j : x_i > \bar{x}_i | u} \mathbb{P}(C_j)$. To compute the most probable induced CP-net we choose the ordering with maximal probability, thus:

$$\max_{j \in \{1,\cdots,m\}} \{[\sum_{C_j : x_i > \bar{x}_i | u} \mathbb{P}(C_j)], [1 - \sum_{C_j : x_i > \bar{x}_i | u} \mathbb{P}(C_j)]\}.$$

To compute the optimal outcome of this CP-net, we choose the first values of the orderings that appear in the CP-table. This is for a generic variable X_i and assignment u, thus is true for all the variables and assignment of their parents. ■

The four CP-voting rules defined above can give different results. For example, consider a PCP-net on two variables X_1 and X_2 with the PCP-tables $(x_1 > \bar{x}_1, 0.6)$, $(x_1 : x_2 > \bar{x}_2, 0.6)$ and $(\bar{x}_1 : x_2 > \bar{x}_2, 0)$. The most probable induced CP-net has $x_1 > \bar{x}_1$, $x_1 : x_2 > \bar{x}_2$ and $\bar{x}_1 : \bar{x}_2 > x_2$, thus $x_1 x_2$ is the optimal outcome. However, the most probable optimal outcome of the PCP-net is $\bar{x}_1 \bar{x}_2$. Therefore, PR_O (respectively, LS_O) will differ from PR_I (resp. LS_I) on a profile P that includes this CP-net. We can also prove the following property:

Theorem 2 *There exists a profile of CP-nets P such that $\{PR_O(P), PR_I(P)\} \cap \{LS_O(P), LS_I(P)\} = \emptyset$.*

Proof. Let us take the following profile P of four CP-nets over two variables A and B:

- $\mathbf{C_1}$ with probability 0.095. C_1 has the edge from X_1 to X_2 and CP-tables: $x_1 > \bar{x}_1$ and $x_1 : x_2 > \bar{x}_2$, $\bar{x}_1 : \bar{x}_2 > x_2$.
- $\mathbf{C_2}$ with probability 0.505. C_2 has the edge from X_1 to X_2 and CP-tables: $x_1 > \bar{x}_1$ and , $x_1 : \bar{x}_2 > x_2$, $\bar{x}_1 : x_2 > \bar{x}_2$.
- $\mathbf{C_3}$ with probability 0.005. C_3 does not have the edge from X_1 to X_2 and has CP-tables: $\bar{x}_1 > x_1$ and $x_2 > \bar{x}_2$.
- $\mathbf{C_4}$ with probability 0.395. C_4 does not have the edge from X_1 to X_2 and has CP-tables: $\bar{x}_1 > x_1$ and $\bar{x}_2 > x_2$.

Proportion gives us the PCP-net $(x_1 > \bar{x}_1, 0.6)$ and $(x_1 : x_2 > \bar{x}_2, 0.51)$, $(\bar{x}_1 : x_2 > \bar{x}_2, 0.1)$, while LS outputs $(x_1 > \bar{x}_1, 0.59)$ and $(x_1 : x_2 > \bar{x}_2, 0.29)$, $(\bar{x}_1 : x_2 > \bar{x}_2, 0)$. Thus we obtain that $PR_O(P) = \bar{x}_1 \bar{x}_2$, $PR_I(P) = x_1 x_2$, $LS_O(P) = x_1 \bar{x}_2$ and $LS_I(P) = x_1 \bar{x}_2$. ■

4 Axiomatic Properties

A first way to compare the four voting rules is by checking if they satisfy various desirable axiomatic properties [16].

- **Anonymity** holds when the result is not sensitive to any permutation of the voters (that is, given a permutation σ on voters, $r(\sigma(P)) = r(P)$).
- **Neutrality** holds if, for any profile P and any permutation σ on alternatives X, then $r(P, X) = r(P, \sigma(X))$.
- **Homogeneity** holds when, for any profile P and any $s \in \mathbb{N}$, we have $r(P) = r(sP)$.
- **Opt-Monotonicity** holds if, given two profiles $P = (C_1, \cdots, C_m)$ and $P' = (C'_1, \cdots, C'_m)$ where C'_i is obtained by C_i by changing the CP-tables so that $r(P)$ is the optimal outcome for C'_i, we have $r(P) = r(P')$
- **Consistency** holds if, given two disjoint profiles P_1 and P_2 such that $r(P_1) = r(P_2)$, we have $r(P_1 \cup P_2) = r(P_1) = r(P_2)$.
- **Participation** holds if, for any profile P and any CP-net C, we have $r(P \cup \{C\}) >_C r(P)$.
- **Consensus** holds if, for any profile $P = (C_1, \cdots, C_m)$, there is no alternative o such that $o >_{C_i} r(P)$, for all $i \in \{1, \cdots, m\}$.

Anonymity and neutrality both hold for all four voting rules. We know PR_I satisfies a stronger version of monotonicity and consistency (hence homogeneity) as it coincides with the sequential voting on O-legal profiles studied by Lang and Xia [11].

Theorem 3 *PR_O satisfies homogeneity.*

Proof. Consider a profile $P = (C_1, f_1), \cdots, (C_k, f_k)$ from which we can get the normalised frequencies $(C_1, \frac{f_1}{m}), \cdots, (C_k, \frac{f_k}{m})$. Considering N times each CP-net, we obtain the following distribution over Nm CP-nets $P' = (C_1, Nf_1), \cdots, (C_k, Nf_k)$ with the following normalised frequencies $(C_1, \frac{Nf_1}{Nm}), \cdots, (C_k, \frac{Nf_k}{Nm})$. The probability of a generic CP-net C_i in P' is $\frac{Nf_i}{Nm} = \frac{f_i}{m}$ which is the same as the probability generated by P. ∎

Theorem 4 *LS_O and LS_I satisfy homogeneity.*

Proof. In the proof of Theorem 3 we add N copies of each CP-net to the original profile resulting in the same probability distribution over the CP-nets. This fact is true for any collection of CP-nets, therefore, we generate the same set of equations to minimize, and thus the same solution (PCP-net). ∎

Theorem 5 *PR_O and PR_I satisfy opt-monotonicity.*

Proof. Let us consider two profiles $P = (C_1, \cdots, C_m)$ and $P' = (C'_1, \cdots, C'_m)$ where C'_i is obtained by C_i by changing the CP-tables so that $r(P)$ is the optimal outcome for C'_i. Let $\mathcal{Q}$ and $\mathcal{Q}'$ represent, respectively, the PCP-nets obtained using the PR aggregation method on P and P'. Let S be the set of rows of the PCP-tables in $\mathcal{Q}$ relevant to $r(P)$. By definition of PR, the only difference between $\mathcal{Q}$ and $\mathcal{Q}'$ is in the probabilities of the rows in S. More specifically, the changes required to obtain C'_i from C_i are such that the probability of the orderings favoring the values assigned in $r(P)$ will be higher in $\mathcal{Q}'$, while those favoring the values opposite to those in $r(P)$ will be lower. Thus, we have that $r(P) = r(P')$.

The result for PR_I can be directly obtained from Theorem 1, that is, by using the equivalence of PR_I with sequential majority. Let us denote with o the optimal outcome of C_i. By replacing C_i with C'_i we increase by one the number of votes for the values that are in $r(P)$ and not in o, we decrease by one those votes for values that are in o and not in $r(P)$ and we leave the vote count the same for the other values (that is, the common ones) unchanged. Thus, since $r(P)$ was the winner for sequential majority given P it will still be the winner given P'. Given Theorem 1 we can conclude that the same holds for PR_I. ∎

Theorem 6 *PR_I and PR_O satisfy participation.*

Proof. We first consider the case $r = PR_I$. Consider a profile $P = (C_1, \cdots, C_m)$ and an additional CP-net C. We have to prove that $r(P \cup \{C\}) >_C r(P)$. PR gives us a PCP-net $\mathcal{Q}$ for P and a PCP-net $\mathcal{Q}'$ for $P \cup \{C\}$. Since P and $P \cup \{C\}$ are O-legal, let X be the first variable according to O such that $r(P \cup \{C\})|_X \neq r(P)|_X$ and let $u = r(P)|_{PA(X)}$. This means that in the most probable induced CP-net of $\mathcal{Q}$ there is a preference statement $u : r(P)|_X > r(P \cup \{C\})|_X$ and in $\mathcal{Q}'$ there is $u : r(P \cup \{C\})|_X > r(P)|_X$. Thus the probability $\mathbb{P}[u : r(P)|_X > r(P \cup \{C\})|_X] \geq 0.5$ in $\mathcal{Q}$, but ≤ 0.5 in $\mathcal{Q}'$. This means that C has the row $u : r(P \cup \{C\})|_X > r(P)|_X$ in X's CP-table. Thus the outcome $r(P \cup \{C\}) \geq_C r(P)$.

A similar reasoning applies to the case $r = PR_O$. In the PCP-tables of $\mathcal{Q}'$, the probabilities of the rows corresponding to the CP-tables of C increase with respect to $\mathcal{Q}$. Thus, for each feature X and each assignment u of its parents, the probability to obtain the first ranked value in the u rows of C CP-tables, increases. This means that it improves the result for C. ■

Theorem 7 *PR_I satisfies consensus and PR_O satisfies consensus over a single feature.*

Proof. We first consider the case $r = PR_I$. Consider profile $P = (C_1, \cdots, C_m)$ and assume there is an alternative o s.t. $o >_{C_i} r(P) \forall i \in \{1, \cdots, m\}$. Since P is O-legal, let X be the first variable according to O such that $o|_X \neq r(P)|_X$. Let u be the assignment to X's parents in o and $r(P)$ ($o|_{PA(X)} = r(P)_{PA(X)}$ since X is the first variable according to O in which they differ). Since $o >_{C_i} r(P) \forall i \in \{1, \cdots, m\}$, it must be that $u : o|_X > r(P)|_X \forall C_i$. Let us now consider the PCP-net $\mathcal{Q}$ obtained from P by PR. It is easy to see that in the PCP-table of X the probability of $u : o|_X > r(P)|_X$ will be strictly higher than the probability of $u : o|_X < r(P)|_X$, which implies that the most probable induced CP-net must have the row $u_i : o|_X > r(P)|_X$. The optimal outcome of the most probable induced CP-net must have the assignment $o|_X$ for the variable X because is ranked first in the table in the u row. But $r(P)|X \neq o|X$ and we have a contradiction.

A similar reasoning applies to the case $r = PR_O$, but in a weaker version. We will prove that, for any profile P, there is no alternative o such that o differs from $r(P)$ on only a single variable and $o >_{C_i} r(P) \forall C_i$. Assume that there were an alternative o such that $o >_{C_i} r(P)$, $\forall C_i$ and o differs from $r(P)$ only on the variable X. The probability of $u : o|_X > r(P)|_X$ in the PCP-table of X is equal to 1 because all the CP-nets in the profile prefer o to $r(P)$. Thus the probability of $r(P)$ is equal to 0, and we have a contradiction. ■

In conclusion, the aggregation method PR, generating the voting rules PR_O and PR_I, satisfies a good number of desirable axiomatic properties. Obtaining results for the LS method is rather hard, given that it is based on numerical optimization. In the next section we will compare the two methods experimentally.

5 Experimental analysis

In this section we test the quality of the outcomes of the four voting rules defined in Section 3.2. We compare the results of these four voting rules with a baseline of random dictatorship named $RAND$, which outputs the optimal outcome of a random CP-net in the profile. We compare these five voting rules using two different scoring functions, each of which is computed using dominance queries on the input profile of CP-nets.

First, given the intractability of the LS method, we introduce the following approximation of LS called $\widetilde{LS}$:

Definition 7 *Let P be a profile of CP-nets. The $\widetilde{LS}$ aggregation method defines a PCP-net as the LS method (Definition 6) but solving the following problem:*

$$\underset{\boldsymbol{q}\in[0,1]^r}{\operatorname{argmin}} \sum_{i=1}^{k} (f_{C_i}(q) - p(C_i))^2$$

where C_i are the k CP-nets observed in the profile P.

The $\widetilde{LS}$ method requires exactly m equations, where m is the number of individuals in the profile, while LS would require an exponential number of equations. Thus, in the experiments we use $\widetilde{LS}_O$ and $\widetilde{LS}_I$.

We now define the two notions of score we will use to compare the outcomes of the voting rules. Let F and G be two CP-voting rules, and let T be a set of O-legal CP-profiles which are randomly generated. Given a profile P, we first define two functions $Dom_>(F,G,P)$ and $Dom_<(F,G,P)$ as follows: $Dom_>(F,G,P)$ returns $True$ if the number of CP-nets in P where the outcome $F(P)$ dominates $G(P)$ is greater than the maximum between the number of CP-nets where the outcomes are incomparable and the number of CP-nets where $G(P)$ dominates $F(P)$; it returns $False$ if $Dom_>(G,F,P) = True$; and it returns $None$ otherwise. $Dom_<(F,G,P)$ returns $True$ if $Dom_>(G,F,P) = True$; it returns $False$ if $Dom_>(F,G,P) = True$; and it returns $None$ otherwise.

We use the following notion of pairwise score:

$$PairScore(F,G) = Dom^T_>(F,G) - Dom^T_<(F,G)$$

where $Dom^T_>(F,G)$ and $Dom^T_<(F,G)$ are:

- $Dom^T_>(F,G) = (\#\{P \in T \mid Dom_>(F,G,P) = True\}) \setminus \#T$
- $Dom^T_<(F,G) = (\#\{P \in T \mid Dom_<(F,G,P) = True\}) \setminus \#T$

Observe that this score belongs to the interval $[-1,1]$. Our second scoring function is inspired by Copeland scoring:

$$CopelandScore(F) = \sum_{G\in V\setminus\{F\}} PairScore(F,G)$$

where $V = \{PR_O, PR_I, \widetilde{LS}_O, \widetilde{LS}_I, RAND\}$. Observe that this score belongs to the interval $[-4,4]$.

In the first set of experiments we compute the $CopelandScore$ of the result[2] of each voting rule, varying the number of agents of the input profiles. We considered 300 O-legal profiles with a fixed number of features ($n = 3$), at most $k = 1$ parent per feature, and varied the number of CP-nets between 1 and 50. Figure 1 plots the $CopelandScore$ for all voting rules. It is clear that according to this measure the best voting rule is PR_I and $RAND$ is the worst. In general the number of individuals in the profile does not significantly infulence the $CopelandScore$ of the voting rule.

In the second set of experiments we vary the number of features of the CP-nets in the profile. We generated 300 profiles each with 50 individual CP-nets where the number of features varies between 1 and 5. Each feature has at most $k = \frac{n}{2}$ parents, where n is the number of features. Figure 2 shows the $CopelandScore$ for all voting rules, while Figure 3 shows the $PairScore$ among different pairs of procedures. As in the first experiment, the best voting rule is PR_I and $RAND$ the worst. We observe that the score of $RAND$ increases when the

[2]In all our experiments ties are broken lexicographically.

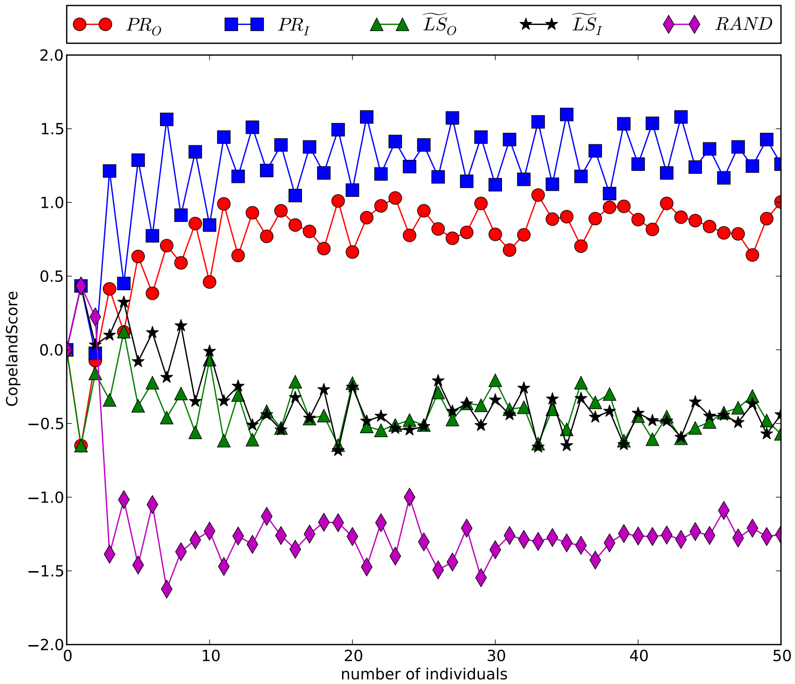

Figure 1: $CopelandScore$, varying the size of the profile.

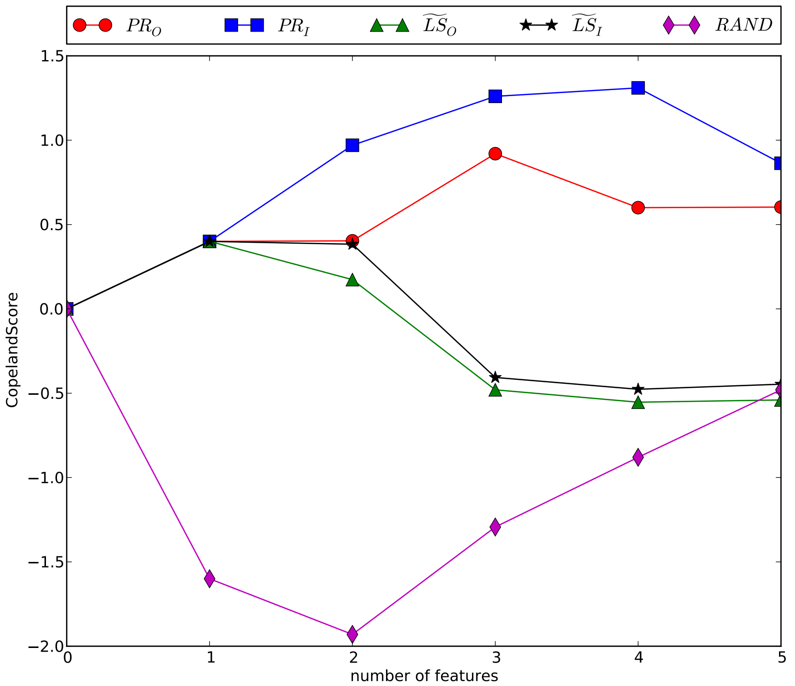

Figure 2: $CopelandScore$, varying the number of features.

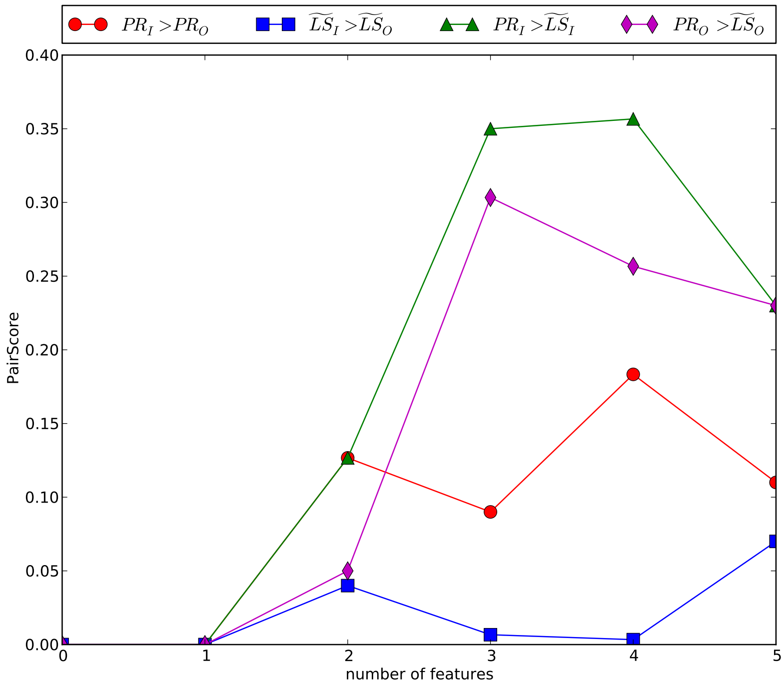

Figure 3: $PairScore$ and number of features.

number of features increases, but $RAND$ is always worse than the other voting rules. One possible explanation is that the number of incomparable outcomes in a CP-net increases when more features are added to the CP-nets. The curves for $PairScore(PR_I, PR_O)$ and $PairScore(\widetilde{LS}_I, \widetilde{LS}_O)$ show that PR_I is always better than PR_O. This same result occurs for $\widetilde{LS}_I$ and $\widetilde{LS}_O$. Hence, considering both the PR and the $\widetilde{LS}$ method, the most probable optimal outcome (O) is worse than the optimal outcome of the most probable induced CP-net (I). Moreover, the curves of $PairScore(PR_O, \widetilde{LS}_O)$ and $PairScore(PR_I, \widetilde{LS}_I)$ show that PR is always better than $\widetilde{LS}$ using either the O or I method.

6 Conclusions and future work

We studied four CP-voting rules which are used to obtain a most preferred outcome from a collection of CP-nets. Our experimental results show that the best method is the PR_I voting rule, which coincides with the result of sequential voting over the input CP-nets. This is not surprising as sequential voting is defined to coincide with the input profile of CP-nets as much as possible. However, our approach obtains this same result through the generation and use of a PCP-net, which can also be used for other collective reasoning tasks, such as dominance queries. In the future we plan to analyse the case of non-binary features, as well as investigating dominance queries in PCP-nets.

Acknowledgements

This work has been partially supported by the strategic project "KIDNEY - Incorporating patients preferences in kidney transplant decision protocols" funded by the University of Padova.

NICTA is funded by the Australian Government through the Department of Communications and the Australian Research Council through the ICT Centre of Excellence Program.

This work is supported by the National Science Foundation, under grant IIS-1107011. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] D. Bigot, H. Fargier, J. Mengin, and B. Zanuttini, 'Probabilistic conditional preference networks', in *Proc. of the 29th International Conference on Uncertainty in Artificial Intelligence (UAI)*, (2013).

[2] S. Bistarelli, U. Montanari, and F. Rossi, 'Semiring-based constraint satisfaction and optimization', *Journal of the ACM (JACM)*, (1997).

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

[4] C. Cornelio, J. Goldsmith, N. Mattei, F. Rossi, and K. B. Venable, 'Updates and uncertainty in CP-nets', in *Proc. of the 26th Australasian Joint Conference on Artificial Intelligence (AUSAI)*, (2013).

[5] C. Domshlak and R.I. Brafman, 'CP-nets: Reasoning and consistency testing', in *Proc. 8th International Conference on Principles of Knowledge Representation and Reasoning (KRR)*, (2002).

[6] C. Domshlak, F. Rossi, K. B. Venable, and T. Walsh, 'Reasoning about soft constraints and conditional preferences: complexity results and approximation techniques', in *Proc. of the 18th International Joint Conference on Artificial Intelligence (IJCAI)*, (2003).

[7] J. Fürnkranz and E. Hüllermeier, *Preference Learning: An Introduction*, Springer, 2010.

[8] J. Goldsmith and U. Junker, 'Preference handling for artificial intelligence', *AI Magazine*, **29**(4), (2009).

[9] J. Goldsmith, J. Lang, M. Truszczynski, and N. Wilson, 'The computational complexity of dominance and consistency in CP-nets', *Journal of Artificial Intelligence Research*, **33**(1), 403–432, (2008).

[10] C. Gonzales and P. Perny, 'GAI networks for utility elicitation.', in *Proc. of the 10th International Conference on Principles of Knowledge Representation and Reasoning (KRR)*, (2004).

[11] J. Lang and L. Xia, 'Sequential composition of voting rules in multi-issue domains', *Mathematical Social Sciences*, **57**(3), 304–324, (2009).

[12] M. Li, Q.B. Vo, and R. Kowalczyk, 'An efficient procedure for collective decision-making with CP-nets', in *Proc. of the 19th European Conference on Artificial Intelligence (ECAI)*, (2010).

[13] M. Li, Q.B. Vo, and R. Kowalczyk, 'Majority-rule-based preference aggregation on multi-attribute domains with CP-nets', in *Proc. of the 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, (2011).

[14] A. Maran, N. Maudet, M.S. Pini, F. Rossi, and K.B. Venable, 'A framework for aggregating influenced CP-nets and its resistance to bribery', in *Proc. of the 27th AAAI Conference on Artificial Intelligence (AAAI)*, (2013).

[15] N. Mattei, M. S. Pini, F. Rossi, and K. B. Venable, 'Bribery in voting with CP-nets', *Annals of Mathematics and Artificial Intelligence*, **68**(1-3), 135–160, (2013).

[16] H. Moulin, *Axioms of Cooperative Decision Making*, Cambridge University Press, 1991.

[17] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, 1988.

[18] F. Rossi, KB Venable, and T. Walsh, 'mCP nets: representing and reasoning with preferences of multiple agents', in *Proc. of the 19th AAAI Conference on Artificial Intelligence (AAAI)*, (2004).

[19] F. Rossi, K.B. Venable, and T. Walsh, *A Short Introduction to Preferences: Between Artificial Intelligence and Social Choice*, Morgan & Claypool Publishers, 2011.

[20] L. Xia, V. Conitzer, and J. Lang, 'Voting on multiattribute domains with cyclic preferential dependencies.', in *Proc. of the 23rd AAAI Conference on Artificial Intelligence (AAAI)*, (2008).

[21] L. Xia, V. Conitzer, and J. Lang, 'Hypercubewise preference aggregation in multi-issue domains', in *Proc. of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*, (2011).

Cristina Cornelio
University of Padua, Italy
Email: cornelio@math.unipd.it

Umberto Grandi
University of Padua, Italy
Email: umberto@math.unipd.it

Judy Goldsmith
University of Kentucky, U.S.
Email: goldsmit@cs.uky.edu

Nicholas Mattei
NICTA and UNSW, Australia
Email: nicholas.mattei@nicta.com.au

Francesca Rossi
University of Padua, Italy
Email: frossi@math.unipd.it

K. Brent Venable
Tulane University and IHMC, U.S
Email: kvenable@tulane.edu