Mathématiques et sciences humaines

Mark Dwyer
Fred R. McMorris
Robert C. Powers

Removal independence and multi-consensus functions

Mathématiques et sciences humaines, tome 148 (1999), p. 31-40
<http://www.numdam.org/item?id=MSH_1999__148__31_0>

Math. Inf. Sci. hum., (37e année, n° 148, *1999*, pp. 31-40)

REMOVAL INDEPENDENCE AND MULTI-CONSENSUS FUNCTIONS

Mark DWYER[1], Fred R. MCMORRIS[2,3], Robert C. POWERS[1]

RÉSUMÉ - Indépendance par suppression et multi-fonctions de consensus.
Vincke et Bouyssou ont montré que, si une procédure d'agrégation de préordres totaux peut retourner plusieurs solutions, alors elle peut satisfaire tous les axiomes du théorème d'Arrow sans être pour autant dictatoriale. Nous étendons cette approche aux hiérarchies utilisées en classification. Dans ce contexte, on obtient des résultats qui peuvent différer de ceux de Vincke et de Bouyssou.

MOTS-CLÉS - *Théorie de la décision, consensus, multi-fonctions de consensus, hiérarchies, théorème d'Arrow*

ABSTRACT - *Work of Vincke and Bouyssou showed that if aggregation procedures on weak orders are allowed to return more than one result, then it might be possible for a procedure to satisfy all the axioms of Arrow's Theorem yet not be dictatorial. This approach is extended from ordered sets to n-trees, which are set-systems used in classifications theory. Results in this context can differ from those of Vincke and Bouyssou.*

KEYWORDS - *Decision theory, consensus, multi-consensus function, hierarchies, Arrow's Theorem*

1. INTRODUCTION

Vincke (1982) and Bouyssou (1992) have given thought provoking interpretations of the classical impossibility result of Arrow (1963). A central point of these papers was that if aggregation procedures on weak orders are allowed to return more than one result and Arrow's axioms are extended to this situation, then "possibility" rather than "impossibility" may occur.

[1]University of Louisville, Louisville, KY 40292, U.S.A.
[2]Illinois Institute of Technology, Chicago, IL 60616-3793, U.S.A.
[3]email: mcmorris@iit.edu

In the present paper we consider a class of hierarchies (called n-trees) rather than the classical weak orders, and develop a program similar to that of Vincke and Bouyssou. An n-tree, viewed as a hierarchical classification, might result after the application of a clustering algorithm to appropriate data. If several of the many available algorithms are used on the same data, the problem of producing an overall summary (consensus) n-tree is clearly one of decision analysis. Since the early 1970's (as a sample, see Adams 1972, Barthélemy et al. 1986, Lapointe and Cucumel 1997, Leclerc and Cucumel 1987) this approach has been used in an area of biology called Numerical Taxonomy, for example. Because of the impact and beauty of Arrow's Theorem in the social sciences, this is often a target theorem in more general contexts such as n-trees, but for n-trees a rich variety of results can occur. Indeed, Barthélemy et al. (1995) have shown that there are at least 9 distinct versions of the independence of irrelevant alternatives axiom of Arrow. As our starting point, we will use the analog of Arrow's Theorem for n-trees proved by Barthélemy et al. (1991) using the key axiom of "removal independence".

2. DEFINITIONS, TERMINOLOGY AND TECHNICAL BACKGROUND

Let S be a finite set with n elements. An *n-tree* on S is a set T of subsets of S satisfying: $S \in T$; $\emptyset \notin T$; $\{x\} \in T$ for all $x \in S$; and $X \cap Y \in \{\emptyset, X, Y\}$ for all $X, Y \in T$. If $X \in T$ with $1 < |X| < n$ then X is called a *nontrivial cluster* of T. $T_\emptyset$ will denote the n-tree with no nontrivial clusters. For X any subset of S, $T|_X$ denotes the n-tree on S whose nontrivial clusters are the nonempty distinct elements of $\{A \cap X\colon A$ is a nontrivial cluster of $T\}$. Another very useful way to realize an n-tree T is by its associated ternary relation (Colonius and Schulze 1981) where $xy|_T z$ if and only if there is an $A \in T$ such that $x, y \in A$, and $z \notin A$. Elements in this ternary relation $|_T$ are called *triads*. The set of all n-trees on S will be denoted by $\mathcal{T}_n$. We let $\mathcal{P}(X)$ denote the set of all subsets of a set X, and $\mathcal{P}_m(X)$ denote the set of all subsets of X with no more than m elements.

A *consensus function* on $\mathcal{T}_n$ is a function $F\colon \mathcal{T}_n^k \to \mathcal{T}_n$, where k is a positive integer and $\mathcal{T}_n^k$ is the k-fold cartesian product. Elements of $\mathcal{T}_n^k$ are called *profiles* and are denoted by $P = (T_1, \ldots, T_k)$, $P' = (T_1', \ldots, T_k')$ and so on. For $X \in \mathcal{P}(S)$ and profile $P = (T_1, \ldots, T_k)$, we let $P|_X = (T_1|_X, \ldots, T_k|_X)$. In the classical theory of Social Choice initiated in Arrow (1963), $\mathcal{T}_n$ is replaced by the set of all weak orders on S, and consensus functions are called *social welfare functions.* As mentioned in the Introduction, there has been a good deal of work studying various forms of consensus functions for tree structures. A nice summary of this work, including more than 90 references, can be found in Leclerc (1998).

Most of the axioms for social welfare functions have analogs. For example, a consensus function F satisfies the *Pareto* condition if, for any profile $P = (T_1, \ldots, T_k)$; $A \in T_i$ for all i, implies that $A \in F(P)$, i.e. $\bigcap T_i \subseteq F(P)$. F is said to be *independent (of irrelevant alternatives)* if, for any two profiles P, P'; $P|_X = P'|_X$ implies that $F(P)|_X = F(P')|_X$. The immediate analog of Arrow's Impossibility Theorem would read: If a consensus function satisfies the Pareto condition and is independent, then it is a dictatorship, where F is a *dictator* consensus function if there exists a j such that for all profiles P; $A \in T_j$ implies $A \in F(P)$. In Barthélemy et al. (1991) an example was given showing that this straightforward translation of Arrow's theorem is not true and that the independence axiom needs to be modified. Indeed, in Barthélemy et al. (1995), several versions of independence that can be defined for $\mathcal{T}_n$ were studied. For a nontrivial subset $X \subset S$ and $T \in \mathcal{T}_n$, let $T|_X - X$ denote the n-tree $T|_X$ without the cluster X. As expected, for a profile P we write $P|_X - X$ for the profile $(T_1|_X - X, \ldots, T_k|_X - X)$. F satisfies *removal independence* when for every $X \subset S$ and profiles P, P'; $P|_X - X = P'|_X - X$ implies that $F(P)|_X - X = F(P')|_X - X$. The main result in Barthélemy et al. (1991) states that if F is a consensus function on $\mathcal{T}_n$ that satisfies the Pareto and Removal Independence conditions, then it is a dictator consensus function. It is this result that we will place in the context developed in Vincke (1982) and Bouyssou (1992) where a consensus function is allowed to be set-valued. Specifically, a function $F : \mathcal{T}_n^k \to \mathcal{P}(\mathcal{T}_n) - \{\emptyset\}$ is a *multi-consensus function* on $\mathcal{T}_n$. A multi-consensus function F is an *m-consensus function* if it maps $F : \mathcal{T}_n^k \to \mathcal{P}_m(\mathcal{T}_n) - \{\emptyset\}$. Thus a 1-consensus function is just an ordinary consensus function.

We now give straightforward generalizations of the axioms of Pareto and Removal Independence to the multi-consensus case labeling them (P) and (RI).

Let F be a multi-consensus function on $\mathcal{T}_n$. Then F satisfies the *Pareto* condition (P) if, for every profile $P = (T_1, \ldots, T_k)$; $A \in T_i$ for every i implies that $A \in T$ for every $T \in F(P)$. F satisfies *removal independence* (RI) if, for every $X \in S$ and profiles P and P'; $P|_X - X = P'|_X - X$ implies that, for each $T \in F(P)$, there exists $T' \in F(P')$ such that $T|_X - X = T'|_X - X$.

As first observed by Vincke (1982) in the context of consensus functions on weak orders, the dictator axiom has two very natural extensions to multi-consensus functions. Obviously this is the case here also. F is a *(strong)dictator* (D) multi-consensus function if there exists a j such that for every profile $P = (T_1, \ldots, T_k)$, there exists a $T \in F(P)$ with $T_j \subseteq T$. F is a *weak dictator* (WD) multi-consensus function if there exists a j such that for every profile $P = (T_1, \ldots, T_k)$, $T_j \subseteq \bigcup_{T \in F(P)} T$. Clearly condition (D) implies condition (WD) and that for 1-consensus functions they both are equivalent to the standard definition of a dictator consensus function.

3. EXAMPLES

In this section we first present an example showing that it is possible to have a multi-consensus function that satisfies (P) and (RI) but not (WD) [and hence also not (D)]. This contrasts with the situation in Vincke (1982) for weak orders, where it is shown that every multi-consensus function on weak orders satisfying Pareto and Independence must satisfy (WD).

EXAMPLE **1** . *For an example we let* $S = \{x_1, x_2, x_3, x_4, x_5\}$, $S_1 = \{x_1, x_2, x_3, x_4\}$, *and* $S_2 = \{x_2, x_3, x_4, x_5\}$. *Set* $M(S_i)$ *to be the set of all 5-trees* T *on* S *where* $T|_{S_i} \cup \{S_i\} = T$ *and* T *has exactly one cluster of size 2 and 3. Note that* $|M(S_i)| = 12$. *Let* $M(S_1) = \{T_1, ..., T_{12}\}$ *and* $M(S_2) = \{T_{13}, ..., T_{24}\}$. *Consider the profile*

$$\widehat{P} = (T_1, ..., T_{24}).$$

Now for $25 \leq i \leq 48$ *set*

$$T_i = T_{i-24} - \{\{x_1, x_2, x_3, x_4\}, \{x_2, x_3, x_4, x_5\}\}.$$

Define $F : (\mathcal{T}_5)^{24} \to \mathcal{P}(\mathcal{T}_5) - \{\emptyset\}$ *by*

$$F(P') = \begin{cases} \{T_{25}, ..., T_{48}\} & \text{if } P' = \widehat{P} \\ \{T'_1, ..., T'_{24}\} & \text{if } P' = (T'_1, ..., T'_{24}) \neq \widehat{P}. \end{cases}$$

Clearly F satisfies (P). To see that F does not satisfy (WD) we need only look at the profile $\widehat{P}$. For every $j \in \{1, ..., 12\}$ and $T \in F(\widehat{P})$, we have $\{x_1, x_2, x_3, x_4\} \in T_j \backslash T$, while for every $j \in \{13, ..., 24\}$ and $T \in F(\widehat{P})$ we have $\{x_2, x_3, x_4, x_5\} \in T_j \backslash T$. To check that F satisfies (RI) requires a tedious but straightforward case by case analysis.

Example 1 used profiles of length 24 for the domain. It would be interesting to find the smallest k for which there exists a multi-consensus function $F : \mathcal{T}_n^k \to \mathcal{P}(\mathcal{T}_n) - \{\emptyset\}$ satisfying (P) and (RI) but not (WD).

In the next section we show that it is impossible to construct a function as in Example 1 if we restrict ourselves to 2-consensus functions. This then raises the question as to whether there is a 2-consensus function satisfying (P) and (RI) but not (D). Our next example shows that this is indeed the case, and parallels the work by Bouyssou (1992) for weak orders.

EXAMPLE **2** . *Let* $S = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ *and* T'_1, T'_2, T'_3, T'_4 *be the 6-trees on* S *defined as follows (only listing the nontrivial clusters):*
$T'_1 = \{\{x_1, x_2, x_3\}, \{x_1, x_2\}, \{x_4, x_5, x_6\}, \{x_4, x_5\}\}$;
$T'_2 = \{\{x_1, x_2, x_3\}, \{x_2, x_3\}, \{x_4, x_5, x_6\}, \{x_5, x_6\}\}$;
$T'_3 = \{\{x_1, x_2, x_3\}, \{x_1, x_2\}, \{x_4, x_5, x_6\}, \{x_5, x_6\}\}$;
$T'_4 = \{\{x_1, x_2, x_3\}, \{x_2, x_3\}, \{x_4, x_5, x_6\}, \{x_4, x_5\}\}$.
Set $\widehat{P} = (T'_1, T'_2)$. *Define* $F : (\mathcal{T}_6)^2 \to \mathcal{P}(\mathcal{T}_6) - \{\emptyset\}$ *by*

$$F(P) = \begin{cases} \{T'_3, T'_4\} & \text{if } P = \widehat{P} \\ \{T_1, T_2\} & \text{if } P = (T_1, T_2) \neq \widehat{P}. \end{cases}$$

Clearly F satisfies (P). To see that F does not satisfy (D) we need only look at the profile P and observe that neither T'_1 nor T'_2 is a subset of T'_3 or T'_4. An even easier case argument shows that F satisfies (RI).

4. A DICTATORSHIP RESULT

The goal of this section is to state carefully and prove our main result which was mentioned in the context of Section 2.

THEOREM **3** . *Let* $|S| = n \geq 4$ *and let* $k \geq 3$. *If* $F : \mathcal{T}_n^k \to \mathcal{P}_2(\mathcal{T}_n)$ *satisfies (P) and (RI), then F satisfies (WD).*

What now follows is a series of technical results, eventually leading to the proof of Theorem 3, although we feel that several of these results are interesting in their own right.

Let $F : \mathcal{T}_n^k \to \mathcal{P}(\mathcal{T}_n) - \{\emptyset\}$ be a multi-consensus function that satisfies (RI) and (P). Since the consensus output $F(P)$ is a set of n-trees we will write $A \in^\vee F(P)$ if $A \in T$ for some $T \in F(P)$. Moreover, we will write $A \in^\wedge F(P)$ if $A \in T$ for every $T \in F(P)$. This notation will be extended to triads. So $ab|c \in^\vee F(P)$ means that $ab|_T c$ for some $T \in F(P)$ and $ab|c \in^\wedge F(P)$ means that $ab|_T c$ for every $T \in F(P)$. If $P|_{\{a,b,c\}} - \{a,b,c\} = P'|_{\{a,b,c\}} - \{a,b,c\}$, then, since F is removal independent, $ab|c \in^\vee F(P)$ implies $ab|c \in^\vee F(P')$. Since this situation occurs frequently we will only write the conclusion: $ab|c \in^\vee F(P)$ implies $ab|c \in^\vee F(P')$. The reader should understand that we are applying the axiom of removal independence and that the hypothesis, $P|_{\{a,b,c\}} - \{a,b,c\} = P'|_{\{a,b,c\}} - \{a,b,c\}$, holds true. Finally, let $P_\emptyset = (T_\emptyset, ..., T_\emptyset)$ and $P_A = (T_A, ..., T_A)$ be notation for constant profiles where $T_A = T_\emptyset \cup \{A\}$.

We will assume throughout this section that $F : \mathcal{T}_n^k \to \mathcal{P}_2(\mathcal{T}_n)$ satisfies (P) and (RI). The first lemma, however, is true even if we replace $\mathcal{P}_2(\mathcal{T}_n)$ by $\mathcal{P}(\mathcal{T}_n) - \{\emptyset\}$.

LEMMA **4** . *If* $ab|c \in^\vee F(P_\emptyset)$ *for some* $a,b,c \in S$, *then* $xy|z \in^\vee F(P_\emptyset)$ *for all* $x,y,z \in S$.

Proof. We are given that $ab|c \in^\vee F(P_\emptyset)$ for some $a,b,c \in S$. Let $z \in S \setminus \{a,b,c\}$. Then $ab|c \in^\vee F(P_\emptyset)$ implies $ab|c \in^\vee F(P_{\{c,z\}})$. Since $\{c,z\} \in^\wedge F(P_{\{c,z\}})$ it follows that $ab|z \in^\vee F(P_{\{c,z\}})$. Now $ab|z \in^\vee F(P_{\{c,z\}})$ implies $ab|z \in^\vee F(P_\emptyset)$. So $ab|z \in^\vee F(P_\emptyset)$ for all $z \in S \setminus \{a,b\}$.

Let $x \in S \setminus \{a,b,c\}$. Then $ab|c \in^\vee F(P_\emptyset)$ implies $ab|c \in^\vee F(P_{\{a,x\}})$. Since $\{a,x\} \in^\wedge F(P_{\{a,x\}})$ it follows that $xb|c \in^\vee F(P_{\{a,x\}})$. Now $xb|c \in^\vee F(P_{\{a,x\}})$ implies $xb|c \in^\vee F(P_\emptyset)$. So $xb|c \in^\vee F(P_\emptyset)$ for all $x \in S \setminus \{b,c\}$.

It now follows that $xy|z \in^\vee F(P_\emptyset)$ for all $x,y,z \in S$. □

PROPOSITION **5** . $F(P_\emptyset) = \{T_\emptyset\}$.

Proof. Assume $F(P_\emptyset) \neq \{T_\emptyset\}$. Then there exists $a,b,c \in S$ such that $ab|c \in^\vee F(P_\emptyset)$. By Lemma 4, $xy|z \in^\vee F(P_\emptyset)$ for all $x,y,z \in S$. In particular, there exist three pairwise incompatible triads in at most two n-trees which is a contradiction. □

The next lemma is a "Co-Pareto condition" on triads.

LEMMA **6** . *Let* $P = (T_1, ..., T_k) \in \mathcal{T}_n^k$. *If* $ab|c \in^{\vee} F(P)$ *for some* $a, b, c \in S$, *then* $ab|c \in T_1 \cup T_2 \cup .... \cup T_k$.

Proof. Assume $ab|c \notin T_1 \cup T_2 \cup .... \cup T_k$. Let

$$\overline{P} = P|_{\{a,b,c\}} - \{a, b, c\} = (\overline{T}_1, ..., \overline{T}_k).$$

Then $ab|c \in^{\vee} F(P)$ implies $ab|c \in^{\vee} F(\overline{P})$. Since $\overline{T_i} \in \{T_\emptyset, T_{\{a,c\}}, T_{\{b,c\}}\}$ for $i = 1, ..., k$, it follows that

$$\overline{P}|_{\{a,b,x\}} - \{a, b, x\} = P_\emptyset|_{\{a,b,x\}} - \{a, b, x\}$$

where $x \in S \setminus \{a, b, c\}$. Since $ab|x \notin^{\vee} F(P_\emptyset)$ it follows that $ab|x \notin^{\vee} F(\overline{P})$. Note that $ab|c \in^{\vee} F(\overline{P})$ and $ab|x \notin^{\vee} F(\overline{P})$ imply $ax|c \in^{\vee} F(\overline{P})$. Let

$$\widetilde{P} = \overline{P}|_{\{a,c,x\}} - \{a, c, x\} = (\widetilde{T}_1, ..., \widetilde{T}_k).$$

Note that $\widetilde{T}_i \in \{T_\emptyset, T_{\{a,c\}}\}$ for $i = 1, ..., k$. Now $ax|c \in^{\vee} F(\overline{P})$ implies $ax|c \in^{\vee} F(\widetilde{P})$. So $ax|b \in^{\vee} F(\widetilde{P})$ or $bx|c \in^{\vee} F(\widetilde{P})$. If $ax|b \in^{\vee} F(\widetilde{P})$, then $ax|b \in^{\vee} F(P_\emptyset)$. If $bx|c \in^{\vee} F(\widetilde{P})$, then $bx|c \in^{\vee} F(P_\emptyset)$. In either case we contradict $F(P_\emptyset) = \{T_\emptyset\}$. □

Let K denote the set $\{1, ..., k\}$. For any $D \subseteq K$ and $x, y \in S$, let

$$P_{\{x,y\};D} = (T_1, ..., T_k)$$

where $T_i = T_{\{x,y\}}$ for $i \in D$ and $T_i = T_\emptyset$ for $i \in K \setminus D$.

LEMMA **7** . *For any* $D \subseteq K$ *and* $x, y \in S$,

$$F(P_{\{x,y\};D}) \subseteq \{T_\emptyset, T_{\{x,y\}}\}.$$

Proof. This result follows immediately from Lemma 6. □

Let $\mathcal{D} = \{D \subseteq K : \{x, y\} \in^{\vee} F(P_{\{x,y\};D})$ for some $x, y \in S\}$. Since F satisfies (P) it follows that $K \in \mathcal{D}$.

LEMMA **8** . *If* M *is a minimal set belonging to* $\mathcal{D}$, *then* $|M| = 1$.

Proof. Assume there exists a minimal set M belonging to $\mathcal{D}$ such that $|M| \geq 2$. Since $M \in \mathcal{D}$ there exist $a, b \in S$ such that $\{a, b\} \in^{\vee} F(P_{\{a,b\};M})$. Let $c \in S \setminus \{a, b\}$ and $j \in M$. Define $P = (T_1, ..., T_k)$ as follows: $T_j = T_{\{a,b\}}$; $T_i = T_{\{a,b,c\}}$ for $i \in M \setminus \{j\}$; $T_i = T_\emptyset$ for $i \in K \setminus M$. Let $d \in S \setminus \{a, b, c\}$ and set $P' = (T'_1, ..., T'_k) = P|_{\{a,b,d\}} - \{a, b, d\}$. Note that $T'_i = T_{\{a,b\}}$ for $i \in M$ and $T'_i = T_\emptyset$ for $i \in K \setminus M$. So $\{a, b\} \in^{\vee} F(P')$. Now $ab|d \in^{\vee} F(P')$ implies that $ab|d \in^{\vee} F(P)$. Since $ax|d \notin T_1 \cup .... \cup T_k$ for all $x \in S \setminus \{a, b, c, d\}$ it follows from Lemma 6 that $ax|d \notin^{\vee} F(P)$ for all $x \in S \setminus \{a, b, c, d\}$. Thus $\{a, b\} \in^{\vee} F(P)$ or $\{a, b, c\} \in^{\vee} F(P)$.

If $\{a, b\} \in^{\vee} F(P)$, then $ab|c \in^{\vee} F(P)$. Now $ab|c \in^{\vee} F(P)$ implies that $ab|c \in^{\vee} F(P_{\{a,b\};\{j\}})$ since $P_{\{a,b\};\{j\}} = P|_{\{a,b,c\}} - \{a, b, c\}$. It follows from the previous lemma that $\{a, b\} \in^{\vee} F(P_{\{a,b\};\{j\}})$. So $\{j\} \in \mathcal{D}$ contrary to the minimality of M.

If $\{a,b,c\} \in^{\vee} F(P)$, then $bc|d \in^{\vee} F(P)$. Now $bc|d \in^{\vee} F(P)$ implies that $bc|d \in^{\vee} F(P_{\{b,c\};M\setminus\{j\}})$ since $P_{\{b,c\};M\setminus\{j\}} = P|_{\{b,c,d\}} - \{b,c,d\}$. It follows from the previous lemma that $\{b,c\} \in^{\vee} F(P_{\{b,c\};M\setminus\{j\}})$. So $M \setminus \{j\} \in \mathcal{D}$. This again contradicts the minimality of M and completes the proof. □

Since $\mathcal{D}$ is nonempty it contains at least one minimal set. For convenience we will assume that $\{1\} \in \mathcal{D}$. So $\{a,b\} \in^{\vee} F(P_{\{a,b\};\{1\}})$ for some $a,b \in S$.

LEMMA 9 . *For any $x,y \in S$, $\{x,y\} \in^{\vee} F(P_{\{x,y\};\{1\}})$.*

Proof. Let $y \in S \setminus \{a,b\}$ and set $P = (T_{\{a,b,y\}}, T_{\{b,y\}}, ..., T_{\{b,y\}})$. If $z \in S \setminus \{a,b,y\}$, then $by|z \in^{\wedge} F(P)$. Now $ab|z \in^{\vee} F(P_{\{a,b\};\{1\}})$ implies that $ab|z \in^{\vee} F(P)$. Putting together $by|z \in^{\wedge} F(P)$ and $ab|z \in^{\vee} F(P)$ we get $ay|z \in^{\vee} F(P)$. The latter implies that $ay|z \in^{\vee} F(P_{\{a,y\};\{1\}})$. So $\{a,y\} \in^{\vee} F(P_{\{a,y\};\{1\}})$ for any $y \in S \setminus \{a,b\}$. If we fix y, then, by symmetry, $\{x,y\} \in^{\vee} F(P_{\{x,y\};\{1\}})$ for any $x \in S$. The result now follows. □

LEMMA 10 . *For any $x,y \in S$ and $D \subseteq K$ with $1 \in D$, $\{x,y\} \in^{\vee} F(P_{\{x,y\};D})$.*

Proof. Let $x,y \in S$ and $D \subseteq K$ with $1 \in D$. Define $P = (T_1, ..., T_k)$ as follows: $T_1 = T_{\{x,y\}}$; $T_i = T_{\{x,y,z\}}$ for $i \in D \setminus \{1\}$; $T_i = T_{\emptyset}$ for $i \in K \setminus D$. Then $xy|z \in^{\vee} F(P)$ since $P|_{\{x,y,z\}} - \{x,y,z\} = P_{\{x,y\};\{1\}}$. Since $yw|z \notin \cup_{i=1}^{k} T_i$ for all $w \in S \setminus \{x,y,z\}$ it follows from Lemma 6 that $yw|z \notin^{\vee} F(P)$ for all $w \in S \setminus \{x,y,z\}$. So $\{x,y\} \in^{\vee} F(P)$. In particular, $xy|w \in^{\vee} F(P)$ where $w \in S \setminus \{x,y,z\}$. Now $xy|w \in^{\vee} F(P)$ implies that $xy|w \in^{\vee} F(P_{\{x,y\};D})$. By Lemma 7, $\{x,y\} \in^{\vee} F(P_{\{x,y\};D})$. □

It follows from the last lemma that for any $D \subseteq K$, with $1 \in D$, $D \in \mathcal{D}$.

LEMMA 11 . *Let $P = (T_1, ..., T_k) \in \mathcal{T}_n^k$. If $ab|c \in T_1$ and $bc|a \notin \cup_{i=1}^{k} T_i$, then $ab|c \in^{\vee} F(P)$.*

Proof. Now $T_1|_{\{a,b,c\}} - \{a,b,c\} = T_{\{a,b\}}$ and $T_i|_{\{a,b,c\}} - \{a,b,c\} \in \{T_{\emptyset}, T_{\{a,b\}}, T_{\{a,c\}}\}$ for $i = 2, ..., k$. Let $\overline{P} = P|_{\{a,b,c\}} - \{a,b,c\}$ and set $\widetilde{P} = \overline{P}_{\{a,b,d\}} - \{a,b,d\}$. Note that $\widetilde{P} = P_{\{a,b\};D}$ with $1 \in D$. By the previous lemma, $\{a,b\} \in^{\vee} F(\widetilde{P})$. So $ab|d \in^{\vee} F(\widetilde{P})$. Now $ab|d \in^{\vee} F(\widetilde{P})$ implies that $ab|d \in^{\vee} F(\overline{P})$. Since $bc|d \notin \cup_{i=1}^{k} \overline{T_i}$ it follows from Lemma 6 that $bc|d \notin^{\vee} F(\overline{P})$. So $ab|c \in^{\vee} F(\overline{P})$. Now $ab|c \in^{\vee} F(\overline{P})$ implies that $ab|c \in^{\vee} F(P)$. □

By symmetry we could replace $bc|a$ in the previous lemma by $ac|b$.

LEMMA 12 . *Let $P = (T_1, ..., T_k) \in \mathcal{T}_n^k$. If $ab|c \in T_1$, then $ab|c \in^{\vee} F(P)$.*

Proof. Let $d \in S \setminus \{a,b,c\}$ and set $P' = (T_1', ..., T_k')$ where $T_1' = T_{\{a,b,d\}}$ and $T_i' = T_i|_{\{a,b,c\}} - \{a,b,c\}$ for $i = 2, ..., k$. If $P'|_{\{a,c,d\}} - \{a,c,d\} = P'' = (T_1'', ..., T_k'')$, then $ad|c \in T_1''$ and $cd|a \notin \cup_{i=1}^{k} T_i''$. By the previous lemma, $ad|c \in^{\vee} F(P'')$. The later implies that $ad|c \in^{\vee} F(P')$. Since $ad|b \notin \cup_{i=1}^{k} T_i'$ it follows from Lemma 6 that $ad|b \notin^{\vee} F(P')$. Now $ad|c \in^{\vee} F(P')$ and $ad|b \notin^{\vee} F(P')$ imply that $ab|c \in^{\vee} F(P')$. Finally, $ab|c \in^{\vee} F(P')$ implies that $ab|c \in^{\vee} F(P)$. □

LEMMA **13** . *The set $\mathcal{D}$ contains at most two minimal sets.*

Proof. By Lemma 8, a minimal set from $\mathcal{D}$ is a singleton set. We know that $\{1\} \in \mathcal{D}$. Assume that $\mathcal{D}$ contains at least two other minimal sets besides $\{1\}$. For convenience we will assume that $\{2\}, \{3\} \in \mathcal{D}$. It follows from the previous lemma that $ab|c, ac|b, bc|a \in^{\vee} F(P)$ where $P = (T_{\{a,b\}}, T_{\{a,c\}}, T_{\{b,c\}}, T_{\emptyset}, ..., T_{\emptyset})$. Since $|F(P)| \leq 2$ we obtain a contradiction. □

We know that $\{1\} \in \mathcal{D}$. By the previous lemma, $\mathcal{D}$ contains at most one other singleton set. This means that there are at least $k-2$ singleton subsets of K that do not belong to $\mathcal{D}$. For convenience we may assume that $\{i\} \notin \mathcal{D}$ for $i = 3, ..., k$. Thus $\{2\}$ may or may not be an element of $\mathcal{D}$.

The next lemma is an improvement of Lemma 6.

LEMMA **14** . *Let $P = (T_1, ..., T_k) \in \mathcal{T}_n^k$. If $ab|c \in^{\vee} F(P)$, then $ab|c \in T_1 \cup T_2$.*

Proof. Now $ab|c \in^{\vee} F(P)$ implies that $ab|c \in^{\vee} F(P|_{\{a,b,c\}} - \{a,b,c\})$. It follows from Lemma 6 that $ab|d \in^{\vee} F(P|_{\{a,b,c\}} - \{a,b,c\})$ where $d \in S \setminus \{a,b,c\}$. Now $ab|d \in^{\vee} F(P|_{\{a,b,c\}} - \{a,b,c\})$ implies that $ab|d \in^{\vee} F(P_{\{a,b\};D})$ where $P_{\{a,b\};D} = [P|_{\{a,b,c\}} - \{a,b,c\}]|_{\{a,b,d\}} - \{a,b,d\}$. It follows from Lemma 7 that $\{a,b\} \in^{\vee} F(P_{\{a,b\};D})$. So $D \in \mathcal{D}$. It follows from the previous lemma that either $1 \in D$ or $2 \in D$. Thus $ab|c \in T_1 \cup T_2$. □

LEMMA **15** . *Let $P = (T_1, ..., T_k) \in \mathcal{T}_n^k$. If $ab|c \in T_1 \cap T_2$, then $ab|c \in^{\wedge} F(P)$.*

Proof. Define $P' = (T'_1, ..., T'_k)$ as follows: $T'_1 = T'_2 = T_{\{a,b\}}$ and $T'_i = T_i|_{\{a,b,c\}} \cup \{\{a,b,c\}\}$ for $i = 3, ..., k$. So $P'|_{\{a,b,c\}} - \{a,b,c\} = P|_{\{a,b,c\}} - \{a,b,c\}$. Note that $ab|d \in T'_i$ for $i = 1, ..., k$. It follows that $ab|d \in^{\wedge} F(P')$. Since $bc|d \notin T'_1 \cup T'_2$ it follows from the previous lemma that $bc|d \notin^{\vee} F(P')$. So $ab|c \in^{\wedge} F(P')$. The later implies that $ab|c \in^{\wedge} F(P)$. □

For the next lemma we will use the following notation. For any $T \in \mathcal{T}_n^k$ and nonempty subset A of S we let

$$\overline{A^T} = \cap\{X \in T : A \subseteq X\}.$$

LEMMA **16** . *Let $P = (T_1, ..., T_k) \in \mathcal{T}_n^k$. Let A be a maximal cluster in T_1 which is not properly contained in a nontrivial cluster in T_2. Then $A \in^{\vee} F(P)$.*

Proof. Let $T \in F(P)$. If $A \in T$, then we're done. If $A \notin T$, then $\overline{A^T} \supset A$. We now assert that $\overline{A^T} = S$.

Assume $\overline{A^T} \subset S$. Let $x \in \overline{A^T} \setminus A$ and $z \in S \setminus \overline{A^T}$. Then $xy|z \in T$ for all $y \in A$. By Lemma 14, $xy|z \in T_1 \cup T_2$ for all $y \in A$. Since A is maximal in T_1, $xy|z \notin T_1$ for all $y \in A$. So $xy|z \in T_2$ for all $y \in A$. This means that T_2 has a cluster that properly contains A contrary to the hypothesis. So $\overline{A^T} = S$.

Let $M = \{X \in T|_A : X \subseteq A$ and X is maximal $\}$. That is, M contains the maximal clusters in $T|_A$ which are subsets of A. Since $\overline{A^T} = S$ it follows that $|M| \geq 2$.

Let $X, X' \in M$. For any $x \in X$, $y \in X'$ and $z \in S \setminus A$, $xy|z \notin T$. (Otherwise, there exists $Y \in T$ such that $x, y \in Y$ and $z \notin Y$. Then $A \cap Y \in T|_A$ intersects two maximal clusters in $T|_A$.) So $xy|z \in T'$ where $T' \in F(P) \setminus \{T\}$. For any $x, x' \in X$ and $z \in S \setminus A$, $xx'|z \in T'$ since $xy|z$ and $x'y|z$ belong to T' where $y \in X'$. It follows that $xy|z \in T'$ for all $x, y \in A$ and $z \in S \setminus A$. So $A \in T'$. Hence $A \in^{\vee} F(P)$. □

LEMMA **17** . *Let* $P = (T_1, ..., T_k) \in \mathcal{T}_n^k$. *Let* A *be a maximal cluster in* T_1 *which is properly contained in a nontrivial cluster* B *in* T_2. *Then* $A \in^{\vee} F(P)$.

Proof. Take B to be the smallest cluster in T_2 that properly contains A and let $P' = P|_B - B = (T_1', ..., T_k')$. Then $A \in T_1'$ is maximal and it is not properly contained in a cluster from T_2'. By the previous lemma, $A \in^{\vee} F(P')$. So $A \in^{\vee} F(P)|_B - B$. So $A = X \cap B$ for some $X \in^{\vee} F(P)$. If there exists $x \in X \setminus A$, then $x \notin B$. Let $y \in A$ and $z \in B \setminus A$. Then $x, y \in X$ and $z \notin X$. So $xy|z \in^{\vee} F(P)$. By Lemma 14, $xy|z \in T_1 \cup T_2$. If $xy|z \in T_1$, then there exists $Z \in T_1$ such that $x, y \in Z$ and $z \notin Z$. So $A \cap Z \neq \emptyset$ and $Z \not\subseteq A$ contrary to the maximality of A. So $xy|z \notin T_1$. If $xy|z \in T_2$, then there exists $Z \in T_2$ such that $x, y \in Z$ and $z \notin Z$. Note that $B \cap Z \neq \emptyset$ and $Z \not\subseteq B$. Also, $B \not\subseteq Z$ since $z \notin Z$. So B and Z can not belong to the same n-tree. In sum, there does not exist $x \in X \setminus A$. So $X \subseteq A$. But $A = X \cap B$ implies that $A \subseteq X$. So $A = X$. Hence $A \in^{\vee} F(P)$. □

It follows from the previous two lemmas that, for any $P = (T_1, ..., T_k)$, $A \in^{\vee} F(P)$ whenever A is a maximal cluster in T_1. We can now complete the proof of Theorem 3.

LEMMA **18** . *Let* $P = (T_1, ..., T_k) \in \mathcal{T}_n^k$. *If* $A \in T_1$, *then* $A \in^{\vee} F(P)$.

Proof. Let $A \in T_1$. If A is a maximal cluster, then, as we just observed, $A \in^{\vee} F(P)$. So we may assume that A is not maximal. Then there exists a chain of $j \geq 2$ (nontrivial) clusters $A_1 \supset A_2 \supset ... \supset A_j$ in T_1 such that A_1 is maximal, A_{i+1} is a proper subset of A_i for $i = 1, ..., j-1$, and $A_j = A$. Further, we want this chain to be maximal in the sense that there does not exist a chain with $j+1$ clusters satisfying the previous properties. From Lemmas 16 and 17, $A_1 \in^{\vee} F(P)$, so our next goal is to show that $A_2 \in^{\vee} F(P)$.

Let $P' = P|_{A_1} - A_1 = (T_1', ..., T_k')$. Note that $A_2 \in T_1'$ is maximal and so $A_2 \in^{\vee} F(P')$ Thus $A_2 \in^{\vee} F(P)|_{A_1} - A_1$. So $A_2 = A_1 \cap X$ for some $X \in^{\vee} F(P)$. If A_1 and X are contained in the same n-tree in the output $F(P)$, then $X = A_2$ and we're done. So we may assume that $A_1 \in T$ and $X \in T'$ where $F(P) = \{T, T'\}$.

We may assume that there exists $x \in X \setminus A_2$. Now for any $y \in A_2$ and $z \in A_1 \setminus A_2$, $xy|z \in^{\vee} F(P) \setminus T_1$. So $xy|z \in T_2$ for all $y \in A_2$ and $z \in A_1 \setminus A_2$. It follows that T_2 contains a cluster Z such that $A_2 \subset Z$ and $Z \cap (A_1 \setminus A_2) = \emptyset$. Note that $ab|c \in T_1 \cap T_2$ for all $a, b \in A_2$ and $c \in A_1 \setminus A_2$. By Lemma 15, $ab|c \in^{\wedge} F(P)$ for all $a, b \in A_2$ and $c \in A_1 \setminus A_2$. In particular, $ab|c \in T$ for all $a, b \in A_2$ and $c \in A_1 \setminus A_2$. Since $A_1 \in T$ it follows that $ab|c \in T$ for all $a, b \in A_2$ and $c \in S \setminus A_1$. In sum, $ab|c \in T$ for all $a, b \in A_2$ and $c \in S \setminus A_2$. Thus $A_2 \in^{\vee} F(P)$.

This argument can be repeated to establish that $A_i \in^{\vee} F(P)$ for $i = 2, ..., j$. In particular, $A = A_j \in^{\vee} F(P)$. □

ACKNOWLEDGEMENTS - This research was supported in part by ONR grants N00014-95-1-0109 and N00014-94-1-0955(AASERT) to F. R. McMorris.

BIBLIOGRAPHY

ADAMS, E.N., "Consensus techniques and the comparison of taxonomic trees", *Systematic Zoology*, 21 (1972), 390-397.

ARROW, K.J., *Social Choice and Individual Values.* 2nd edn., New York, John Wiley (1963).

BARTHÉLEMY, J.P., LECLERC, B. and MONJARDET, B., "On the use of ordered sets in problems of comparison and consensus of classifications", *Journal of Classification*, 3 (1986), 187-224.

BARTHÉLEMY, J.P. and McMORRIS, F.R., "On an independence condition for consensus n-trees", *Applied Mathematics Letters*, 2 (1989), 75-78.

BARTHÉLEMY, J.P., McMORRIS, F.R. and POWERS, R.C., "Independence conditions for consensus n-trees revisited", *Applied Mathematics Letters*, 2 (1991), 43-46.

BARTHÉLEMY, J.P., McMORRIS, F.R. and POWERS, R.C., "Stability conditions for consensus functions defined on n-trees", *Mathematical and Computer Modelling*, 22 (1995), 79-87.

BOUYSSOU, D., "Democracy and efficiency: A note on 'Arrow's theorem is not a surprising result'", *European Journal of Operational Research*, 58 (1992), 427-430.

COLONIUS, H. and SCHULZE, H.H., "Tree structures for proximity data", *The British Journal of Mathematical and Statistical Psychology*, 34 (1981), 167-180.

LAPOINTE, F.J. and CUCUMEL G., "The average consensus procedure: combination of weighted trees containing identical or overlapping sets of objects", *Systematic Zoology*, 46 (1997), 306-312.

LECLERC, B. and CUCUMEL, G., "Consensus en classification: une revue bibliographique", *Mathématiques Informatique et Sciences Humaines*, 100 (1987), 109-128.

LECLERC, B., "Consensus of classifications: the case of trees", in *Advances in Data Science and Classification*, A. Rizzi, M. Vicki and H.-H. Bock, Eds., Springer-Verlag, Berlin (1998), 81-90.

VINCKE, P., "Arrow's Theorem is not a surprising result", *European Journal of Operational Research.*, 10 (1982), 22-25.