Consensus of Trees – Desirable Properties and Computational Methods

SAMPATH KANNAN* and Z. SWEEDYK
*Department of Computer and Information Sciences, University of Pennsylvania, Philadelphia, P. A. 19104, USA.
  email: kannan@cis.upenn.edu

Introduction

The problem of finding a "consensus" object for a collection of mathematical objects has a long and rich history[3, 14, 4]. A consensus function is a mapping $c : \mathcal{R}^* \to \mathcal{R}$, where $\mathcal{R}$ is a finite set and $\mathcal{R}^* = \bigcup_k \mathcal{R}^k$. An element of $\mathcal{R}^*$ is called a *profile* or *input profile*. We will also consider *n-consensus functions* for $\mathcal{R}$, which map $\mathcal{R}^n$ to $\mathcal{R}$. *Generalized consensus functions* for $\mathcal{R}$, which map $\mathcal{R}^*$ to $2^{\mathcal{R}}$ (*i.e.* the power set of $\mathcal{R}$), are also commonly studied but we will not discuss them in this paper.

The following concrete example illustrates the definition above and is also the principal example of interest in this paper. In this example, $\mathcal{R}$ is the set of rooted trees with leaves labelled bijectively from the set $\{1, 2,..., n\}$. A consensus function in this situation is required to take as input a set of rooted trees and produce as output a single rooted tree where all trees have the same leaf labels. The difficulty in this consensus problem (as in all others) is that we need to produce an output that "captures the information" in the input while being restricted to be a tree. In the sequel we will formalize the desirable properties of general consensus functions and then return to the particular consensus problem for rooted trees.

Perhaps the first mathematical formulation of consensus functions is due to the economist, K. Arrow[3]. He considered the special case where $\mathcal{R}$ consists of all total orders on some ground set U. An application, for example, is where U represents possible options for governmental expenditure, an element of $\mathcal{R}$ represents a ranking of these options, a profile represents the respective rankings of the citizens, and a consensus function chooses a ranking of options based on the citizen's wishes. Arrow proved a now-famous impossibility theorem; he defined a reasonable set of desirable properties for *n*-consensus functions and showed that they were impossible to achieve.

Since Arrow's seminal work other authors have extended the axiomatic treatment of consensus to other mathematical objects (sets $\mathcal{R}$) and proved analogues of the impossibility theorem. Barthélemy and Janowitz[4] describe a "meta-theory" of consensus that handles more general mathematical objects such as partial orders and trees. Neumann[14] focuses on consensus functions on trees, desirable properties of these function, and construction methods that satisfy various subsets of these properties. Our focus in this survey paper will also be on consensus functions on rooted trees. Several authors have proposed particular tree consensus methods[1, 5, 6, 11, 15, 16, 8].

In the present paper we begin by reviewing the axiomatic framework for consensus methods in Section 2. Our discussion is closely based on the descriptions in [4], and [14]. In Section 3 we describe several important and widely-used consensus functions for trees. We also study the axiomatic properties of these functions. Finally in Section 4 we describe variants of the consensus problem that have been studied as well as extensions that are worth studying. We also discuss possible applications of consensus techniques in computational biology and other areas.

2. Axiomatic Framework for Consensus Methods

2.1. *Arrow's Theorem*

We begin by describing Arrow's impossibility result. Our description is adapted from [7, pages 523–524].

Arrow's result is restricted to n-consensus functions over $\mathcal{R}$, where $\mathcal{R}$ is the set of total orders over some ground set $U = \{u_1,..., u_m\}$. For this case Arrow posed the following question: is there an n-consensus function $c : \mathcal{R}^n \rightarrow \mathcal{R}$ that satisfies the following (reasonable) properties?

Positive Association of Social and Individual Orderings: For any u, $v \in U$, for any profile $P \in \mathcal{R}^n$, if $u < v$ in $c(P)$ then $u < v$ in $c(P^u)$ where P^u is an arbitrary profile obtained from P by (possibly) increasing the preference for u; *i.e.* if $P = (a_1,..., a_n)$ then $P^u = (a'_1,..., a'_n)$ satisfies the following: if $x < y$ in a_j and $y < x$ in a'_j then $y = u$.

Independence of Irrelevant Alternatives: For any distinct u, v, $w \in U$, for any profile $P \in \mathcal{R}^n$, if $v < w$ in $c(P)$ then $v < w$ in $c(\widetilde{P}^u)$, where $\widetilde{P}^u$ is an arbitrary profile obtained from P by (possibly) changing the preference for u; *i.e.* if $P = (a_1,..., a_n)$ then $P^u = (a'_1,..., a'_n)$ satisfies the following: if $x < y$ in a_j and $y < x$ in a'_j then $x = u$ or $y = u$.

Individual's Sovereignty: For any distinct u, $w \in U$ there is some profile P such that $u < w$ in $c(P)$.

Nondictatorship: There is no j such that $u < v$ in a_j forces $u < v$ in $c(a_1,..., a_n)$ regardless of $a_1,..., a_{j-1}, a_{j+1},..., a_n$.

Theorem 1 (Arrow) *For $n \geq 2$ and $k \geq 3$ there is no n-consensus function that satisfies all of the above properties.*

Arrow's theorem can be generalized to *weak orders*; see [4] for details. A number of papers have considered consensus functions with reduced requirements, while others have proved analogues of this theorem for consensus of other types of mathematical objects.

2.2. *Meta-Consensus Theory*

Arrow's theorem applies to consensus functions for total orders but one might naturally talk about consensus functions for other mathematical objects such as partial orders, rooted trees, partitions of a set, etc. Barthélemy and Janowitz define an axiomatic structure for a broader class of objects. In this section we describe these axioms; the content of this subsection is largely drawn from[4].

The axiomatic properties described by Arrow explicitly rely on the fact that $c\mathcal{R}$ is a set of total orders of the ground set U. An $a \in \mathcal{R}$ contains certain *information*, such as $u < v$, that may or may not be reflected in $c(P)$ when a is part of the profile P. In developing a more generalized framework for consensus, it is necessary to identify the information in the elements of $\mathcal{R}$ that should influence the consensus function. Barthélemy and Janowitz formalize this notion by associating with $\mathcal{R}$ a set $\mathcal{B}$, called the *bricks* of $\mathcal{R}$, and also a mapping $b : \mathcal{R} \rightarrow 2^n$. Thus an element $a \in \mathcal{R}$ has certain *bricks* (or information) associated with it (via b). For example, when $\mathcal{R}$ is the set of total orders on U, we can define the set $\mathcal{B}$ to be $U \times U$ and for $a \in \mathcal{R}$, $b(a) = \{(u, v) \mid u < v$ under the total order $a\}$. Since we only consider consensus functions, as opposed to generalized consensus functions, $b(c(P))$ is well-defined.

Now we are ready to describe various properties of consensus functions as defined.

Stability, Neutrality, and Monotone Neutrality: A consensus function c on $\mathcal{R}$ is

1. **stable** provided that for all $\alpha \in c\mathcal{B}$, for all $k > 0$, for all profiles $P_1 = (a_1,..., a_k)$ and $P_2 = (a'_1,...a'_k)$

   If $\{l : \alpha \in b(a_l)\} = \{l : \alpha \in b(a'_l)\}$ then $\alpha \in b(c(P_1)) \Leftrightarrow \alpha \in b(c(P_2))$.

2. **neutral** provided that for all $\alpha, \beta \in \mathcal{B}$, for all $k > 0$, for all profiles $P_1 = (a_1,..., a_k)$ and $P_2 = (a'_1,...,a'_k)$

   If $\{l : \alpha \in b(a_l)\} = \{l : \beta \in b(a'_l)\}$ then $\alpha \in b(c(P_1)) \Rightarrow \beta \in b(c(P_2))$.

3. **monotone neutral** provided that for all $\alpha, \beta \in \mathcal{B}$, for all $k > 0$, for all profiles $P_1 = (a_1,..., a_k)$ and $P_2 = (a'_1,...,a'_k)$

   If $\{l : \alpha \in b(a_l)\} \subseteq \{l : \beta \in b(a'_l)\}$ then $\alpha \in b(c(P_1)) \Rightarrow \beta \in b(c(P_2))$.

Note that monotone neutrality implies neutrality implies stability.

Paretian: A consensus rule c on $\mathcal{R}$ is **paretian** provided that for all $\alpha \in \mathcal{B}$, for all $k > 0$, for all profiles $P = (a_1,..., a_k)$
If $\alpha \in \bigcap_{j=1}^{k} b(a_j)$ then $\alpha \in b(c(P))$.

Symmetric: A consensus rule c on cR is **symmetric** if and only if $b(c(P))$ is invariant under any permutation of P.

We now state some characterization theorems for consensus rules in terms of the previous properties.

Theorem 2

1. *For weak orders and total orders, the only stable and paretian consensus functions are the dictatorships; that is, for each integer k there exists an integer p_k k such that for each $P = \left(a_{i_1},...a_{i_k}\right) \in \mathcal{R}^k, c(P) = a_{p_k}$.*

2. *For partial orders and partitions of finite sets, the only stable and paretian consensus functions are the oligarchic rules; that is, for each integer k there exists a non-empty subset I_k of indices between 1 and k such that for each $P = \left(a_{i_1},...,a_{i_k}\right) \in \mathcal{R}^k,$*

$c(P) = \bigcap_{l=1}^{k} a_{i_l}$. *(Here we are viewing each partial order as a binary relation or a set of ordered pairs. Intersection of partial orders is therefore the usual set intersection).*

3. *For rooted trees the only neutral k-consensus functions are characterized as follows: There are subsets $D_1,..., D_m$ of $\{1, 2,..., k\}$ such that $D_i \cap D_j \neq \emptyset$ for all i, j and for any profile $P = (a_1,..., a_k)$, $\alpha \in b(c(P))$ if and only if $\{l : \alpha \in b(a_l)\} = D_i$ for some i.*

Part 1 above is yet another version of Arrow's theorem due to Arrow[3]. Part 2 for partial orders is due to Mas-Collel and Sonnenschein[10] and for set partitions is due to Mirkin[13]. Part 3 is due to Neumann[14].

Corollary 1

1. *For weak orders and total orders, there is no consensus function that is stable, paretian and symmetric.*

2. *For partial orders and partitions of finite sets, the only consensus function that is stable, paretian and symmetric is the unanimity rule; that is, $c(P) = \bigcap_{l=1}^{k} a_{i_l}$.*

2.3. *Consensus of Rooted Trees*

In this section we consider the consensus problem when $\mathcal{R}$ consists of rooted trees. The contents of this section are drawn mainly from[14].

Let $U = \{1, 2,..., m\}$. Let $c\mathcal{R}$ be the set of all root (labelled) trees with leaves U. We define the bricks $\mathcal{B}$ associated with $\mathcal{R}$ as the the subsets of U; *i.e.*, $\mathcal{B} = 2^U$. In this special case, the bricks are more commonly referred to as *clusters*. For a tree $T \in \mathcal{R}$, $b(T)$ contains each $V \subseteq U$ such that some rooted subtree of T contains exactly the leaves in V. We will assume, without loss of generality, that internal nodes in the trees of $\mathcal{R}$ have more than one child. Thus $b(T)$ is a one-to-one function. Note that for any T, $b(T)$ contains the singleton sets $\{i\}$, $i = 1,..., m$.

A common monotone neutral consensus function is the "majority rule" where a cluster is present in the consensus tree if and only if it is present in a majority of the trees in the profile. Neumann argues that monotone neutral rules preserve very little information in the situation where there are minor differences between the clusters of the trees in the profile. (See for example, Figure 1, taken from Neumann[14]).

Therefore he proposes a new definition.

Faithful: A consensus function c is *faithful* provided for any positive k, for any profile $P = (T_1,..., T_k)$ and any choice of cluster $X_j \in b(T_i)$, $i = 1, 2,..., k$, $c(P)$ contains a cluster V such that
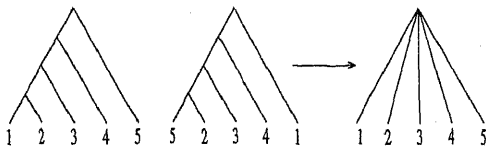

FIG. 1. Example to show "information loss" of majority rule.

$$\bigcap_{i=1}^{k} X_i \subseteq V \subseteq \bigcup_{i=1}^{k} X_i$$

Notice that a faithful consensus function is paretian. Neumann draws a connection between paretian consensus functions satisfying the axiom of independence of irrelevant alternatives and faithful consensus functions. We shall refer to these axioms as Axiom (P) (Paretian), Axiom (I) (Independence of Irrelevant Alternatives) and Axiom (F) (Faithfulness). We first extend the definition of Axiom (I) for rooted trees. For $V \subseteq U$ and $T \in \mathcal{R}$, $T|_V$ refers to the homeomorphic subtree of T obtained by deleting all leaves in $U \setminus V$ and collapsing any degree 2 internal nodes that are created.

Axiom (I) - Independence of Irrelevant Alternatives for Trees: For $V \subset U$, $c(P|_V)$ contains a cluster W if and only if $c(P)$ contains a cluster W' such that $W \subseteq W' \subseteq u/v$.

Theorem 3. If c is a consensus function on rooted trees that satisfies Axioms (P) and (I) then it satisfies Axiom (F).

Neumann points out that there are faithful, paretian consensus functions that do not satisfy Axiom (I). Thus faithfulness is a weakening of the independence of irrelevant alternatives axiom. Neumann leaves open the question of whether a consensus function satisfying Axioms (P) and (I) must in fact be dictatorial.

2.4. A Closer Look at Axiom (I)

In this section we prove some new results that further restrict the class of non-dictatorial consensus functions that satisfy (P) and (I). This goes part way towards answering the open question of Neumann.

Axiom (I) suggests that we can characterize a consensus rule by its behavior on each triple of elements in U. In other words, we consider all possible restrictions of the profile to three elements of U and find the structure of the output tree on these three elements. The information obtained from all of these triples would then characterize the consensus tree.

There are four possible shapes of a tree with three leaves x, y, and z shown in Figure 2.

We will refer to these shapes by the shorthand (xy, z), (x, yz), (xz, y), (xyz) respectively. The first three shapes are called *resolved* while the last is called *unresolved*. Thus, for example, the profile $(xy, z)^k$ refers to a profile of k trees where all of them have the shape that gives x and y a more recent common ancestor than x and z or y and z. For such an input, a pretian consensus function should yield an output that also has the shape (xy, z). We say that a consensus function c is Paretian with respect to resolved triples if for any x, y, $z \in U$, for any profile $P \in \mathcal{R}^k$ such that if $(xy, z)^k = P|_{\{x, y, z\}}$ then $(xy, z) = c(P)|_{\{x, y, z\}}$.
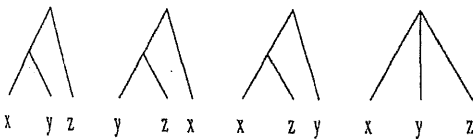

FIG. 2.

The above definitions of axioms (I) and (P) are made with the view that clusters are the bricks of rooted trees. Since we want to talk about trees as constituted of shapes of triples, we need to "translate" these axioms to such a view and the following lemmas do this.

Lemma 1: *If a consensus function satisfies Axiom (I) and (P), then it is paretian with respect to resolved triples.*

Proof: Let k be any positive integer and let $P = (T_1,..., T_k)$ be any profile in $\mathcal{R}^k$. Let $\{x, y, z\}$ be a set of three elements such that $P|_{\{x, y, z\}} = (xy, z)^k$. Since $\{x, y\}$ is a cluster that is present in all the trees in the profile, $\{x, y\}$ must be present as a cluster in $c(P|_{\{x, y, z\}})$. Axiom (I) then says that some cluster B between $\{x, y\}$ and $U \setminus \{z\}$ must be present in $c(P)$. The existence of such a cluster guarantees that the shape of $\{x, y, z\}$ in $c(P)$ is (xy, z). Thus we have shown that if the shape (xy, z) is present in all the trees of the profile, it is also present in the output which is exactly axiom (P) for triples.

Lemma 2. *Let $c: \mathcal{R}^* \to \mathcal{R}$ be a consensus function. Then $c(P|_A)$ is well-defined for any profile P and for any $A \subset U$ such that $|A| = 3$. Furthermore, c is completely determined by its behavior on such restricted profiles.*

Proof: The first part follows from axiom (I). The second statement is true because there is at most one tree consistent with the set of shape specifications for all triples.

Thus, any rule that satisfies axiom (I) is defined by a set of elementary rules characterizing the behavior of the rule on triples. This is the motivation behind the so-called *local consensus rules* proposed by Kannan, Warnow, and Yooseph[8]. However, the implication is only one-way, i.e., any rule that satisfies (I) is characterized in terms of its behavior on triples but not every rule which is characterized in terms of its behavior on triples satisfies axiom (I). Thus local consensus rules afford an interesting alternative to faithful rules and this is discussed in the next section.

The definition of neutrality with respect to clusters is just a special case of the definition of neutrality in general. A rule is said to be non-neutral on triples if there are profiles P and P' in $\mathcal{R}^k$ and two ordered triples of elements $A = (x, y, z)$ and $B = (a, b, c)$ such that under the mapping $x \to a$, $y \to b$, $z \to c$, the profile $P|A$ is transformed into $P'|_B$ but $c(P|_A)$ is not transformed to $c(P'|_B)$.

Lemma 3. *Suppose c is non-neutral on triples. Then there are ordered triples (x, y, z) and (u, y, z) and identical restricted profiles on these triples such that c on these restricted profiles produces different shapes for the respective triples.*

Proof: Consider a graph whose vertex set is all ordered triples. Two vertices (x, y, z) and (u, v, w) are connected if and only if for every restricted profile on (x, y, z) and the corresponding profile for (u, v, w) obtained by replacing p by x, q by y, and r by z, c produces identical shapes for the respective triples. Since c is non-neutral on triples, this graph has more than one component. However, this means that there must be two nodes that differ in only one component that are in different components of this graph. By permuting the order of the triples in these two nodes if necessary we can find triples to differ only in the first component that are not in the same component of the graph.

We are now ready to prove:

Theorem 4. *Any consensus rule c satisfying (I) and (P) that is non-neutral on triples must be dictatorial.*

Proof: By Lemma 3 there are ordered triples (a, c, d) and (b, c, d) for which there are a pair of "identical" profiles P and P' (with k arguments for some $k > 0$) that cause c to produce non-identical shapes. Let U, the set of elements be $\{a, b, c, d\}$ and construct a profile in T_4^K as follows:

Start with $T_1 = T_2 = \cdot = T_k = (ab, c)$. "Inject" d into each tree in order to make the triples (a, c, d) and (b, c, d) have the (identical) profiles P and P' respectively. (Note that this can always be done.)

By the fact that c is paretian on resolved triples, the output tree must have the shape (ab, c) on the triple (a, b, c). Then, no matter where d is injected into the output tree, the shapes of (a, c, d) and (b, c, d) will be identical contradicting the fact that they are supposed to be different.

Neumann proves that any non-dictatorial consensus function satisfying (I) and (P) must be non-neutral on clusters. We show, on the other hand, that any non-dictatorial consensus function satisfying (I) and (P) must be neutral on triples. The combined restrictions on such consensus functions are very strong and further support the conjecture that any consensus function satisfying (I) and (P) must, in fact, be dictatorial.

3. Particular Tree Consensus Rules and Their Properties

In this section we review a number of tree consensus rules from the literature. Primarily these rules view clusters as the bricks making up rooted trees. There are a few exceptions of which we will describe two — the Adams' consensus rule that views objects called *nestings* as the building blocks of trees and the local consensus rules that view triples as the building blocks. The discussion of the last subsection should provide some motivation for why triple-based consensus rules might be important.

3.1. *Cluster-based Consensus Rules*

Before we discuss cluster-based consensus rules it is useful to state some well-known facts about clusters and trees.

Fact 1. *Let T be a rooted tree and $A, B \in C(T)$ (the cluster encoding of T). Then $A \cap B \in \{A, B, \phi\}$.*

Fact 2. *A set of clusters S is said to be **compatible** if and only if there is a tree T such that $C(T) = S$. A set of clusters is compatible if and only if for every pair of sets $A, B \in S$, $A \cap B \in \{A, B, \phi\}$.*

3.1.1. *Neutral Rules*

Strict Consensus: For a profile $P = (T_1, T_2,..., T_k) \in T_n^k$ the strict consensus rule $sc(P)$ produces a tree T such that $C(T) = \bigcap_{i=1}^{k} C(T_i)$.

Notice that $sc(P)$ is defined for every profile P by the facts about compatibility of clusters. The strict consensus rule is paretian and monotone neutral. However, it does not satisfy axiom (I) nor even axiom (F). The critique of the strict consensus function is that it is not very informative — minor differences between the clusters of the trees in the profile can lead to the output tree not containing any clusters at all.

Threshold Consensus: The strict consensus is a special case of a family of consensus rules called threshold consensus rules. These rules are best defined as k-consensus rules for each positive integer k. Let $1 \leq t(k) \leq k$. A $t(k)$-**threshold k-consensus rule** c is defined as follows: For a profile $P = (T_1,..., T_k) \in \mathcal{T}_n^k$, $c(P)$ contains those clusters that are present in at least $t(k)$ of the input trees.

Unfortunately, threshold consensus rules are not necessarily well-defined when $t(k) \leq k/2$ since the set of clusters that must be present in the output may prove to be incompatible. However when $t(k) > k/2$, the threshold consensus always exists. If A and B are two clusters that must be present in the output tree, then each of them occurs in at least $t(k)$ of the trees in the profile and since $t(k) > k/2$ there must be some tree in the input profile that contains both A and B as clusters. Thus A and B are compatible. Since such compatibility holds for each pair of clusters in the output, the set of output clusters are compatible and c is well-defined.

The case of $t(k) = k$ is exactly the strict consensus. Another case of importance is $t(k) = (k + 1)/2$ and this threshold consensus rule is called the *majority consensus* rule.

The case $t(k) = k/2$ leads to some analysis of theoretical interest. For k odd, this is still the majority consensus. For k even, it is possible that the set of clusters that are required to be in the output are not compatible. The so called **median procedure** chooses an output tree that contains all the clusters that occur in strictly greater than half of the trees in the input profile as well as a subset of the clusters that occur in exactly half of the input trees. The problem of deciding whether there is a binary tree of this form is NP-complete[12].

Consider a metric space on $\mathcal{T}_n$ where $d(T_1, T_2)$ is defined to be the cardinality of the symmetric difference of $C(T_1)$ and $C(T_2)$ for $T_1, T_2 \in \mathcal{T}_n$. A consensus problem in this context is, given a profile $(T_1,..., T_k)$, find a tree T that minimizes $\sum_{i=1}^{k} d(T_i, T)$. It is easy to see that any tree produced by the median procedure is optimal for this criterion. Although this metric space seems too coarse to produce informative trees, the idea of posing the consensus problem as an optimization problem in a suitable metric space is attractive and worthy of further exploration.

Further generalization of threshold rules: Neumann[14] provides a nice characterization of neutral rules on rooted trees. Associated with any neutral k-consensus rule c is a set $\mathcal{D}(c)$ of subsets of $\{1, 2,..., k\}$ such that if $X, Y \in \mathcal{D}(c)$ then $X \cap Y \neq \varnothing$. Furthermore, for a given profile $P = (T_1,..., T_k) \in \mathcal{T}_n^k$, a cluster A is in $c(A)$ if and only if $\{i | A \in T_i\} \in \mathcal{D}(c)$. This is possibly the last word on neutral rules.

3.1.2. Faithful Rules

Neumann presents a common framework for deriving all known faithful rules. However, this does not represent a characterization of faithful rules and such a characterization is an interest-

ing open problem. We describe Neumann's framework in somewhat different terms. It is easy to argue following Neumann that all the rules defined by our framework are faithful. It also appears that all of the faithful rules obtainable in Neumann's framework can be obtained by our scheme. One advantage of our scheme is that while Neumann's framework can produce some rules that are not faithful, every rule produced here is faithful and well-defined.

Let T be a rooted tree with leaf set S. An *expansion* of T is obtained by applying the following operation any number of times (possibly 0 times): Choose an edge (u, v) in the current tree and create a new node w in the middle of this edge. That is, make w a child of u and the parent of v. Thus, the expanded tree T' differs from T only in that it (possibly) has some degree 2 internal nodes. Let u be the parent of v in T. Suppose in T' we have a path $u = a_0, a_1,..., a_1 = v$ between u and v. Recall that $N(u)$ and $N(v)$ are the clusters associated with u and v respectively and notice that the conversion of T to T' does not change these clusters. For some I such that $0 \le i < 1$ we treat $N(u)$ as the cluster associated with $a_0, a_1,..., a_i$ and $N(v)$ as the cluster associated with $a_{i,1},..., a_1$. Consider this assignment of clusters to the new nodes part of the expansion process.

For any rooted tree T (including trees with internal nodes of degree 2) the *depth* of a node v (denoted $d(v)$) in T is the length of the path from the root to v. We will also define the depth of the cluster $N(v)$ to be the same as the depth of v. We can now define a family of faithful rules:

Given a profile $P = (T_1, T_2,..., T_k)$ the first step in the rule is an expansion of each of the trees in the profile to create a new profile $P' = (T_1', T_2',..., T_k')$. $C(c(P))$ (the cluster encoding of the consensus tree) is defined to be

$$\bigcup_{d=0}^{\infty} \left\{ \left(\bigcap_{i=1}^{k} X_i \,|\, X_i \in C(T_i), d(X_i) = d \right) \right\}$$

The proof of Neumann[14] can be used to prove that every rule defined in this manner is faithful and well-defined.

3.2. *Nestings and the Adams Consensus*

Given a rooted tree T with leaf set $N = \{1, 2,..., n\}$, for any $A \cup N$ define $lca_T(A)$ to be the least common ancestor in T of the elements in A. Using the notation "$v < u$" to denote the fact that node u lies on the path from node v to the root and $u \ne v$, we can define a partial order on the least common ancestors of subsets of N.

Adams[2] defines a *nesting relation*, R_T on $\mathcal{P}(N) \times \mathcal{P}(N)$ as follows:

$$R_T = \{(A, B) : A \subset B, lca_T(A) < lca_T(B)$$

Based on this relation he defines a consensus rule c as follows: Given a profile P, $c(P) = T$ where T is the unique tree satisfying the following two properties:

$$(\forall_p \in P)(A, B) \in R_p \Rightarrow (A, B) \in R_T (\forall A, B \in C(T), (A, B) \in R_T \Rightarrow (A, B) \in R_p \forall_p \in P$$

Neumann states that the Adams consensus method is faithful. Although, the original paper describing this method[1] gave an exponential time procedure for computing this consensus function, subsequently other authors have shown that it can be constructed very efficiently[9].

Adams argues that nestings serve the role of basic informational blocks or bricks better than rooted triples. However, the more efficient computation of the Adams consensus is achieved by an algorithm that views triples as the bricks suggesting that a triple-based view is at least as powerful as a nesting based view.

3.3. *Triples and Local Consensus*

As explained previously, looking at consensus rules restricted to rooted trees with 3 leaves captures a lot of information about these rules. In particular if the consensus rule satisfies axiom (I), information about its behavior on such trees completely determines the rule.

Local consensus rules are a family of triple-based rules and were defined by Kannan, Warnow, and Yooseph[8]. We will use the term *triple profiles* to indicate profiles where the leaf set of each tree is a set of three elements. A local consensus rule is defined by a set of *strong* and *weak* constraints. Both types of constraints map triple profiles to particular shapes of these triples. Let $A = \{x, y, z\}$ and let C be a constraint that maps a triple profile P_A to a particular shape S on A. Such a constraint is *satisfied* by a consensus rule, c, if given any profile P such that $P|_A = P_A$, $c(P)$ contains the shape S. All local consensus rules require that all strong constraints be satisfied while a maximal number of weak constraints must be satisfied.

The generality of the definition of local consensus rules leads to a rich class of such rules. In particular, several of the rules we have discussed including the strict consensus and the Adams consensus can be viewed as particular local consensus rules or minor variations of such rules[8,9]. The generality also means that while many of the rules are easy to compute in a unified framework, some are NP-hard[8]. In addition, the class of local consensus rules contains neutral rules, faithful rules, and rules that are not well-defined on all profiles. The particular choice of rule can be made by the user based on particular assumptions about the data set at hand.

Of particular interest are the local consensus rules that are not well-defined on all profiles. Traditionally, such partially-defined consensus rules have not been studied at all. However such rules may be attractive where the data set is suspect, and the non-existence of a consensus tree points to problems in the data.

Discussion and Conclusions

In this paper we have surveyed some of the literature on consensus functions with a focus on consensus functions on rooted trees. Such consensus functions are widely used in the construction of evolutionary or phylogenetic trees on data sets. The need for such functions in this application arises for a number of reasons.

- Data from different genes may support different trees for the same set of species.

- Different algorithms with different objective functions may support different trees for the same set of species.

- Different optimal or near-optimal trees may arise from the same data set using the same objective function.

The choice in all these cases is either to find a consensus of all the potential trees or to try to find a subset of the species on which all the trees in the profile are in agreement. The latter approach leads to the maximum agreement subtree problem which also has a large literature not reviewed in this paper.

Axiom (I) does not seem very appropriate in the biological context. A number of authors have argued that in the phylogeny context the shape of a triple (a, b, c) in the consensus tree must be influenced by the position of other species. The neutrality axiom is not perfect either. It is possible that the existence of certain clusters is confirmed by evidence not present in the profile of trees provided to the consensus algorithm. In such cases it may be proper to infer this cluster in the output on scantier evidence than required for some other cluster. Thus the search continues for appropriate axioms for consensus functions of rooted trees, at least for the application to phylogenies.

One possible solution is as follows. The current axioms view the existence or non-existence of clusters as a 0-1 event. However, the process of evolution is continuous and the trees used to model this process can be assigned weights on the edges (corresponding to time) to reflect this continuity. If the profile consists of weighted trees it may be possible to talk about the *degree of support* for a particular cluster in a particular weighted tree. It may then be possible to define appropriate weighted notions of axiom (I) and the neutrality axiom. It is even conceivable that there are reasonable consensus functions that satisfy axiom (I) as well as the neutrality axiom in this context.

Quite rightly, the focus of papers on consensus functions has so far been on the mathematics rather than the computer science. With a reasonable understanding of mathematical limitations it is now worthwhile to look at the computational complexity of consensus functions as done, for example, in[6,12,8].

References

1. ADAMS, E. N. III.   Consensus techniques and the comparison of taxonomic trees. *Syst. Zool.*, 1972, **21**, 390–397.

2. ADAMS, E. N. III.   N-trees as Nestings: Complexity, Similarity and Consensus. *J. Classification*, 1986, **3**, 299–317.

3. ARROW, K. P.,   Social Choice and Individual Values. 2nd ed., John Wiley.

4. BARTHÉLEMY, J. P. AND JANOWITZ, M. F.,   A Formal Theory of Consensus. *SIAM J. Disc. Math.*, 1991, **4**, 305–322.

5. BARTHÉLEMY, J. P. AND MONJARDET, B.,   The median procedure in cluster analysis and social choice theory, *Math. Social Sci.*, 1981, **1**, 235–267.

6. DAY, W. H. E.,   Optimal Algorithms for Comparing Trees with Labeled Leaves. *J. Classification*, 1985, **2**, 7–28.

7. KEENEY, R. L. AND RAIFFA, H.   Decisions with Multiple Objectives — Preferences and Value Tradeoffs, Cambridge University Press, Cambridge 1993.

8. KANNAN, S., WARNOW, T. AND YOOSEPH, S.   Computing the Local Consensus, *Proc. 6th ACM-SIAM Symp. Disc. Algorithms*, 1995.

9. KANNAN, S., WARNOW, T. AND YOOSEPH, S.,   Efficiently Computing the Adams Consensus, to be submitted.

10. MAS-COLLEL, A. AND SONNENSCHEIN, H.,   General Possibility Theorems for Group Decision. *Rev. Economic Stud.*, 1972, **39**, 185–192.

11. MCMORRIS, F. R. AND NEUMANN, D.   Consensus Functions Defined on Trees, *Math. Social Sci.*, 1983, **4**, 131–136.

12. MCMORRIS, F. R. AND STEEL, M.,   The Complexity of the Median Procedure for Binary Trees. Proceedings of the 4th Conference of the International Federation of Classification Societies, Paris, 1993, to be published in the Series "Studies in Classification, Data Analysis, and Knowledge Organization" by Springer-Verlag.

13. MIRKIN, B. G.   *On the Problem of Reconciling Partitions*, in Quantitative Sociology. I, International Perspectives on Mathematical and Statistical Modeling, H.M. Bullock, A. Aganbegian, H.M. Borodkin, R. Boudon, and V. Capecchi, eds., Acad. Press, New York, 441–449, 1975.

14. NEUMANN, D. A.   Faithful Consensus Methods for *n*-Trees, *Math. Biosci.*, 1983, **63**, 271–287.

15. JAMES ROHLF, F.,   Consensus Indices for Computing Classifications, *Math. Biosci.*, 1982, **59**, 131–144.

16. SHAO, K.,   *Consensus Methods in Numerical Taxononmy.*, Ph.D. Dissertation, State University of New York, Stony Brook, New York.