

Biological Aggregation at the Interface Between Theory and Practice

William H. E. Day

Port Maitland, Nova Scotia B0W 2V0, Canada
(e-mail: whday@istar.ca)

Abstract. To understand evolutionary processes better, biologists use aggregation methods to estimate evolutionary relationships; yet properties of the methods are sometimes so imprecisely defined, and their interrelationships so poorly understood, that useful formal results may be difficult to obtain. To address this problem I describe a strategy for modeling aggregation methods and studying their properties. The approach accommodates impossibility results for aggregating rankings, nonhierarchical classifications, hierarchies, and phylogenies. It remains to formulate other relevant models of biological aggregation and to characterize methods for solving biological problems of agreement and synthesis.

Keywords: Aggregation, Agreement, Axiom, Consensus, Impossibility, Synthesis.

The axiomatic method is, strictly speaking, nothing but this art of drawing up texts whose formalization is straightforward in principle. As such it is not a new invention; but its systematic use as an instrument of discovery is one of the original features of contemporary mathematics. — Nicolas Bourbaki [Bourbaki, 1968, p. 8]

1 Aggregation problems in biology

Mathematical models of aggregation have long been used in systematic or evolutionary biology [Day and McMorris, 2003]. Given a sequence of trees that estimate phylogenetic relationships among species, for example, one wants to develop methods to synthesize these trees into a single large phylogenetic supertree [Steel *et al.*, 2000, Wilkinson *et al.*, 2004]. If estimating supertrees is an exemplar of biological aggregation, the following questions pertain.

What is a supertree? Most biologists understand biological supertrees and their use to estimate evolutionary history, while mathematicians wish to know no more about supertrees than is necessary to construct appropriate models. Here I assume that supertrees and other relevant objects are defined so that their essential features are expressible by sets of elementary structures.

What is a supertree rule? I describe an abstract framework in which aggregation can be modeled and concepts investigated. Given a profile (sequence) of objects: an agreement rule returns an object having only features in common agreement among the profile's objects, a consensus rule returns

an object best representing the profile's objects, and a synthesis rule returns a composite of the profile's objects.

What biologically relevant properties should supertree rules exhibit? Properties of aggregation constrain the formal model so as to improve its capability to approximate a biological process [Wilkinson *et al.*, 2004]. Here I ignore issues of practicality and computational complexity since analyses of time and space resources are best left to computer scientists. I am particularly interested in axioms that if satisfied by an aggregation rule may increase one's confidence in the biological relevance of that rule's results.

Can supertree rules exhibit particular sets of desirable properties? What properties do known supertree rules exhibit? Since little has yet been done to answer such questions, here I simply mention some biologically interesting impossibility results and some open problems concerning the axiomatics of biological aggregation rules.

2 Aggregation models

For 30 years researchers have striven to develop consensus rules for biological applications. Although inappropriate for investigating agreement or synthesis, consensus rules are a useful point of reference. There is a set of voters. Each voter votes by specifying an object. A consensus rule C accepts a profile of objects and returns a unique consensus object that in some sense best represents the profile. A simple model requires that C be a function $C : \mathcal{X}^k \rightarrow \mathcal{X}$, where \mathcal{X} is a set of *objects* such as those in table 1 and \mathcal{X}^k is

\mathcal{X} is the set of all ...
\mathcal{E} Nonhierarchical classifications or partitions of S , each being a set of nonempty classes or subsets of S that are pairwise disjoint and that include every element of S .
\mathcal{O} Rankings of S , each being a partition of S the classes of which are linearly ordered from most to least preferred.
\mathcal{H} Rooted trees, each with n leaves, such that the root vertex has degree at least 2, every other interior vertex has degree at least 3, and every leaf is labeled with a distinct element of S .
\mathcal{P} Unrooted trees, each with n leaves, such that no vertex has degree 2 and every leaf is labeled with a distinct element of S .

Table 1. Objects defined in terms of S , $n = |S| > 0$

the set of all profiles (k -tuples) of \mathcal{X} . C is further specified by a set K of k *indices* to name the voters, a set S of n *labels* or species names with which to describe objects, *encoding* functions to represent objects in meaningful ways, and *reduction* functions to reveal the structure of objects. Since the concepts

of object, index, label, encoding, and reduction appear naturally in models of agreement, consensus, or synthesis, the incremental strategy in table 2 can be used to study them.

1. Begin with the basic concepts of object, index, and label.
2. Design an aggregation model. Specify axioms and use them to prove things.
3. Add a concept of encoding. Specify axioms and use them to prove things.
4. Add a concept of reduction. Specify axioms and use them to prove things.
5. Add other relevant concepts. Specify axioms and use them to prove things.
6. Repeat steps 2–5 for related aggregation models.

Table 2. Strategy to investigate aggregation

To specify models let $K = \{1, \dots, k\}$, $S = \{s_1, \dots, s_n\}$, and for every $X \subseteq S$ let \mathcal{X}_X be a set of objects defined in terms of each and every label of X . For every $X \subseteq S$ let $\mathcal{X}_{[X]} = \bigcup_{Y \subseteq X} \mathcal{X}_Y$ where $\mathcal{X}_X \subseteq \mathcal{X}_{[X]}$. An object of \mathcal{X}_X has the label set X , but an object of $\mathcal{X}_{[X]}$ may have as its label set any subset of X ; thus \mathcal{H}_S is the set of hierarchies having exactly n leaf labels and $\mathcal{H}_{[S]}$ is the set of hierarchies having at most n leaf labels. \mathcal{X} , K , S then yield

$$C : \mathcal{X}_S^k \longrightarrow \mathcal{X}_{[S]}, \text{ a model of agreement,} \quad (1)$$

$$C : \mathcal{X}_S^k \longrightarrow \mathcal{X}_S, \text{ a model of consensus, and} \quad (2)$$

$$C : \mathcal{X}_{[S]}^k \longrightarrow \mathcal{X}_S, \text{ a model of synthesis.} \quad (3)$$

The essence of consensus is that profile objects and consensus result have the same label set S . Agreement (1) is more general than consensus (2) since, although the domains are identical, an agreement result's label set may be a proper subset of S . Synthesis (3) is more general than consensus (2) since, although the codomains are identical, the label set of any synthesis profile object may be a proper subset of S .

Models (1)–(3) can be modified into rules that accept profiles of varying lengths or return more than one aggregated object. Let $\mathcal{X}^* = \bigcup_{k \geq 1} \mathcal{X}^k$ be the set of all profiles of finite positive length and call any aggregation rule with domain \mathcal{X}^* a *complete* rule. Let $2^{\mathcal{X}} \setminus \{\emptyset\}$ be the set of all nonempty subsets of \mathcal{X} and call any aggregation rule with codomain $2^{\mathcal{X}} \setminus \{\emptyset\}$ a *multiaggregation* rule. Thus a complete multisynthesis rule is modeled by a function

$$C : \mathcal{X}_{[S]}^* \longrightarrow 2^{\mathcal{X}_S} \setminus \{\emptyset\}. \quad (4)$$

3 Aggregation axioms

The axioms in table 3 address issues of impartiality (whether rules favor one label or index more than another), delegation of authority (whether determining outcomes resides with proper subsets of indices), optimality (how rules

Line	Axiom	Concept	Reference
(5)	<i>S-Ntr</i> : Neutrality of Labels	Label	[Steel <i>et al.</i> , 2000]
(6)	<i>Sym</i> : Symmetry of Indices	Index	[Steel <i>et al.</i> , 2000]
(7)	<i>Prj</i> : Projection	"	[Barthélemy <i>et al.</i> , 1991]
(8)	<i>Dct</i> : Weak Dictatorship	Encoding	[Arrow, 1963]
(9)	<i>Olg</i> : Oligarchy	"	[Mirkin, 1975]
(10)	<i>PO</i> : Pareto Optimality	"	[Arrow, 1963]
(11)	<i>SO</i> : Strong Optimality	"	[Steel <i>et al.</i> , 2000]
(12)	<i>RC</i> : Reduction Consistency	Reduction	[Wilkinson <i>et al.</i> , 2004]
(13)	<i>Ind</i> : Independence	"	[Arrow, 1963]
(14)	<i>Dsp</i> : Display	"	[Steel <i>et al.</i> , 2000]
(15)	<i>Agr</i> : Agreement	"	[Day and McMorris, 2003]

Table 3. Axioms and their related concepts

behave in the presence of object agreement), contexture (how rules respond to changes in structure or composition), and resolvability (how rules preserve relationships between objects).

To motivate axioms I give informal prose descriptions, but to specify axioms I define them using the logical symbols for negation (\neg), conjunction (\wedge), disjunction (\vee), implication (\implies), equivalence (\iff), and universal (\forall) and existential (\exists) quantification. Since axioms may apply to more than one model, in their definitions I assume as little as possible about the model's form: unless stated otherwise let it be a function $C : \mathcal{X}^k \longrightarrow \mathcal{Y}$ for $\mathcal{X}, \mathcal{Y} \subseteq \mathcal{X}_{[S]}$ where $\mathcal{X} \subseteq \mathcal{Y}$ and/or $\mathcal{Y} \subseteq \mathcal{X}$. Such an axiom might be relevant to any of the models (1)–(3).

Let f and g be functions, let x be an element in g 's domain, and let $g(x)$ be in f 's domain; then I reduce notational clutter by writing fgx instead of $f(g(x))$. Thus $C\sigma P = C(\sigma(P))$ as in (6). To specify objects of $P \in \mathcal{X}^k$ let $P = (T_1, \dots, T_k)$ as in (7).

3.1 Basic axioms

Three axioms treat objects as atomic and indivisible.

***S-Ntr*: Neutrality of Labels.** Let $\phi : S \longrightarrow S$ be a function that permutes the labels in S . Let $\phi : \mathcal{X} \longrightarrow \mathcal{X}$ permute the labels of an object: for every $T \in \mathcal{X}$, ϕT is the object obtained by using ϕ to permute the labels of T . Let $\phi : \mathcal{X}^k \longrightarrow \mathcal{X}^k$ permute the labels in every object of a profile: $(\forall P \in \mathcal{X}^k)(\phi P = (\phi T_1, \dots, \phi T_k))$. Although three functions are named ϕ , context shows which ϕ pertains. *Motivation:* If a profile P is described by a data matrix in which each row represents a label then the aggregation of P should be insensitive to the relative order of P 's rows (labels). Put another way, for every P and every S -permutation ϕ , the relabeling by ϕ of the aggregation of P should equal the aggregation of the profile in which P 's objects are relabeled by ϕ . *Axiom:*

$$(\forall P \in \mathcal{X}^k)(\forall S\text{-permutations } \phi)(C\phi P = \phi CP) \quad (5)$$

Sym: Symmetry of Indices. Let $\sigma : K \rightarrow K$ be a function that permutes the indices in $K = \{1, \dots, k\}$. Now σ can permute objects in a profile by permuting the indices of the objects in that profile, i.e., let $\sigma : \mathcal{X}^k \rightarrow \mathcal{X}^k$ be a function such that $(\forall P \in \mathcal{X}^k)(\sigma P = (T_{\sigma_1}, \dots, T_{\sigma_k}))$. Although two functions are named σ , context shows which σ pertains. *Motivation:* If a profile P is described by a data matrix in which each column represents an object then the aggregation of P should be insensitive to the relative order of P 's columns (objects). Put another way, for every P and every K -permutation σ , the aggregation of P should equal the aggregation of the profile in which the positions of P 's objects are permuted by σ . *Axiom:*

$$(\forall P \in \mathcal{X}^k)(\forall K\text{-permutations } \sigma)(CP = C\sigma P) \quad (6)$$

Prj: Projection (Strong Dictatorship). *Motivation:* In nontrivial oligarchies and dictatorships the power to control aggregation is shared unequally by voters. In a strong dictatorship, for some index j and every profile P , the aggregation of P is the j^{th} object of P . *Axiom:*

$$(\exists j \in K)(\forall P \in \mathcal{X}^k)(CP = T_j) \quad (7)$$

Thus if P is a point in a k -dimensional space then C projects P onto a single dimension.

3.2 Axioms using object encodings

Whereas in section 3.1 objects were atomic and indivisible, now let every object $T \in \mathcal{X}$ be a set of elementary structures that are defined using the labels of S . Specifically let E_S be a complete set of elementary structures defined using the labels of S , and let r denote an encoding by which every $T \in \mathcal{X}$ is a well-defined subset of E_S . The encodings in table 4 may be familiar to biologists; the axioms in sections 3.2 and 3.3 assume that such an encoding has been applied.

\mathcal{X}	r	Using r , $T \in \mathcal{X}$ is a	Reference
\mathcal{O}	w	weak order	[Arrow, 1963]
\mathcal{E}	e	equivalence relation	[Mirkin, 1975]
\mathcal{H}	c	set of clusters	[Margush and McMorris, 1981]
\mathcal{H}	t	set of triads	[Colonus and Schulze, 1981]
\mathcal{H}	n	set of nestings	[Adams III, 1986]
\mathcal{P}	s	set of splits	[Buneman, 1971]
\mathcal{P}	q	set of quartets	[Colonus and Schulze, 1981]

Table 4. Encodings (r) to represent objects as sets of elementary structures

Dct: Weak Dictatorship. *Motivation:* In a weak dictatorship, for some index j and every profile P , the aggregation of P contains as a subset the j^{th} object of P . *Axiom:*

$$(\exists j \in K)(\forall P \in \mathcal{X}^k)(T_j \subseteq CP) \quad (8)$$

Olg: Oligarchy. *Motivation:* Oligarchy extends the strong dictatorial concept to forms of aggregation in which ruling power is shared by a set of individuals: for some index set V and every profile P , the aggregation of P is the set intersection of the objects of P that are specified by V . *Axiom:*

$$(\exists V \subseteq K)(\forall P \in \mathcal{X}^k)(\cap_{j \in V} T_j = CP) \quad (9)$$

An oligarchy of one individual is a strong dictator; an oligarchy of k individuals is a form of rule by unanimity.

PO: Pareto Optimality. *Motivation:* Proposals may require for adoption the unanimous support of a society's members. For every profile P the aggregation of P should include those elementary structures (i.e., proposals) that are in every object of P (i.e., are supported by every member). *Axiom:*

$$(\forall P \in \mathcal{X}^k)(\cap_{i \in K} T_i \subseteq CP) \quad (10)$$

SO: Strong Optimality. *Motivation:* Instead of requiring unanimous support, proposals may be adopted if they are unopposed by conflicting proposals. With hierarchies represented by sets of triads (see figure 1), for every

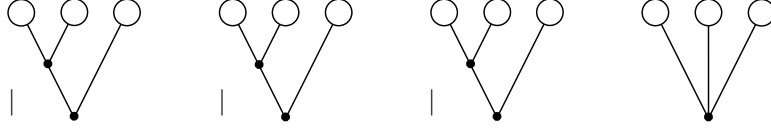


Fig. 1. Triads for representing hierarchies.

profile P and every three labels x, y, z , if $xy|z$ is in some object of P but neither $xz|y$ nor $yz|x$ are in P 's objects, then $xy|z$ should be in the aggregation of P . *Axiom:*

$$\begin{aligned} &(\forall P \in \mathcal{H}^k)(\forall x, y, z \in S)(\\ &[(\exists j \in K)(xy|z \in T_j) \wedge (\forall i \in K)(xz|y \notin T_i \wedge yz|x \notin T_i)] \\ &\implies xy|z \in CP) \end{aligned} \quad (11)$$

With phylogenies represented by sets of quartets the axiom becomes

$$\begin{aligned} &(\forall P \in \mathcal{P}^k)(\forall w, x, y, z \in S)(\\ &[(\exists j \in K)(wx|yz \in T_j) \wedge (\forall i \in K)(wy|xz \notin T_i \wedge wz|xy \notin T_i)] \\ &\implies wx|yz \in CP) \end{aligned}$$

In such cases SO is stronger than PO in the sense that $SO \implies PO$.

3.3 Axioms using object encodings and reductions

Let an encoding (as in table 4) be applied so that every object in \mathcal{X} is represented by a set of elementary structures. Like an X-ray machine, reduction penetrates the surfaces of such objects to reveal hidden structure. For every $X \subseteq S$, let the function $\xi_X : \mathcal{X}_{[S]} \rightarrow \mathcal{X}_{[X]}$ reduce objects on subsets of S to objects on subsets of X : for every $T \in \mathcal{X}_{[S]}$, $\xi_X T$ is the object obtained by suppressing in T the structure associated with $S \setminus X$. Thus if T were a graph G with vertex set S then $\xi_X T$ might be the subgraph of G that is induced by X . Also let $\xi_X : \mathcal{X}_{[S]}^k \rightarrow \mathcal{X}_{[X]}^k$ reduce profiles rather than single objects: for every $X \subseteq S$ then $(\forall P \in \mathcal{X}_{[S]}^k)(\xi_X P = (\xi_X T_1, \dots, \xi_X T_k))$. Although two functions are named ξ_X , context shows which ξ_X pertains.

RC: Reduction Consistency. *Motivation:* The order in which reduction and aggregation functions are applied ought not to matter: for every profile P and subset X of labels, the aggregation of the reduction of P to X by ξ should equal the reduction to X by ξ of the aggregation of P . *Axiom:*

$$(\forall P \in \mathcal{X}_{[S]}^k)(\forall X \subseteq S)(C\xi_X P = \xi_X CP) \quad (12)$$

Ind: Independence. Profiles $P, P' \in \mathcal{X}_{[S]}^k$ are called equal, i.e., $P = P'$, if and only if $(\forall i \in K)(T_i = T'_i)$. *Motivation:* For all profiles P and P' and every subset X , if P and P' are equal when reduced to X by ξ then the aggregations of P and P' must be equal when reduced to X by ξ . *Axiom:*

$$(\forall P, P' \in \mathcal{X}_{[S]}^k)(\forall X \subseteq S)(\xi_X P = \xi_X P' \implies \xi_X CP = \xi_X CP') \quad (13)$$

Ind, which Arrow [Arrow, 1963] called independence of irrelevant alternatives, imposes on aggregation rules a form of context insensitivity. *Ind* is weaker than reduction consistency in the sense that $RC \implies Ind$. Some researchers have confounded *Ind* with *RC*, a result perhaps unsurprising since [Arrow, 1963, pp. 26–27] motivated his definition of *Ind* with examples of both *RC* and *Ind* ([McLean, 1995, p. 108]).

Dsp: Display. An object T is said to resolve an object T' if T' can be obtained from T by a sequence of simplifying elementary transformations. For partitions an elementary transformation forms the union of two classes of the previous partition; for rankings those two classes must be adjacent in the previous linear order. For hierarchies or phylogenies an elementary transformation contracts an interior edge by identifying its endpoints and deleting the resulting loop. For every $T, T' \in \mathcal{X}_{[S]}$, T is said to display T' if, for some $X \subseteq S$, $\xi_X T = T'$ or $\xi_X T$ resolves T' . An object also can display a profile: for every object T and profile P , T is said to display P if T displays T_i for every $i \in K$. *Motivation:* For every profile P if some object displays P then the aggregation of P should display P . *Axiom:*

$$(\forall P \in \mathcal{X}_{[S]}^k)[(\exists T \in \mathcal{X}_{[S]})(T \text{ displays } P) \implies (CP \text{ displays } P)] \quad (14)$$

Agr: Agreement. For every $P \in \mathcal{X}_{[S]}^k$ let $D(P)$ be the set of all non-trivial objects (i.e., those having nontrivial elementary structures) that are displayed by every $T_i \in P$. *Motivation:* For every profile P if some nontrivial object is displayed by P then the aggregation of P should be nontrivial and should be displayed by P . *Axiom:*

$$(\forall P \in \mathcal{X}_{[S]}^k)(D(P) \neq \emptyset \implies CP \in D(P)) \quad (15)$$

4 Problems at the interface

Researchers have used the axiomatic approach to prove impossibility results (table 5) for models (1)–(3) of aggregation, a result being called impossible if an undesirable property, e.g., *Dct* or *Olg*, follows from desirable properties, e.g., *Ind* and *PO*. Consequently the following questions may be relevant when assessing the efficacy of such models for biological aggregation.

Model	\mathcal{X}	r	Impossibility Result	Reference
consensus	\mathcal{O}	w	$Ind \wedge PO \implies Dct$	[Arrow, 1963, p. 97]
consensus	\mathcal{E}	e	$Ind \wedge PO \iff Olg$	[Mirkin, 1975, p. 446]
consensus	\mathcal{H}	c	$Ind \wedge PO \iff Prj$	[Barthélemy <i>et al.</i> , 1992, p. 63]
agr,con,syn	\mathcal{H}	t	$SO \iff \neg SO$	[Steel <i>et al.</i> , 2000, p. 367]
consensus	\mathcal{P}	q	$Ind \wedge PO \iff Prj$	[McMorris and Powers, 1993, p. 54]
consensus	\mathcal{P}	q	$S-Ntr \wedge PO \implies \neg Sym$	[Steel <i>et al.</i> , 2000, p. 366]
synthesis	\mathcal{P}	q	$S-Ntr \wedge Dsp \implies \neg Sym$	[Steel <i>et al.</i> , 2000, p. 364]
agreement	\mathcal{P}	q	$S-Ntr \implies \neg Agr$	[Day and McMorris, 2003, p. 108]

Table 5. Impossibility results for aggregation models (1)–(3), the representation of objects in \mathcal{X} being determined by the encoding r . For many other such results see [Day and McMorris, 2003] and references therein.

Are we using the right axioms? [Wilkinson *et al.*, 2004] argue that elusive properties of input trees involving tree size, tree shape, or the location or size of conflicting structures may adversely bias methods to build supertrees. How should such properties be included in formal studies of aggregation models? Even devising adequate definitions of such properties may be problematic. Would some particular encoding provide a natural setting in which such properties could be investigated? Since the strategy in table 2 is simplistic, using it to guide the analysis of such models may be ineffective or infeasible.

Engaging but specific problems exist. How strong is *S*-neutrality? For agreement, consensus, or synthesis rules on phylogenies, characterize those rules that satisfy *S-Ntr*. Since independence (*Ind*) imposes a strong concept of context insensitivity, could it be replaced by biologically useful concepts of context sensitivity?

Are we solving the right problems? Impossibility results encourage mathematicians to explore the boundary areas between feasible and infeasible aggregation rules. Biologists might be more excited by axiomatic characterizations of actual or ideal aggregation rules for biologically relevant objects.

Are we using the right models? Since much is known about complete multiconsensus median rules [McMorris *et al.*, 2000, McMorris *et al.*, 2003], do such axiomatic results generalize to the complete multisynthesis model (4)? Do the concepts of agreement, consensus, synthesis, multiaggregation, and completeness yield useful aggregation models for biological applications? Although an extensive literature on biologically relevant consensus rules exists [Day and McMorris, 2003], axiomatic investigations of agreement and synthesis rules are just beginning [Steel *et al.*, 2000] and show great promise.

Have we the right perspective? If objects are complex structures, one can exploit that complexity to study the interrelationships among objects; but if objects are taken to be atomic and indivisible, one must use object interrelationships to study the basic properties of sets of objects. Would it be useful to investigate agreement or synthesis models from an order theoretic perspective, as was done for consensus models by [Monjardet, 1990] and [Leclerc and Monjardet, 1995]?

For some readers this paper may have little of biological interest since its biological relevance emerges only by specifying undefined terms, e.g., object, and open-ended concepts, e.g., encoding or reduction. If then the axioms or models prove to be inappropriate for analyzing biological problems, perhaps biologists and mathematicians would collaborate to refine the approach.

Acknowledgements. My views have been influenced by [Steel *et al.*, 2000], who investigate supertrees axiomatically and whose clarity of exposition is exemplary, and by [Wilkinson *et al.*, 2004], who examine biologically desirable properties of synthesis rules for supertrees. I am indebted to Robert C. Powers for criticizing preliminary drafts of this paper.

References

- [Adams III, 1986]E. N. Adams III. n -Trees as nestings: Complexity, similarity, and consensus. *Journal of Classification*, 3(2):299–317, 1986.
- [Arrow, 1963]K. J. Arrow. *Social Choice and Individual Values*. Number 12 in Cowles Foundation for Research in Economics at Yale University: Monographs. Wiley, New York, second edition, 1963. Reprinted by Yale University Press (New Haven) in 1978.
- [Barthélemy *et al.*, 1991]J. P. Barthélemy, F. R. McMorris, and R. C. Powers. Independence conditions for consensus n -trees revisited. *Applied Mathematics Letters*, 4(5):43–46, 1991.
- [Barthélemy *et al.*, 1992]J. P. Barthélemy, F. R. McMorris, and R. C. Powers. Dictatorial consensus functions on n -trees. *Mathematical Social Sciences*, 25(1):59–64, December 1992.

- [Bourbaki, 1968]N. Bourbaki. *Theory of Sets*. Number 1 in Elements of Mathematics. Addison-Wesley, Reading, Massachusetts, 1968.
- [Buneman, 1971]P. Buneman. The recovery of trees from measures of dissimilarity. In F. R. Hodson, D. G. Kendall, and P. Tautu, editors, *Mathematics in the Archaeological and Historical Sciences*, pages 387–395. Edinburgh University Press, Edinburgh, 1971.
- [Colonus and Schulze, 1981]H. Colonus and H. H. Schulze. Tree structures for proximity data. *British Journal of Mathematical & Statistical Psychology*, 34(2):167–180, 1981.
- [Day and McMorris, 2003]W. H. E. Day and F. R. McMorris. *Axiomatic Consensus Theory in Group Choice and Biomathematics*. Number 29 in SIAM Frontiers in Applied Mathematics. Society for Industrial and Applied Mathematics, Philadelphia, 2003.
- [Leclerc and Monjardet, 1995]B. Leclerc and B. Monjardet. Latticial theory of consensus. In W. A. Barnett, H. Moulin, M. Salles, and N. J. Schofield, editors, *Social Choice, Welfare, and Ethics: Proceedings of the Eighth International Symposium in Economic Theory and Econometrics*, International Symposia in Economic Theory and Econometrics, chapter 6, pages 145–160. Cambridge University Press, Cambridge, 1995.
- [Margush and McMorris, 1981]T. Margush and F. R. McMorris. Consensus n -trees. *Bulletin of Mathematical Biology*, 43(2):239–244, 1981.
- [McLean, 1995]I. McLean. Independence of irrelevant alternatives before Arrow. *Mathematical Social Sciences*, 30(2):107–126, October 1995.
- [McMorris and Neumann, 1983]F. R. McMorris and D. A. Neumann. Consensus functions defined on trees. *Mathematical Social Sciences*, 4(2):131–136, April 1983.
- [McMorris and Powers, 1993]F. R. McMorris and R. C. Powers. Consensus functions on trees that satisfy an independence axiom. *Discrete Applied Mathematics*, 47(1):47–55, 16 November 1993.
- [McMorris *et al.*, 2000]F. R. McMorris, H. M. Mulder, and R. C. Powers. The median function on median graphs and semilattices. *Discrete Applied Mathematics*, 101(1–3):221–230, 15 April 2000.
- [McMorris *et al.*, 2003]F. R. McMorris, H. M. Mulder, and R. C. Powers. The median function on distributive semilattices. *Discrete Applied Mathematics*, 127(2):319–324, 2003.
- [Mirkin, 1975]B. G. Mirkin. On the problem of reconciling partitions. In H. M. Blalock, A. Aganbegian, F. M. Borodkin, R. Boudon, and V. Capecchi, editors, *Quantitative Sociology: International Perspectives on Mathematical and Statistical Modelling*, Quantitative Studies in Social Relations, chapter 15, pages 441–449. Academic Press, New York, 1975.
- [Monjardet, 1990]B. Monjardet. Arrowian characterizations of latticial federation consensus functions. *Mathematical Social Sciences*, 20(1):51–71, August 1990.
- [Steel *et al.*, 2000]M. A. Steel, A. W. M. Dress, and S. Böcker. Simple but fundamental limitations on supertree and consensus tree methods. *Systematic Biology*, 49(2):363–368, June 2000.
- [Wilkinson *et al.*, 2004]M. Wilkinson, J. L. Thorley, D. Pisani, F. J. Lapointe, and J. O. McInerney. Some desiderata for liberal supertrees. In O. R. P. Bininda-Emonds, editor, *Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life*, number 4 in Computational Biology, chapter 10, pages 227–246. Kluwer Academic Publishers, Boston, Massachusetts, 2004.