Journal of Mathematical Biology manuscript No.
(will be inserted by the editor)

Cladistic hypotheses as degree of equivalence relational structures: implications for three-item statements

Valentin Rineau · Stéphane Prin

Received: date / Accepted: date

Abstract Three-item statements, as minimal informative rooted binary phylogenetic trees on three items, are the minimal units of cladistic information. Their importance for phylogenetic reconstruction, consensus and supertree methods relies on both (i) the fact that any cladistic tree can always be decomposed into a set of three-item statements, and (ii) the possibility, at least under some conditions, to build a new cladistic tree by combining all or part of the three-item statements deduced from several prior cladistic trees. In order to formalise such procedures, several k -adic rules of inference, i.e., rules that allow us to deduce at least one new three-item statement from exactly k other ones, have been identified. However, no axiomatic background has been proposed, and it remains unknown if a particular k -adic rule of inference can be reduced to more basic rules. In order to solve this problem, we propose here to define three-item statements in terms of degree of equivalence relations. Given both the axiomatic definition of the latter and their strong connection to hierarchical classifications, we establish a list of the most basic properties for three-item statements. With such an approach, we show that it is possible to combine five three-item statements from basic rules although they are not combinable only from dyadic rules. Such a result suggests that all higher k -adic rules are well reducible to a finite set of simpler rules.

Keywords mathematical phylogenetics · hierarchical classifications · evolutionary trees · inference rules · cladistic analysis · systematics

Mathematics Subject Classification (2010) 03B10 · 03E20 · 92F05

These authors contributed equally to this work

Rineau Valentin

Center for Theoretical Study, Charles University & Czech Academy of Sciences, Jilská 1,
110 00 Praha 1, Czech Republic.

E-mail: valentin.rineau@gmail.com

Prin Stéphane

Romainville, France

1 Introduction

The increase in biological data as well as in processing capacities has made it possible to generate more and more phylogenetic trees on various taxonomic groups. New problems have arisen with the increasing number of trees concerning the manipulation and amalgamation of several trees into a unified hypothesis in order to make their multitude intelligible, and, with them, new phylogenetic methods of tree reconstruction and combination (Bininda-Emonds 2004). Supertree-like methods, including species tree estimation under the multispecies coalescent (e.g., ASTRAL; Zhang et al. 2018) have become central for phylogeny estimation in the past ten years. As a result, a growing number of algorithms have been designed to solve supertree-like problems (Nguyen et al. 2012; Sevillya et al. 2016; Fleischauer and Böcker 2017). The development of these methods and practices implies more than ever to have a relevant formal description of phylogenetic trees as mathematical structures on which biologists can rely.

Among the first to investigate the combinatorial properties of hierarchical n -trees and/or unrooted n -trees, Dobson (1974) introduced a symbolism for quartets (i.e., minimal informative unrooted subtrees). Using the four-points conditions, she then deduced as corollaries two inference rules for quartets. After that, Dekker (1986) proposed a more systematic approach to inference rules for quartets. She introduced the concept of k -adic rules of inference, i.e., rules such that given k quartets, at least one other follows. She then proved the existence of not only two dyadic rules of inference, but also several higher-order k -adic rules of inference irreducible to the former. Dekker eventually conjectured from this result that such higher-order rules actually are infinite. Subsequently, Bryant and Steel (1995) transposed Dekker's work to three-item statements (minimal informative rooted subtrees; 3is hereafter) and proved Dekker's conjecture for both 3is and quartets. Note that their approach, although not axiomatic, is formal and abstract. Using a non-formal and example-based approach, Wilkinson et al. (2004) finally inferred three dyadic rules for 3is from the two dyadic rules for quartets proposed by Dekker and explicitly interpreted these rules as dependency rules. These works show that understanding the problem of amalgaming larger trees from combination of subtrees as a problem of k -adic rules lead to an infinite complexity, that is to say, there are an infinite quantity of higher-order rules of combination irreducible to a finite set of dyadic rules of combination.

Interestingly, none of the previous work is axiomatic, i.e., none of them consists to start from a consistent, necessary, and sufficient primitive rules from which all the other rules could be deduced. As a result, it is hard to determine whether or not the higher rules really are irreducible to more basic rules. Especially, one cannot determine whether the fact that higher order rules are irreducible to dyadic rules means that these higher-order rules are irreducible to more basic rules. In other words, it is hard to determine whether dyadic rules are the only possible most basic rules. In this paper, we propose to address this issue for the specific case of rooted trees and 3is with a different framework.

On the one hand we propose to define 3is in terms of specific ternary relations on a set called “degree of equivalence relations”. The idea here is to rely on both the axiomatic properties of these relations and their bijective connexion to hierarchical classifications. On the other hand, we propose to illustrate the fecundity of such an approach by considering an example of five compatible but pairwise non-combinable 3is (Wilkinson pers. com.).

Usually, 3is are encountered in cladistic phylogenetics. Indeed, the latter consists of establishing the degree of kinship relationship *sensu* Prin (2016), also called “phylogenetic relationship” (Hennig 1965; Nelson 1994), “relationship of common ancestry” (Nelson 1970, 1972, 1973), or “cladistic relationship” (Camin and Sokal 1965; Hull 1970) between the n taxonomic units¹ (TU hereafter) of a given set. The degree of kinship relationship is a ternary biological relationship of kinship that consists to say that given three biological objects of the same kind (organisms or taxa), two of them are closely related to each other than either is to the third. But any 3is is precisely an assertion that, given three distinct objects of a same kind (e.g., TU), states that two of them are closely related to each other than either is to the third. Hence the possibility to deduce a consistent set of 3is from any cladistic tree, i.e., any hierarchical n -tree formalising a given cladistic hypothesis. By accepting that degree of equivalence relations formalise the degree of kinship relationship, that suggests the possibility to formally define 3is in terms of degree of equivalence relations.

In order to achieve this aim, this paper is structured in the following seven steps: (i) we provide an axiomatic definition of degree of equivalence relations, (ii) we recall the axiomatic definition of hierarchical classifications, (iii) we address the mathematical equivalence between degree of equivalence relations and hierarchical classifications, (iv) we draw the consequences on the mathematical structures of cladistic hypotheses, (v) we define 3is in terms of degree of equivalence relations and investigate some consequences of it, (vi) we generalise some basic rules seen before to higher-order rules, and then (vii) we apply these consequences for a specific set of compatible but pairwise non-combinable 3is.

2 Degree of equivalence relation

Let $\langle A, T \rangle$ be an ordered pair. In this case, $\langle A, T \rangle$ is also a *degree of equivalence relational structure* if and only if (a) A is a non-empty set and (b) T is a *degree*

¹ Given a phylogenetic/cluster analysis, a TU is any taxon defined prior to this analysis and taken as an object by this analysis. Usually, phylogenetic analyses seek to establish kinship relationships of this or that kind (e.g., the ancestor-descendant relationship or the degree of kinship relationship) between several mutually disjoint TU. These taxa can be contrasted to hypothetical taxa (HT), i.e., taxa highlighted by phylogenetic reconstruction (e.g., inclusive clades in cladistic phylogenetics).

of equivalence relation on A , i.e., a ternary relation on A ($T \subseteq A^3$) such that for all $w, x, y, z \in A$:²

1. $\langle x, y, z \rangle \in T \rightarrow [\langle y, x, z \rangle \in T \wedge \langle x, z, y \rangle \notin T]$ (1-2 symmetry & 2-3 asymmetry)
2. $x \neq y \rightarrow \langle x, x, y \rangle \in T$ (1-2 conditional reflexivity)
3. $[\langle x, y, w \rangle \in T \wedge \langle y, z, w \rangle \in T] \rightarrow \langle x, z, w \rangle \in T$ (1-2 transitivity)
4. $\langle w, x, z \rangle \in T \rightarrow [\langle w, x, y \rangle \in T \vee \langle w, y, z \rangle \in T]$ (2-3 negative transitivity)

Interestingly, these axiomatic properties imply several derived properties. Below are some of them (here too, for all $w, x, y, z \in A$):

5. $\langle x, y, z \rangle \in T \rightarrow \langle z, y, x \rangle \notin T$ (1-3 asymmetry)
6. $\langle x, y, x \rangle \notin T \wedge \langle x, y, y \rangle \notin T$ (1-3 & 2-3 anti-reflexivity)
7. $\langle x, x, y \rangle \in T \leftrightarrow x \neq y$ (1-2 relative identity)
8. $\langle x, x, x \rangle \notin T$ (anti-reflexivity)
9. $[\langle w, x, y \rangle \in T \wedge \langle w, y, z \rangle \in T] \rightarrow [\langle w, x, z \rangle \in T \wedge \langle x, y, z \rangle \in T]$ (2-3 enhanced transitivity)
10. $[\langle x, w, y \rangle \in T \wedge \langle y, w, z \rangle \in T] \rightarrow [\langle x, w, z \rangle \in T \wedge \langle x, y, z \rangle \in T]$ (1-3 enhanced transitivity)
11. $[\langle w, x, y \rangle \in T \wedge \langle x, y, z \rangle \in T] \rightarrow [\langle w, x, z \rangle \in T \wedge \langle w, y, z \rangle \in T]$ (crossed transitivity)
12. $[\langle w, x, y \rangle \in T \wedge \langle y, z, w \rangle \in T] \rightarrow [\langle w, x, z \rangle \in T \wedge \langle y, z, x \rangle \in T]$ (pair separation)

Hereafter, the set A will be considered only as finite. Now, let $\mathcal{D}_A$ be the set of all degree of equivalence relations on A . Then let $T \in \mathcal{D}_A$ and let $\langle x, y, z \rangle \in T$. In this case, $\langle x, y, z \rangle$ is trivial (or non-informative) if $x = y$ and non-trivial (or informative) if $x \neq y$. It follows a single trivial degree of equivalence relation T_{A0} on A such that

$$T_{A0} =_{df} \{\langle x, x, y \rangle \in A^3 \mid x \neq y\}$$

i.e.,

$$T_{A0} =_{df} \{\langle x, y, z \rangle \in A^3 \mid x = y \wedge x \neq z\}$$

Given the properties above, it is not hard to see that $T \supseteq T_{A0}$ for any $T \in \mathcal{D}_A$. In corollary, any $T \in \mathcal{D}_A$ is non-trivial if and only if $T \supset T_{A0}$. Let us now define the set

$$I_T =_{df} \{\langle x, y, z \rangle \in T \mid x \neq y\}$$

of all non-trivial 3-tuples of T . In this case, one has (i) $I_T \notin \mathcal{D}_A$, (ii) $T = T_{A0} \cup I_T$, and (iii) $T_{A0} \cap I_T = \emptyset$.

To conclude, note that the symmetry-based properties of T imply that there are always three incompatible pairs of non-trivial 3-tuples for all distinct $x, y, z \in A$: $\{\langle x, y, z \rangle, \langle y, x, z \rangle\}$, $\{\langle x, z, y \rangle, \langle z, x, y \rangle\}$, and $\{\langle y, z, x \rangle, \langle z, y, x \rangle\}$.

² For previous axiomatic approaches on these relations, see Colonius and Schulze (1981) and McMorris and Powers (2003).

In other words, either $\{\langle x, y, z \rangle, \langle y, x, z \rangle\} \subseteq T$ and $\{\langle x, z, y \rangle, \langle z, x, y \rangle, \langle y, z, x \rangle, \langle z, y, x \rangle\} \cap T = \emptyset$, or $\{\langle x, z, y \rangle, \langle z, x, y \rangle\} \subseteq T$ and $\{\langle x, y, z \rangle, \langle y, x, z \rangle, \langle y, z, x \rangle, \langle z, y, x \rangle\} \cap T = \emptyset$, or $\{\langle y, z, x \rangle, \langle z, y, x \rangle\} \subseteq T$ and $\{\langle x, y, z \rangle, \langle y, x, z \rangle, \langle x, z, y \rangle, \langle z, x, y \rangle\} \cap T = \emptyset$, or $\{\langle x, y, z \rangle, \langle y, x, z \rangle, \langle x, z, y \rangle, \langle z, x, y \rangle\} \cap T = \emptyset$.

3 Hierarchical classification

Let $\langle A, H \rangle$ be an ordered pair. In this case, $\langle A, H \rangle$ is a finite *hierarchical classificatory structure* if and only if (a) A is a finite non-empty set, and (b) H is a *hierarchical classification*³ of A , i.e., a subset of $\wp(A) \setminus \{\emptyset\}$ such that

13. $H \supseteq \{X \subseteq A \mid X = A \vee |X| = 1\}$
14. $(\forall X, Y \in H) (X \cap Y = \emptyset \vee X \subseteq Y \vee Y \subseteq X)$

Hereafter, the elements of H will be called its “clusters”. Now, let $\mathcal{H}_A$ be the set of all possible hierarchical classifications of A . Then let $H \in \mathcal{H}_A$ and $X \in H$. In this case, X is trivial (or non-informative) if $X = A$ and/or $|X| = 1$, and non-trivial (or informative) if $2 \leq |X| < |A|$. Any non-empty set A thus implies one and only one trivial hierarchical classification H_{A0} of A such that

$$H_{A0} =_{df} \{X \subseteq A \mid X = A \vee |X| = 1\}$$

i.e.,

$$H_{A0} =_{df} \{A\} \cup \{\{x\} \mid x \in A\}$$

Here too, the properties above imply $H \supseteq H_{A0}$ for any $H \in \mathcal{H}_A$. In corollary, any $H \in \mathcal{H}_A$ is non-trivial if and only if $H \supset H_{A0}$. Then considers the set

$$I_H =_{df} \{X \in H \mid 2 \leq |X| < |A|\}$$

of all non-trivial clusters of H . In this case, (i) $I_H \notin \mathcal{H}_A$, (ii) $H = H_{A0} \cup I_H$, and (iii) $H_{A0} \cap I_H = \emptyset$.

To conclude, note that property 14 implies an incompatibility between the possible non-trivial clusters of H analogous to the incompatibility between the non-trivial 3-tuples of T . Indeed, suppose for all distinct $x, y, z \in A$ and any $X \in H$ that there is at least one $Y \in H$ such that $\{x, y\} \subseteq Y$ and $z \notin Y$. In this case, $X \supset Y$ if either $\{x, z\} \subseteq X$ or $\{y, z\} \subseteq X$.

³ Although some authors also use this term in biology (e.g., Ridley 2004: 472), these classifications are primarily known in mathematics as total hierarchies (Benzecri 1972: 30), hierarchies (Leclerc 1985a, 1985b; Vach 1994: 61; McMorris and Powers 2003: 42), hierarchies of parts (Barthélemy and Guénoche 1991: 22-23), or even n -trees for the specific case where A is finite (Margush and McMorris 1981: 240; Adams 1986: 304; McMorris and Powers 2003: 42). A related (and more general) concept is that of tree of subsets sensu Estabrook and McMorris (1980: 369). See Leclerc (1985a, 1985b) for an extensive study of them.

4 Degree of equivalence relation and hierarchical classification

Remarkably, there is a mathematical equivalence between (the) degree of equivalence relations (on A) and (the) hierarchical classifications (of A) (Coloni and Schulze 1981; McMorris and Powers 2003). This equivalence can be described as follows (Prin 2016: 2nd appendix). (1) Let $T \in \mathcal{D}_A$ and $\{x, y\} \subseteq A$. Then, define the T -relative *degree of equivalence class* of x and y as the set $[x \& y]_T$ of the elements z of A such that x and y are not T -equivalent compared to z , i.e.,

$$[x \& y]_T =_{df} \{z \in A \mid \langle x, y, z \rangle \notin T\}$$

In this case, the T -relative *propinquity set* of A , i.e., the set

$$A//T =_{df} \{[x \& y]_T \mid \{x, y\} \subseteq A\}$$

i.e.,

$$A//T =_{df} \{X \subseteq A \mid (\exists x, y \in A) X = [x \& y]_T\}$$

of all possible degree of equivalence classes induced by T is necessarily a hierarchical classification of A . Hence, there is a function $f : \mathcal{D}_A \rightarrow \mathcal{H}_A$ such that $f(T) = A//T$ for any $T \in \mathcal{D}_A$. (2) Let $H \in \mathcal{H}_A$. In this case, the ternary relation T_H on A such that for any $\{x, y, z\}$, $\langle x, y, z \rangle \in T_H$ if and only if (i) $\langle x, y, z \rangle \in A^3$ and (ii) there is at least one $X \in H$ containing x and y but not z , i.e.,

$$(\forall x, y, z) [\langle x, y, z \rangle \in T_H \leftrightarrow [\langle x, y, z \rangle \in A^3 \wedge (\exists X \in H) (\{x, y\} \subseteq X \wedge z \notin X)]]$$

i.e.,

$$T_H = \{\langle x, y, z \rangle \in A^3 \mid (\exists X \in H) (\{x, y\} \subseteq X \wedge z \notin X)\}$$

is necessarily a degree of equivalence relation on A . Thus, there is a function $g : \mathcal{H}_A \rightarrow \mathcal{D}_A$ such that $g(H) = T_H$ for any $H \in \mathcal{H}_A$. (3) Let $\langle T, H \rangle \in \mathcal{D}_A \times \mathcal{H}_A$. In this case, not only $g(f(T)) = T$, but also $f(g(H)) = H$. Hence, f and g are both injective (which implies both $|\mathcal{D}_A| \leq |\mathcal{H}_A|$ and $|\mathcal{H}_A| \leq |\mathcal{D}_A|$) and so, bijective too (since $|\mathcal{D}_A| = |\mathcal{H}_A|$). (4) Finally, let $\langle T, H \rangle \in \mathcal{D}_A \times \mathcal{H}_A$ again. The preceding implying that $f(T) = H$ if and only if $g(H) = T$, it follows that f and g are converses to each other. Degree of equivalence relations and hierarchical classifications are therefore just two ways for saying that given three objects, either two of them are equivalent compared to the third, or there is no possibility to separate one of them from the other two. In particular, T_{A0} and H_{A0} correspond to each other, i.e., $f(T_{A0}) = H_{A0}$ and $g(H_{A0}) = T_{A0}$.

5 Consequences on cladistic hypotheses

Following Prin (2016), we adopt the view according to which any phylogenetic hypothesis can basically be formalised as a finite simple k -ary relational structure, i.e., an ordered pair

$$\langle A, R_\rho(A) \rangle$$

such that: (a) A is a finite non-empty set of mutually disjoint TU; and (b) $R_\rho(A)$ is a k -ary relation on A , i.e., a subset of A^k such as

- a) $R_\rho(A) = \{\langle x_1, x_2, \dots, x_k \rangle \in A^k \mid \rho(x_1, x_2, \dots, x_k)\}$, and
- b) the k -adic propositional function “ $\rho(x_1, x_2, \dots, x_k)$ ” corresponds to the complete predicative formulation of a given k -ary concept of kinship in our language (e.g., “ x is an ancestor of y ” for the ancestor-descendant relationship).⁴

Such a general schema can easily be applied to cladistic phylogenetic hypotheses, i.e., phylogenetic hypotheses expressing the degree of kinship relationship between the TU of a given set. For this, one needs both (i) a predicative formulation of the degree of kinship relationship, and (ii) to know its formal properties. For the predicative formulation, let us recall that this relationship of kinship is a ternary one that consists to say that given three biological entities (organisms and their parts or taxa and their properties), two of them are more closely related than either is to the third (or, if you prefer, that two of them are closely related in respect to the third). Thus, possible propositional functions can be either “ x is more closely related to y than to z ” or “ x is (closely) related to y compared to z ”, abbreviated “ $\delta(x, y, z)$ ” hereafter. As for its formal properties, it has been argued in Prin (2016) that this relationship mathematically corresponds to degree of equivalence relations. Consequently, any cladistic phylogenetic hypothesis D can basically be formalised as a finite non-trivial degree of equivalence relational structure, i.e., an ordered pair

$$\langle A, T_\delta(A) \rangle$$

such that: (a) A is a finite non-empty set of at least three disjoint TU (i.e., $|A| \geq 3$); and (b) $T_\delta(A)$ is a non-trivial degree of equivalence relation on A , a subset of A^3 such as:

- a) $T_\delta(A) = \{\langle x, y, z \rangle \in A^3 \mid \delta(x, y, z)\}$, and
- b) there are at least three distinct $x, y, z \in A$ such that either $\{\langle x, y, z \rangle, \langle y, x, z \rangle\} \subseteq T$ and $\{\langle x, z, y \rangle, \langle z, x, y \rangle, \langle y, z, x \rangle, \langle z, y, x \rangle\} \cap T = \emptyset$, or $\{\langle x, z, y \rangle, \langle z, x, y \rangle\} \subseteq T$ and $\{\langle x, y, z \rangle, \langle y, x, z \rangle, \langle y, z, x \rangle, \langle z, y, x \rangle\} \cap T = \emptyset$, or $\{\langle y, z, x \rangle, \langle z, y, x \rangle\} \subseteq T$ and $\{\langle x, y, z \rangle, \langle y, x, z \rangle, \langle x, z, y \rangle, \langle z, x, y \rangle\} \cap T = \emptyset$.

From this first formalisation follow at least three other ones. Indeed, if D is formalizable as $\langle A, T_\delta(A) \rangle$, it is then formalizable as a finite hierarchical classificatory structure $\langle A, H_\delta(A) \rangle$ such that $H_\delta(A) = A / T_\delta(A)$. But if true, that also means that we can formalise D as a finite (binary) relational structure of strict order $\langle H_\delta(A), R_\supset[H_\delta(A)] \rangle$ such that

⁴ This formalisation should be understood as a minimal one. Indeed, phylogenetic hypotheses can be more complex by the incorporation of other kinds of information (e.g., branch lengths for tree-like and network-like phylogenetic hypotheses). However, even such complex phylogenetic hypotheses always involve a finite non-empty set of (biological) objects and a phylogenetic relationship (of a given arity) on them.

$$R_{\supset}[H_{\delta}(A)] = \{\langle X, Y \rangle \in H_{\delta}(A) \times H_{\delta}(A) \mid X \supset Y\}$$

And finally, the latter implies the possibility to formalise D as a covering relational structure, i.e., an ordered pair $\langle H_{\delta}(A), R_{\supset}[H_{\delta}(A)] \rangle$ such that the covering relation $R_{\supset}[H_{\delta}(A)]$ is nothing but the transitive reduction of $R_{\supset}[H_{\delta}(A)]$, i.e., the greatest anti-transitive binary relation on $H_{\delta}(A)$ included in $R_{\supset}[H_{\delta}(A)]$. Note that the latter can be reinterpreted as a simple directed graph $\langle V, E \rangle$ where each vertex $v \in V$ bijectively corresponds to one cluster $X \in H_{\delta}(A)$, and each directed edge $e \in E$ bijectively corresponds to one ordered pair $\langle X, Y \rangle \in R_{\supset}[H_{\delta}(A)]$ (i.e., one immediate proper inclusion link $X \supset Y$ between the two distinct clusters $X, Y \in H_{\delta}(A)$). But $R_{\supset}[H_{\delta}(A)]$ is also a fully ramified hierarchical strict order relation as it satisfies three important properties (other than anti-reflexivity and transitivity): (i) it has a minimum element (i.e., a single origin), which is its exhaustive cluster (A itself); (ii) it has no convergent ramification (i.e., no reticulation); and (iii) it has a complete divergent ramification. As a result, $\langle H_{\delta}(A), R_{\supset}[H_{\delta}(A)] \rangle$ is also a hierarchical n -tree, i.e., a directed tree such that its single root, its n leaves, and its (internal) nodes respectively (and bijectively) corresponds to the exhaustive cluster of $H_{\delta}(A)$ (A itself), the n terminal clusters of $H_{\delta}(A)$ (the n singletons $\{x\}$ such that $x \in A$), and the non-trivial clusters of $H_{\delta}(A)$. As we shall see below, this last formalisation justifies why one can represent cladistic hypotheses by dendrograms.

We can now illustrate this approach with a very simple theoretical example. Let $A = \{a, b, c, d\}$ be a set of four distinct TU and let

$$D = ((ab)cd)$$

be a A -relative cladistic hypothesis saying that a and b are more closely related than either is to c and d (or, if you prefer, saying that a and b are closely related compared to c and d). In this case, the application of “ $\delta(x, y, z)$ ” into A (i.e., the searching for the 3-tuples $\langle x, y, z \rangle \in A^3$ for which “ $\delta(x, y, z)$ ” is true) implies the following statements (and nothing else): “ $\delta(a, a, b)$ ”, “ $\delta(a, a, c)$ ”, “ $\delta(a, a, d)$ ”, “ $\delta(b, b, a)$ ”, “ $\delta(b, b, c)$ ”, “ $\delta(b, b, d)$ ”, “ $\delta(c, c, a)$ ”, “ $\delta(c, c, b)$ ”, “ $\delta(c, c, d)$ ”, “ $\delta(d, d, a)$ ”, “ $\delta(d, d, b)$ ”, “ $\delta(d, d, c)$ ”, “ $\delta(a, b, c)$ ”, “ $\delta(b, a, c)$ ”, “ $\delta(a, b, d)$ ”, “ $\delta(b, a, d)$ ”. But given that $T_{\delta}(A) = \{\langle x, y, z \rangle \in A^3 \mid \delta(x, y, z)\}$, it follows that

$$\begin{aligned} T_{\delta}(A) = \{ & \langle a, a, b \rangle, \langle a, a, c \rangle, \langle a, a, d \rangle, \langle b, b, a \rangle, \langle b, b, c \rangle, \langle b, b, d \rangle, \\ & \langle c, c, a \rangle, \langle c, c, b \rangle, \langle c, c, d \rangle, \langle d, d, a \rangle, \langle d, d, b \rangle, \langle d, d, c \rangle, \\ & \langle a, b, c \rangle, \langle b, a, c \rangle, \langle a, b, d \rangle, \langle b, a, d \rangle \} \end{aligned}$$

i.e.,

$$T_{\delta}(A) = T_{A0} \cup \{\langle a, b, c \rangle, \langle b, a, c \rangle, \langle a, b, d \rangle, \langle b, a, d \rangle\}$$

Consequently, $T_{\delta}(A)$ is a non-trivial degree of equivalence relation on A and so, $\langle A, T_{\delta}(A) \rangle$ is well a finite non-trivial degree of equivalence relational structure. Hence our first (and more basic) formalisation of D .

Then, let us recall that for all $x, y \in A$, the $T_\delta(A)$ -relative degree of equivalence class of x and y is the set $[x \& y]_{T_\delta(A)} =_{df} \{z \in A \mid \langle x, y, z \rangle \notin T_\delta(A)\}$. One therefore obtains: $[a \& a]_{T_\delta(A)} = \{a\}$, $[b \& b]_{T_\delta(A)} = \{b\}$, $[c \& c]_{T_\delta(A)} = \{c\}$, $[d \& d]_{T_\delta(A)} = \{d\}$, $[a \& c]_{T_\delta(A)} = [c \& a]_{T_\delta(A)} = [a \& d]_{T_\delta(A)} = [d \& a]_{T_\delta(A)} = [b \& c]_{T_\delta(A)} = [c \& b]_{T_\delta(A)} = [b \& d]_{T_\delta(A)} = [d \& b]_{T_\delta(A)} = [c \& d]_{T_\delta(A)} = [d \& c]_{T_\delta(A)} = \{a, b, c, d\}$, and $[a \& b]_{T_\delta(A)} = [b \& a]_{T_\delta(A)} = \{a, b\}$. But given that $A//T_\delta(A) =_{df} \{X \subseteq A \mid (\exists x, y \in A) X = [x \& y]_{T_\delta(A)}\}$, it follows that

$$A//T_\delta(A) = \{\{a\}, \{b\}, \{c\}, \{d\}, \{a, b, c, d\}, \{a, b\}\}$$

i.e.,

$$A//T_\delta(A) = H_{A0} \cup \{\{a, b\}\}$$

By defining $H_\delta(A) = A//T_\delta(A)$, one can therefore conclude that $H_\delta(A)$ is a non-trivial hierarchical classification of A and so, that $\langle A, H_\delta(A) \rangle$ is well a finite non-trivial hierarchical classificatory structure. Hence our second (and more familiar) formalisation of D .

After that, consider the set

$$R_\supset[H_\delta(A)] = \{\langle X, Y \rangle \in H_\delta(A) \times H_\delta(A) \mid X \supset Y\}$$

Given the extensional definition of $H_\delta(A)$, one obtains

$$R_\supset[H_\delta(A)] = \{\langle \{a, b, c, d\}, \{a\} \rangle, \langle \{a, b, c, d\}, \{b\} \rangle, \langle \{a, b, c, d\}, \{c\} \rangle, \langle \{a, b, c, d\}, \{d\} \rangle, \langle \{a, b, c, d\}, \{a, b\} \rangle, \langle \{a, b\}, \{a\} \rangle, \langle \{a, b\}, \{b\} \rangle\}$$

This set being a relation of hierarchical strict order on $H_\delta(A)$, it follows that $\langle H_\delta(A), R_\supset[H_\delta(A)] \rangle$ is well a finite relational structure of hierarchical strict order. Hence our third formalisation of D . And finally, one can easily calculate the transitive reduction $R_\sqsubset[H_\delta(A)]$ of $R_\supset[H_\delta(A)]$ by removing all the ordered pairs of the latter entailed by transitivity, i.e.,

$$R_\sqsubset[H_\delta(A)] = \{\langle \{a, b, c, d\}, \{c\} \rangle, \langle \{a, b, c, d\}, \{d\} \rangle, \langle \{a, b, c, d\}, \{a, b\} \rangle, \langle \{a, b\}, \{a\} \rangle, \langle \{a, b\}, \{b\} \rangle\}$$

Such a set is a covering relation on $H_\delta(A)$ and so, $\langle H_\delta(A), R_\sqsubset[H_\delta(A)] \rangle$ is well a finite covering relational structure or, if you prefer, a hierarchical 4-tree such that its 4 leaves bijectively correspond to the 4 elements of A . Hence our fourth formalisation of D .

We now turn with the dendritic representations of $\langle H_\delta(A), R_\sqsubset[H_\delta(A)] \rangle$ and $\langle H_\delta(A), R_\supset[H_\delta(A)] \rangle$. The first kind of dendritic representation that interest us in this paper is that of “arrow diagram”. Given a finite simple binary relational structure, i.e., an ordered pair $\langle A, R \rangle$ such that (i) A is finite non-empty set, and (ii) R is a binary relation on A , i.e., a set such that $R \subseteq A^2$, its arrow diagram is a picture made up of dots and arrowed lines such that (iii) the dots bijectively represent the members of A , meaning that there is exactly one dot $f(x)$ for any $x \in A$, and (iv) for all $x, y \in A$, there is exactly one arrowed line joining $f(x)$ to $f(y)$ if and only if $\langle x, y \rangle \in R$. Figure 1a illustrates the

Fig. 1 Dendritic representations of the order-like formalisations of the cladistic hypothesis $D = ((ab)cd)$. **a** Arrow diagram of the relational structure of hierarchical strict order $\langle H_\delta(A), R_\supset[H_\delta(A)] \rangle$ such that $A = \{a, b, c, d\}$, $H_\delta(A) = H_{A0} \cup \{\{a, b\}\}$, and $R[H_\delta(A)] = \{\langle X, Y \rangle \in H_\delta(A) \times H_\delta(A) \mid X \supset Y\}$. **b** Arrow diagram of the corresponding covering relational structure $\langle H_\delta(A), R_\supset[H_\delta(A)] \rangle$, i.e., Hasse diagram of $\langle H_\delta(A), R_\supset[H_\delta(A)] \rangle$. **c** Dendrogram of D ; note that the latter can be obtained from the previous Hasse diagram from the following simplifications: deletion of the dots representing the clusters of $H_\delta(A)$, replacement of the arrowed lines by non-arrowed lines, replacement of the label " $\{x\}$ " by the label " x " for any x of A , and deletion of all the other labels.

arrow diagram of $\langle H_\delta(A), R_\supset[H_\delta(A)] \rangle$. Then, suppose that $\langle A, R \rangle$ is also a finite relational structure of strict order. In this case, its arrow diagram can be simplified into a "Hasse diagram" by the removing of all arrowed line corresponding to transitivity. In other words, the Hasse diagram of such an order is nothing but the arrow diagram of its corresponding covering relational structure. Figure 1b shows a diagram which is both the Hasse diagram of $\langle H_\delta(A), R_\supset[H_\delta(A)] \rangle$ and the arrow diagram of $\langle H_\delta(A), R_\supset[H_\delta(A)] \rangle$. Finally, one can obtain a dendrogram from a given Hasse diagram for the specific case where the order $\langle A, R \rangle$ is also hierarchical from the following simplifications: (a) deletion of the dots representing the element of A , (b) replacement of the arrowed lines by non-arrowed lines, and (c) deletion of all the labels except those of the terminal dots. Note also that for the specific case where $\langle A, R \rangle$ represents the inclusive order of a given hierarchical classificatory structure $\langle S, A \rangle$, all the labels " $\{x\}$ " (which represent the singletons of A) are replaced by labels " x " for all $x \in S$. Figure 1c depicts the dendrogram of both $\langle H_\delta(A), R_\supset[H_\delta(A)] \rangle$ and $\langle H_\delta(A), R_\supset[H_\delta(A)] \rangle$ that can be deduced from the arrow/Hasse diagram of figure 1b.

6 Three-item statements

In phylogenetics, a 3is, also called a triad (Adams 1986: 302), a triplet (Vach 1994; Wilkinson 1994: 344), or still a rooted triple (Bryant 2003: 3) is an assertion saying that given three distinct objects (of the same kind), two of them are more closely related than either is to the third. More abstractly, that consists of saying that two of them are equivalent compared to the third (given a context or a parameter).

For example, a possible 3is for the three distinct elements a , b , and c of A is the one asserting that a and b are equivalent compared to c . If appropriate,

one writes “ $(ab)c$ ”, “ $c(ab)$ ”, or still “ $ab|c$ ”. Note that the order between a and b is irrelevant, i.e., $(ab)c = (ba)c$. Indeed, “ a and b are equivalent compared to c ” actually corresponds to the conjunction of “ a is equivalent to b compared to c ” and “ b is equivalent to a compared to c ”. But if true, then that also means for any $T \in \mathcal{D}_A$ and all distinct $x, y, z \in A$ that $(xy)z$ is deductible from T if and only if $\{\langle x, y, z \rangle, \langle y, x, z \rangle\} \subseteq T$.

Now let three distinct $x, y, z \in A$. In this case, the 3is $(xy)z$ can be defined as follows:

$$(xy)z =_{df} \{\langle x, x, y \rangle, \langle x, x, z \rangle, \langle y, y, x \rangle, \langle y, y, z \rangle, \langle z, z, x \rangle, \langle z, z, y \rangle, \langle x, y, z \rangle, \langle y, x, z \rangle\}$$

i.e.,

$$(xy)z =_{df} T_{\{x, y, z\}0} \cup \{\langle x, y, z \rangle, \langle y, x, z \rangle\}$$

where $T_{\{x, y, z\}0}$ denotes the trivial degree of equivalence relation on $\{x, y, z\}$. For any $\langle T, H \rangle \in \mathcal{D}_A \times \mathcal{H}_A$, and for all distinct $x, y, z \in A$, such a definition implies the following consequences: (i) $(xy)z = (yx)z$, (ii) for all distinct $u, v, w \in A$, $(uv)w = (xy)z$ if and only if $\{u, v\} = \{x, y\}$ and $w = z$, (iii) $(xy)z$ is deductible from T if and only if $(xy)z \subseteq T$ or, if you prefer, if and only if $z \notin [x \& y]_T$, (iv) $(xy)z$ is deductible from H if and only if there is at least one $X \in I_H$ such that $\{x, y\} \subseteq X$ and $z \notin X$, and (v) the deduction of $(xy)z$ from T is equivalent to the deduction of $(xy)z$ from H if and only if $f(T) = H$ and $g(H) = T$.

Then, suppose that T implies k 3is $t_1, t_2, \dots, t_k$. In this case, one has (vi)

$$T = T_{A0} \cup \bigcup_{i=1}^k t_i$$

and (vii) either

$$\bigcup_{i=1}^k t_i \supset T_{A0}$$

for the case where for all distinct $x, y \in A$, there are at least one 3is $t_i \subseteq T$ and one $z \in A$ such that $t_i \in \{(xy)z, (xz)y, (yz)x\}$, or

$$T_{A0} \cap \bigcup_{i=1}^k t_i \notin \left\{ \emptyset, T_{A0}, \bigcup_{i=1}^k t_i \right\}$$

otherwise. In other words, the 3is deductible from T always correspond to the non-trivial part of the latter (T implies at least one 3is if and only if $T \neq T_{A0}$), i.e.,

$$\bigcup_{i=1}^k t_i \supset I_T$$

For the cladistic hypothesis D seen above, one can therefore write

$$T_\delta(A) = T_{A0} \cup (ab)c \cup (ab)d$$

Now let $\langle T, H \rangle \in \mathcal{D}_A \times \mathcal{H}_A$, and let three distinct $x, y, z \in A$. In this case, $(xy)z$ is part-whole compatible with T and/or H if and only if $(xy)z$ is respectively deducible from T and/or H . Naturally, this implies that the part-whole compatibility of $(xy)z$ with T is equivalent to the part-whole compatibility of $(xy)z$ with H if and only if $f(T) = H$ and $g(H) = T$. Then let $(uv)w$ and $(xy)z$ be two 3is. In this case, $(uv)w$ and $(xy)z$ are pairwise compatible if and only if there is at least one non-trivial degree of equivalence relation T on $\{u, v, w\} \cup \{x, y, z\}$ such that $(uv)w$ and $(xy)z$ are both part-whole compatible with T . But if true, then that also means that $(uv)w$ and $(xy)z$ are pairwise compatible if and only if there is at least one non-trivial hierarchical classification H of $\{u, v, w\} \cup \{x, y, z\}$ such that $(uv)w$ and $(xy)z$ are both part-whole compatible with H . It follows that $(uv)w$ and $(xy)z$ are pairwise incompatible if and only if $\{u, v, w\} = \{x, y, z\}$ and $(uv)w \neq (xy)z$. By contraposition, that also means that $(uv)w$ and $(xy)z$ are pairwise compatible if and only if either $\{u, v, w\} \neq \{x, y, z\}$ or $(uv)w = (xy)z$. Finally, let $(xy)z$ be a 3is and let $T \in \mathcal{D}_A$. In this case, $(xy)z$ is compatible by aggregation with T if and only if there is at least one non-trivial degree of equivalence relation T_a on $A \cup \{x, y, z\}$ such that (i) $T_a \supseteq T$ and (ii) $(xy)z$ is part-whole compatible with T_a . Thus, compatibility by aggregation generalises both part-whole compatibility and pairwise compatibility.

Another consequence of the previous definitions of 3is is that the properties seen above for any $T \in \mathcal{D}_A$ can be rewritten in terms of 3is. For all distinct $w, x, y, z \in A$, this gives:

15. $(xy)z \subseteq T \rightarrow \neg[(xz)y \subseteq T \vee (yz)x \subseteq T]$
16. $(wx)z \subseteq T \rightarrow [(wx)y \subseteq T \vee (wy)z \subseteq T]$
17. $[(xy)w \subseteq T \wedge (yz)w \subseteq T] \rightarrow (xz)w \subseteq T$
18. $[(wx)y \subseteq T \wedge (wy)z \subseteq T] \rightarrow [(wx)z \subseteq T \wedge (xy)z \subseteq T]$
19. $[(wx)y \subseteq T \wedge (xy)z \subseteq T] \rightarrow [(wx)z \subseteq T \wedge (wy)z \subseteq T]$
20. $[(wx)y \subseteq T \wedge (yz)w \subseteq T] \rightarrow [(wx)z \subseteq T \wedge (yz)x \subseteq T]$

Among them, property 15, which is in fact another statement of the pairwise incompatibility property seen above, follows from properties 1 and 5. In the same way, property 16 is entailed by property 4, property 17 by property 3, property 18 by properties 9 or 10, property 19 by property 11, and property 20 translates property 12.

We can also write these properties less formally. Respectively, this gives (here too, for all distinct $w, x, y, z \in A$):

15. If $(xy)z$, then not $(xz)y$ and not $(yz)x$
16. If $(wx)z$, then $(wx)y$ and/or $(wy)z$
17. If $(xy)w$ and $(yz)w$, then $(xz)w$
18. If $(wx)y$ and $(wy)z$, then $(wx)z$ and $(xy)z$
19. If $(wx)y$ and $(xy)z$, then $(wx)z$ and $(wy)z$
20. If $(wx)y$ and $(yz)w$, then $(wx)z$ and $(yz)x$

Note that properties 17 to 20 correspond to Dekker's (1986) dyadic rules established for fully ramified undirected n -trees when they are reformulated for hierarchical n -trees by Wilkinson et al. (2004: 994) and completed by Rineau (2017) and Rineau et al. (2020). To be more precise, property 17 corresponds to the rooted variant of first Dekker's dyadic rule, and properties 18 to 20 correspond to the rooted variants of the second one.

Now, consider the following writing rule: for all distinct $x, y, z \in A$, there is one ordered pair of parentheses "(" and ")" that contain x and y but not z – i.e., one may write " $(\dots x \dots y \dots) \dots z \dots$ " where " $\dots$ " stands for any other object of A – if and only if $\{\langle x, y, z \rangle, \langle y, x, z \rangle\} \subseteq T$. In other words, any instance of " $(\dots x \dots y \dots) \dots z \dots$ " means that there is at least one $X \in I_H$ such that $\{x, y\} \subseteq X$ and $z \notin X$. In this case, one can rewrite properties 16 to 20 in terms of four-item statements (4is; Fig. 2), i.e., (still for any four distinct objects $w, x, y, z \in A$):

16. If $(wx)z$, then $(wx)yz$ and/or $(wxy)z$ (insertion property)

Fig. 2 Illustration of 3is properties in terms of 4is: **a** insertion property, **b** in-out combination, **c** 1st asymmetric combination, **d** 2nd asymmetric combination, **e** symmetric combination, **f** in-in combination.

17. If $(xy)w$ and $(yz)w$, then $(xyz)w$ (in-out combination)
18. If $(wx)y$ and $(wy)z$, then $((wx)y)z$ (1^{st} asymmetric combination)
19. If $(wx)y$ and $(xy)z$, then $((wx)y)z$ (2^{nd} asymmetric combination)
20. If $(wx)y$ and $(yz)w$, then $(wx)(yz)$ (symmetric combination)

Let us make some words about the name of these properties that are given in parenthesis. On the one hand, property 16 is called "insertion property" because it actually translates the fact that there are seven possibilities for inserting the object y as a fourth leaf on the hierarchical 3-tree $((wx)z)$, i.e., (i) $((wy)x)z$, (ii) $((wxy)z)$, (iii) $((w(xy))z)$, (iv) $((wx)y)z$, (v) $((wx)yz)$, (vi) $((wx)(yz))$, and (vii) $((wx)z)y$. On the other hand, properties 17 to 20 are combination properties because they allow us to deduce a 4is from two 3is sharing two items. More precisely, property 17 is called "in-out combination" because of the distribution of these two shared items, and the names of the properties 18 to 20 follow from the shape of the obtained 4is.

Until here, all the combinatory relationships between 3is have been deduced only from basic and derived properties of degree of equivalence relations. It is however possible to deduce some additional properties by taking into consideration the mathematical equivalence between these relations and hierarchical classifications. For example, let $\langle T, H \rangle \in \mathcal{D}_A \times \mathcal{H}_A$ such that $H = f(T)$ and $T = g(H)$, and let four arbitrary distinct elements $w, x, y, z \in A$. Then, suppose that $(wx)y \subseteq T$ and $(wx)z \subseteq T$. In this case, there are (at least) two clusters X and Y in I_H such that: (i) $\{w, x\} \subseteq X$ and $y \notin X$, and (ii) $\{w, x\} \subseteq Y$ and $z \notin Y$. Thus, $X \cap Y \supseteq \{w, x\}$ and so $X \cap Y \neq \emptyset$. But if true, then $X \subseteq Y$ and/or $Y \subseteq X$, since $H \in \mathcal{H}_A$. As a result, there is necessarily at least one $Z \in I_H$ such that: (i) $\{w, x\} \subseteq Z$ and $\{y, z\} \cap Z = \emptyset$, and (ii) $Z = X$ and/or $Z = Y$. By reconsidering the previous parenthesis-based symbolism, we can now write this property in terms of 4is:

21. If $(wx)y$ and $(wx)z$, then $(wx)yz$ (in-in combination)

As for property 17, its name follows from the distribution of the shared items between $(wx)y$ and $(wx)z$.

7 Higher order rules

From these basic combination rules can be derived higher-order combination rules. In a non-exhaustive way, properties 17, 21, 18, and 19 can respectively be generalised as follows (Vautrin 2019):

22. If $(x_1x_2)y$, $(x_2x_3)y$, ..., and $(x_{k-1}x_k)y$, then $(x_1x_2...x_k)y$ (generalised in-out combination)
23. If $(xy)z_1$, $(xy)z_2$, ..., and $(xy)z_k$, then $(xy)z_1z_2...z_k$ (generalised in-in combination)
24. If $(x_1x_2)x_3$, $(x_1x_3)x_4$, ..., and $(x_1x_{k-1})x_k$, then $((...(((x_1x_2)x_3)x_4)...x_{k-1})x_k)$ (generalised 1^{st} asymmetric combination)

25. If $(x_1x_2)x_3, (x_2x_3)x_4, \dots$, and $(x_{k-2}x_{k-1})x_k$, then
 $((\dots(((x_1x_2)x_3)x_4)\dots)x_{k-2})x_{k-1})x_k$ (generalised 2^{st} asymmetric combination)

Property 20 can be generalised in two main ways:

26. If $(x_1x_2\dots x_k)y_1, (x_1x_2\dots x_k)y_2, \dots, (x_1x_2\dots x_k)y_n$, and $(y_1y_2\dots y_n)x_1, (y_1y_2\dots y_n)x_2, \dots, (y_1y_2\dots y_n)x_k$, then $(x_1x_2\dots x_k)(y_1y_2\dots y_n)$ (apical generalised symmetric combination)
27. If $(x_1y_1)x_2, (x_1y_1)x_3, \dots, (x_1y_1)x_k$, and $(x_2y_2)x_1, (x_2y_2)x_3, \dots, (x_2y_2)x_k$ and $(x_ky_k)x_1, (x_ky_k)x_2, \dots, (x_ky_k)x_{k-1}$, then $(x_1y_1)(x_2y_2)\dots(x_ky_k)$ (basal generalised symmetric combination)

To give an idea of the range of rules that can be inferred from the 3is rules, one can also mix the previous rules. Let us consider for example two compatible trees with only one informative node having a common cluster (including the same terminals). If the common cluster is the root, then there are three possibilities:

22. If $(x_1x_2\dots x_k)x_{k+1}\dots x_l$ and $(x_{k+1}\dots x_l)x_1x_2\dots x_k$, then $(x_1x_2\dots x_k)(x_{k+1}\dots x_l)$
23. If $(x_1x_2\dots x_k)x_{k+1}\dots x_m$ and $(x_{k+1}\dots x_l)x_1x_2\dots x_kx_{l+1}\dots x_m$, then $(x_1x_2\dots x_k)(x_{k+1}\dots x_l)x_{l+1}\dots x_m$
24. If $(x_1x_2\dots x_kx_{k+1}\dots x_l)x_{l+1}\dots x_m$ and $(x_1x_2\dots x_k)x_{k+1}\dots x_m$, then $((x_1x_2\dots x_k)x_{k+1}\dots x_l)x_{l+1}\dots x_m$

If the common cluster is the root for one tree but not for the other, then there is only one possibility:

22. If $(x_1x_2\dots x_kx_{k+1}\dots x_l)x_{l+1}\dots x_m$ and $(x_1x_2\dots x_k)x_{k+1}\dots x_l$, then $((x_1x_2\dots x_k)x_{k+1}\dots x_l)x_{l+1}\dots x_m$

If the common cluster is not the root in both trees, then there is only one possibility:

22. If $(x_1x_2\dots x_k)x_{k+1}\dots x_l$ and $(x_1x_2\dots x_k)x_{l+1}\dots x_m$, then $(x_1x_2\dots x_k)x_{k+1}\dots x_lx_{l+1}\dots x_m$

The set of these higher-order rules allows the following definition, namely that any pair of trees sharing a cluster in common (by all the terminals it contains) can be unambiguously combined into a tree.

8 Application

Bryant and Steel (1995: 448, theorem 3) have shown that there is an infinite number of 3is k -adic rules irreducible to l -adic rules such that $2 \leq l < k$ for any integer $k > 3$. Based on an example of five compatible but pairwise non-combinable 3is, we show here how it is possible, at least for this example, to reduce this 5-adic rule to the 3is properties stated previously.

Fig. 3 Illustration of the set of five 3is non-pairwise combinable by application of basic dyadic combination properties, but combinable as a whole into larger trees. **a** the five initial 3is, **b** the two cladistic trees implied by them.

To simplify the reading, we will not detail all of the 3is generated by the combinations, but only the intermediate trees generated by the inference rules (thus, a combination can take up 3is from a tree previously inferred). Moreover, higher-order inference rules stated in the above part are also used for simplicity.

Let $\{a, b, c, d, e, f, g\}$ be a set of seven elements and $\{(ab)g, (dg)e, (cd)a, (df)b, (ef)a\}$ be a set of five triplets (Fig. 3a; Wilkinson pers. com.). Let us start from the following property:

- If $(dg)e$ then $(adg)e$ and/or $(dg)ae$ (2-3 negative transitivity)

We begin by demonstrating that knowing $\{(ab)g, (dg)e, (cd)a, (df)b, (ef)a\}$, then $(adg)e$ is incompatible with them. If $(adg)e$ is not possible, then $(dg)ae$ is necessarily true.

8.1 If $(gda)e$ is true, then

- If $(ad)e$ and $(cd)a$, then $((cd)a)e$ (asymmetric combination)
- If $(ad)e$ and $(ef)a$, then $(ad)(ef)$ (symmetric combination)
- If $(cd)a$ and $(ad)f$, then $((cd)a)f$ (asymmetric combination)
- If $(ef)a$ and $(ac)f$, then $(ef)(ac)$ (symmetric combination)
- If $(cd)a$ and $(cd)e$ and $(cd)f$, then $(cd)ae$ (generalized in-in combination)
- If $(ef)a$ and $(ef)c$ and $(ef)d$, then $(ef)acd$ (generalized in-in combination)
- If $(cda)e$ and $(cda)f$, then $(cda)ef$ (combination by common node)
- If $(cd)ae$ and $(cda)ef$ and $(ef)acd$, then $((cd)a)(ef)$ (combination by common node)

addition of g

- If $(ef)d$ and $(dg)e$, then $(ef)(gd)$ (symmetric combination)
- If $(cd)e$ and $(dg)e$, then $(cdg)e$ (in-out combination)
- If $(ef)g$ and $(cg)e$, then $(ef)(cg)$ (symmetric combination)
- If $(ef)ag$ and $(ef)c$ and $(ef)d$ and $(ef)g$, then $(ef)acd$ (generalized in-in combination)

- If $(cg)f$ and $(ad)f$ and $(cd)f$, then $(acd)g$ (generalized in-out combination)
- If $(dg)e$ and $(ad)e$ and $(cd)e$, then $(acd)g$ (generalized in-out combination)
- If $(acd)g$ and $(acd)g$, then $(acd)ef$ (combination by common node)
- If $(acd)ef$ and $(ef)acd$, then $(ef)(acd)$ (combination by common node)

addition of b

- If $(ab)g$ and $(ag)e$, then $((ab)g)e$ (asymmetric combination)
- If $(ab)g$ and $(ag)f$, then $((ab)g)f$ (asymmetric combination)
- If $(ab)e$ and $(ac)e$, then $(abc)e$ (in-out combination)
- If $(ad)f$ and $(ab)f$, then $(abd)f$ (in-out combination)

But $(abd)f$ implies $(bd)f$, which contradicts $(df)b$ (incompatibility property). Thus $(ad)g$ is false and so, $(dg)ae$ is necessarily true (disjunctive syllogism).

8.2 If $(dg)a$ is true, then

- If $(dg)a$ and $(ab)g$, then $(dg)(ab)$ (symmetric combination)
- If $(dg)a$ and $(cd)a$, then $(cd)g$ (in-out combination)
- If $(cd)a$ and $(ab)d$, then $(ab)(cd)$ (symmetric combination)
- If $(dg)b$ and $(cd)b$, then $(cd)g$ (in-out combination)
- If $(cd)g$ and $(gcd)b$, then $(cd)g$ (combination by common node)
- If $(ab)c$ and $(ab)d$ and $(ab)g$, then $(ab)cdg$ (generalized in-in combination)
- If $(cd)g$ and $(ab)cdg$, then $(cd)g(ab)$ (combination by common node)

addition of f

- If $(df)b$ and $(ab)d$, then $(ab)(df)$ (symmetric combination)
- If $(df)a$ and $(dg)a$, then $(df)g$ (in-out combination)
- If $(df)b$ and $(dg)b$, then $(df)g$ (in-out combination)
- If $(df)b$ and $(dg)b$ and $(cd)b$, then $(cdf)g$ (combination by common node)
- If $(df)a$ and $(dg)a$ and $(cd)a$, then $(cdf)g$ (combination by common node)
- If $(cdf)g$ and $(cdf)g$, then $(cdf)ab$ (combination by common node)
- If $(ab)c$ and $(ab)d$ and $(ab)f$ and $(ab)g$, then $(ab)cdfg$ (generalized in-in combination)
- If $(cdf)ab$ and $(ab)cdfg$, then $(ab)(cdfg)$ (combination by common node)

addition of e

- If $(ef)a$ and $(ab)f$, then $(ab)(ef)$ (symmetric combination)
- If $(df)b$ and $(dg)b$ and $(cd)b$ and $(ef)b$, then $(cdef)g$ (generalized in-in combination)
- If $(df)a$ and $(dg)a$ and $(cd)a$ and $(ef)a$, then $(cdef)g$ (generalized in-in combination)
- If $(cdef)g$ and $(cdef)g$, then $(cdef)ab$ (combination by common node)

- If $(ab)c$ and $(ab)d$ and $(ab)e$ and $(ab)f$ and $(ab)g$, then $(ab)cdefg$ (generalized in-in combination)
- If $(dfgcfe)ab$ and $(ab)cdefg$, then $(ab)(cdefg)$ (combination by common node)
- If $(gd)e$ and $(eg)a$, then $((gd)e)a$ (asymmetric combination)
- If $(gd)e$ and $(eg)b$, then $((gd)e)b$ (asymmetric combination)
- If $((gd)e)a$ and $((gd)e)b$, then $((gd)e)ab$ (combination by common node)
- If $((gd)e)ab$ and $(ab)gde$, then $((gd)e)(ab)$ (combination by common node)

In summary, if $(ab)g$, $(gd)e$, $(cd)a$, $(df)b$, and $(ef)a$, then $(ab)(cdefg)$ and $((dg)e)(ab)$ (Fig. 3).

Consequently, only a finite and few numbers of primary inference rules are needed, all of the other being deductible. One can see that the key property allowing this is the 2-3 negative transitivity. The same result can be obtained empirically by coding the five 3is in a parsimony matrix (Nelson and Platnick 1991) and resuming the obtained optimal trees using a reduced cladistic consensus (Wilkinson 1994). Such a result suggests that using the 2-3 negative transitivity property could allow us to reduce all k -adic rules of combination for 3is to a finite set of more basic rules. Given the strong connexion between rooted trees and unrooted trees (Bryant and Steel 1995: 448-449), one can expect that our proposal for 3is is transposable to quartets.

Acknowledgements We thank Mark Wilkinson for sharing with us a stimulating example of five three-item statements, and Mike Steel for sending us Dekker's remarkable dissertation. We are also grateful to Paul Zaharias for his helpful comments on the manuscript. Finally, we thank Rachel Vautrin, René Zaragüeta, Paul Zaharias, and all the members of the Paris working group on 3ia and Cladistics for fruitful discussions.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Adams EN (1986) N-Trees as Nestings: Complexity, Similarity and Consensus. *J of Classif* 3:299-317
2. Barthélemy JP, Guénoche A (1991) Trees and proximity representations. John Wiley & Sons, Chichester
3. Benzécri JP (1972) Description mathématique des classifications. *Rev de Stat Appl* 20:23-56
4. Bininda-Emonds OR (2004) The evolution of supertrees. *Trends in ecol and evol* 19:315-322
5. Bryant D (2003) A classification of consensus methods for phylogenetics. In: Janowitz MF, Lapointe FJ, McMorris FR, Mirkin B, Robert FS (eds) Bioconsensus. DIMACS series in discrete mathematics and theoretical computer sciences 61. American Mathematical Society, Providence, pp 163-184
6. Bryant D, Steel M (1995) Extension Operations on Sets of Leaf-Labelled Trees. *Adv in Appl Math* 16:425-453

7. Camin JH, Sokal RR (1965) A method for deducing branching sequences in phylogeny. *Evol* 311-326
8. Colonius H, Schulze HH (1981) Tree structures for proximity data. *Br J of Math and Stat Psychol* 34:167-180
9. Dekker MCH (1986) Reconstruction methods for derivation trees. Master thesis, Vrije Universiteit Amsterdam
10. Dobson AJ (1974) Unrooted Trees for Numerical Taxonomy. *J of App Probab* 11: 32-42
11. Estabrook GF, McMorris FR (1980) When One Estimate of Evolutionary Relationships is a Refinement of Another? *J of Math Biol* 10:367-373
12. Fleischauer M, Böcker S (2017) Bad Clade Deletion supertrees: a fast and accurate supertree algorithm. *Mol biol and evol* 34:2408-2421
13. Hennig W (1965) Phylogenetic systematics. *Annu rev of entomol* 10:97-116
14. Hull DL (1970) Contemporary systematic philosophies. *Annu rev of Ecol and Syst* 1:19-54
15. Leclerc G (1985a) Les hiérarchies de parties et leur demi-treillis. *Math et sci hum* 89:5-34
16. Leclerc G (1985b) La comparaison des hiérarchies: indices et métriques. *Math et sci hum* 92:5-40
17. Margush T, McMorris FR (1981) Consensus n-Trees. *Bull of Math Biol* 43:239-244
18. McMorris FR, Powers RC (2003) The Arrovian Program from weak orders to hierarchical and tree-like relations. In: Janowitz MF, Lapointe FJ, McMorris FR, Mirkin B, Robert FS (eds) *Bioconsensus. DIMACS series in discrete mathematics and theoretical computer sciences* 61. American Mathematical Society, Providence, pp 37-45
19. Nelson G (1970) Outline of a theory of comparative biology. *Syst Zool* 19:373-384
20. Nelson G (1972) Phylogenetic relationship and classification. *Syst Zool* 21:227-231
21. Nelson G (1973) Classification as an expression of phylogenetic relationships. *Syst Zool* 22:344-359
22. Nelson G (1994) Homology and systematics. In: Hall BK (ed) *Homology: the hierarchical basis of comparative biology*. Academic Press, San Diego, pp 101-149
23. Nelson G, Ladiges PY (1991) Three-area statements: standard assumptions for biogeographic analysis. *Syst Biol* 40:470-485
24. Nelson G, Platnick NI (1991) Three-Taxon Statement: a more precise use of parsimony? *Cladistics* 7:351-366
25. Nguyen N, Mirarab S, Warnow T (2012) MRL and SuperFine+MRL: new supertree methods. *Algorithms for mol biol* 7:1-13
26. Ridley M (2004) *Evolution*, 3rd edition. Blackwell, Malden.
27. Rineau V (2017) Un nouveau regard cladistique sur l'anatomie comparée, la phylogénie, la systématique et la paléocologie des rudistes (Bivalvia, Hippuritida). PhD thesis, Université Pierre et Marie Curie.
28. Rineau V, Zaragüeta R, Bardin J (2020) Information Content of Trees: Three-taxon Statements Inference Rules and Dependency. *bioRxiv* 2020.06.08.141515
29. Prin S (2016) The relational view of phylogenetic hypotheses and what it tells us on the phylogeny/classification relation problem. In: Williams D, Schmitt M, Wheeler Q (eds) *The Future of Phylogenetic Systematics: The Legacy of Willi Hennig*. Cambridge University Press and Systematics Association, Cambridge, pp. 431-468
30. Sevillya G, Frenkel Z, Snir S (2016) Triplet MaxCut: a new toolkit for rooted supertree. *Methods in ecol and evol* 7:1359-1365
31. Vach W (1994) Preserving Consensus Hierarchies. *J of Classif* 11:59-77
32. Vautrin R (2019) Des arbres aux super-arbres : règles de combinaison des cladogrammes. Master thesis, Sorbonne Université
33. Wilkinson M (1994) Common Cladistic Information and its Consensus Representation: Reduced Adams and Reduced Cladistic Consensus Trees and Profiles. *Syst Biol* 43:343-368
34. Wilkinson M, Cotton JA, Thorley JL (2004) The Information Content of Trees and Their Matrix Representations. *Syst Biol* 53:989-1001
35. Zhang C, Rabiee M, Sayyari E, Mirarab S (2018) ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC bioinform* 19:15-30