The 5th Congress of Romanian Mathematicians, Romania, University of Pitesti, 22-28 iunie 2003

A NEW CONCEPTION ABOUT NUMERICAL TAXONOMY APPLIED IN NATURAL SCIENCES

Conf. Dr. Liviu Dragomirescu

University of Bucharest, Faculty of Biology

S U M M A R Y

Most authors agree that numerical taxonomy (n.t.) provides an objective classification. Nevertheless, S. Watanabe [15] proved that no classification can be objective. This is why, in my opinion, classifications must be performed without n.t. by authentic experts in biology, ecology, behavioral sciences etc. Instead of the presently used algorithms, based on different arbitrary "logics", we must build algorithms that model the proper logic of the corresponding field of study [2-9].

We axiomatically defined the homogeneity concept within this new theoretical development. For biology and ecology, we introduced and validated "homogeneities" specific to the respective subfields using real data.

The homogeneity concept generalizes to sets of n operational taxonomical units the similarity concept from n.t. applied in most of the recent studies [11]. We introduced several homogeneities (H^*, h^*, M_1^* etc.) [5]. The homogeneity H^* gave remarkable results in the classification of fish species [2]. h^* was used by Scaalje & Beus and became the first measure of variability available to molecular biology (measuring the reliability of electrophoresis) [13]. M_1^* proved to be suitable for measuring the homogeneity of a plant community in a given area [8].

For any homogeneity we apply only one clustering method, that of Buser & Baroni-Urbani's (B&B-U) [1], but modified in two steps, so as to yield a single result almost every time [5], respectively a necessarily unique result [6]. The B&B-U method with the first modification yielded the first "numerical demonstration" of correct ecological thinking, assuming that "structure determines functions, which are reflected in biomass" [8]. Applying the same method to χ^2 probability in a contingency table leads classifications with statistical signification for each class. These classifications are useful in serological anthropology [7] and in simultaneous comparison of several treatments efficiency [9].

Conceptually, we linked "homogeneity" to the "cohesion" defined by Watanabe, noticing that they represent just conjugated measures. Consequently, we can build "in mirror" homogeneities for cohesions and conversely. Therefore, we obtain a theoretical link between agglomerative and divisive algorithms that process the entire information without deforming it (contrarily to common algorithms) [3].

To strengthen this concept using the practical results enumerated, the necessity to study, in the same frame, ultrametric structures, subadditive and supraadditive measures appears obvious [3].

The 5th Congress of Romanian Mathematicians, Romania, University of Pitesti, 22-28 iunie 2003

R E Z U M A T

În mod usual, se consideră că taxonomia numerică (t.n.) obiectivizează clasificarea. S. Watanabe [15] a demonstrat însă că nici-o clasificare nu poate fi obiectivă. În consecință, în opinia mea, consider că t.n. trebuie să producă clasificările făcute fără t.n. de către specialiştii autentici din biologie, ecologie, etologie etc. Pentru aceasta, în locul algoritmilor utilizați în prezent şi care sunt bazați de diverse "logici" arbitrare, trebuie construiți algoritmi care să modeleze "logica" proprie specialității respective [2, 9].

Am definit axiomatic conceptul de omogenitate, iar pentru biologie şi ecologie am introdus şi validat pe date reale, "omogenitați" specifice subdomenilor respective.

Conceptul de omogenitate generalizează la mulțimi de n unități de clasificat, conceptul de similaritate din t.n. practicată în majoritatea studiilor recente [11]. Am introdus mai multe omogenități (H^*, h^*, M_1^* etc.) [5]. Omogenitatea H^* a dat un rezultat remarcabil în sistematica peştilor [2]. h^* a fost preluată de Scaalje & Beus, devenind prima măsură a variabilității pusă la dispoziția biologiei moleculare (pentru măsurarea fiabilității electroforezei) [13]. M_1^* s-a dovedit a fi potrivită pentru măsurarea omogenității unei comunități de plante dintr-un areal dat [8].

Oricărei omogenități i se aplică doar o singura metodă de grupare, cea a lui Buser & Baroni-Urbani (B&B-U) [1], amendată însă în doi paşi, asfel încât să producă rezultat unic aproape întotdeauna [5], respectiv, în mod necesar [6]. Metoda B&B-U cu primul amendament a produs prima "demonstrație numerică" a justeței gândirii ecologice care presupune că "structura determină funcțiile, iar acestea se citesc în biomase" [8]. Aplicând aceeaşi metodă probabilității lui χ^2 într-o tabelă de contingență se obțin clasificări cu semnificație statistică pentru fiecare clasă. Aceste clasificări s-au dovedit utile în antroplogia serologică [7] şi în compararea simultană a eficacității mai multor tratamente [9].

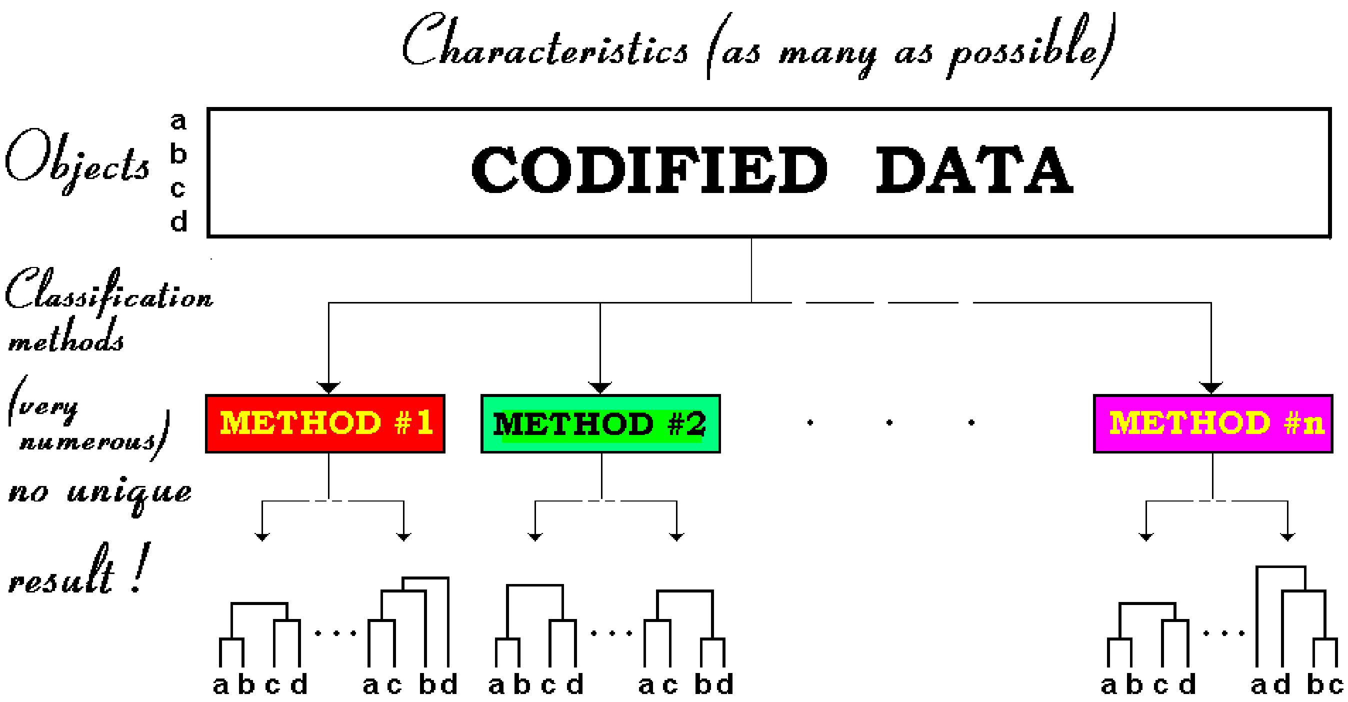
În plan axiomatic, am legat conceptul de "omognitate" cu cel de "coeziune" definit de Watanabe, observând că sunt măsuri conjugate. În consecință, se pot construi "în oglindă" omogenități pentru coeziuni şi reciproc. Totodată, se obține o legătură teoretică între algoritmii aglomerativi şi cei divizivi care prelucrează întreaga informație fără deformare (în mod contrar, algoritmilor comuni) [3].

Pentru consolidarea acestei concepții cu rezultatele practice enumerate, am semnalat necesitatea studierii în acelaşi cadru a structurilor ultrametrice şi a măsurilor subaditive şi supraaditive [3].

The 5th Congress of Romanian Mathematicians, Romania, University of Pitesti, 22-28 iunie 2003

CONCEPTIONS ON THE CONSTRUCTION AND APPLICATION OF NUMERICAL TAXONOMY (NT)

THE ANGLO-SAXON CONCEPTION (Sokal & Sneath, 1963, 1973)



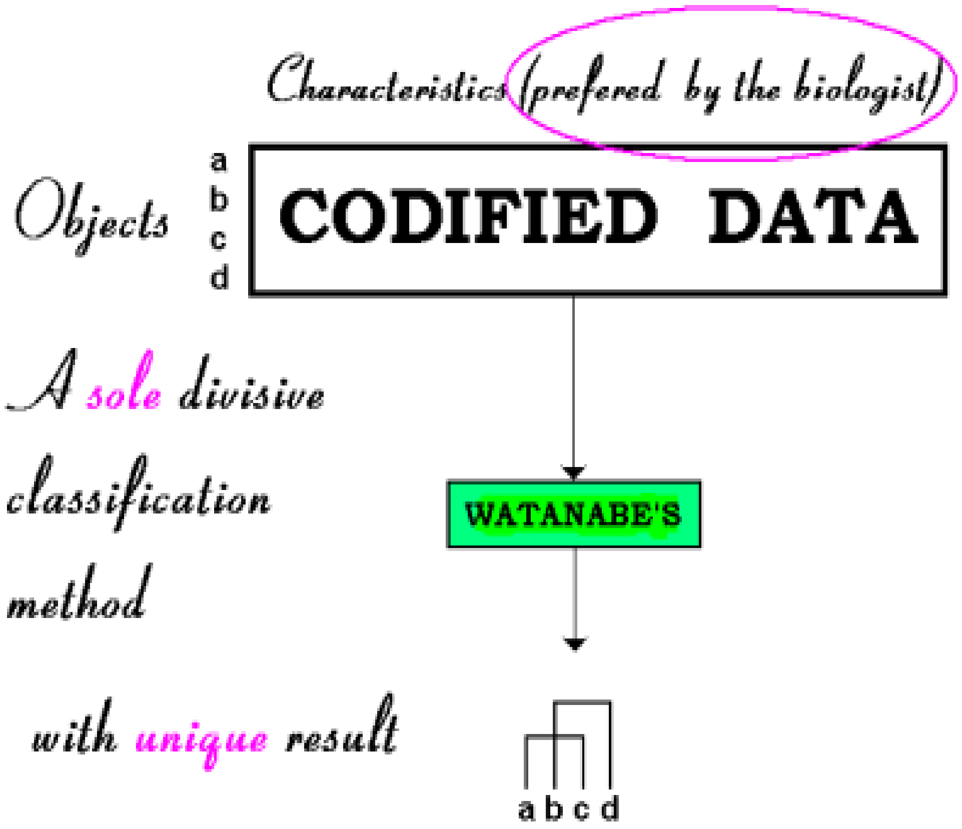
- The oldest conception, still fashionable today.
- NT = „the grouping by numerical methods of taxonomic units" (objects) „into taxa on the basis of their character states."

↖ „The greather the content of information in the taxa of a classification and the more characters on which it is based, the better a given classification will be."

← Uses a vast **literature of methods („CLASSICAL METHODES")**, which do NOT produce UNIQUE results.

- Main aims of NT methods are repetability and **objectivity**.
- NT „is viewed and practiced as an **empirical science**."

THE JAPANESE CONCEPTION (Watanabe, 1969)



- „Logic is insufficient to create concepts, classifications.” (Watanabe’s theoreme) ⇒ An **extralogic** viewpoint is needed to obtain concepts, trivial classifications.

↖ This viewpoint is expressed **only** in the list of the specialist’s preferred characteristics for the description of the objects to be classified.

↙ The object-characteristic table thus obtained is processed through a SOLE DIVISIVE CLASSIFICATION METHOD based on the concept of *Cohesion* which is a supra-additive set function, denoted by *c,* i.e.:

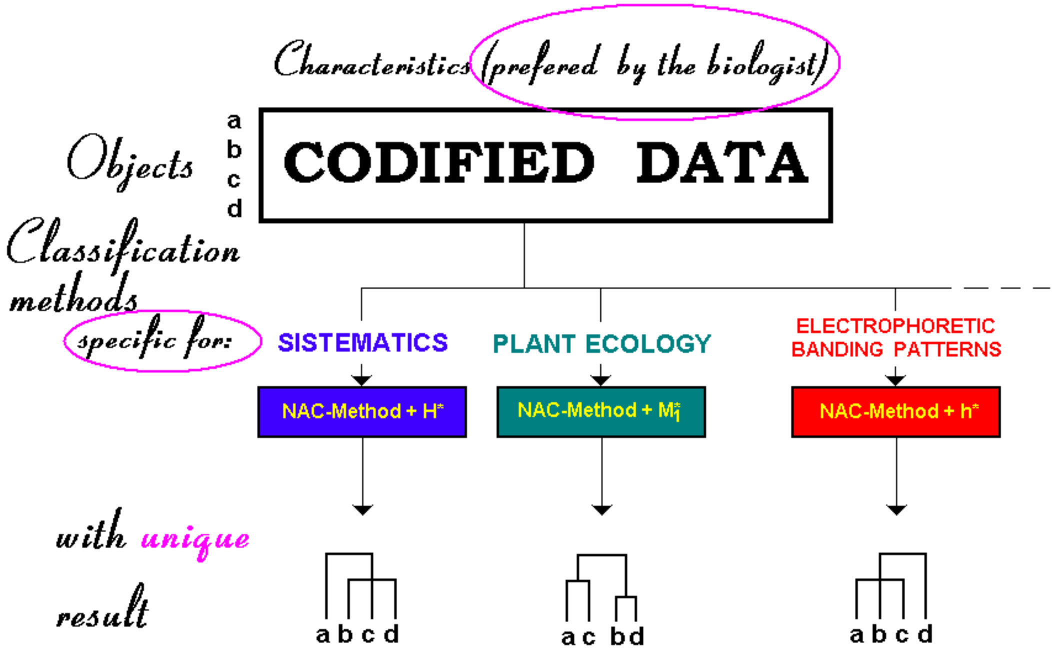
$$c\left(\bigcup_{k=1}^{r} A_k\right) \geq \sum_{k=1}^{r} c(A_k)$$

$(\forall)\{A_k\}_{k=1,2,\ldots,r}$ - family of mutually disjoint sub-sets,

since by “splitting” a whole $\bigcup_{k=1}^{r} A_k$ into several parts ($\{A_k\}_{k=1,2,\ldots,r}$), cohesion may be lost.

← The algorythm produces a UNIQUE RESULT.

THE PROPOSED CONCEPTION

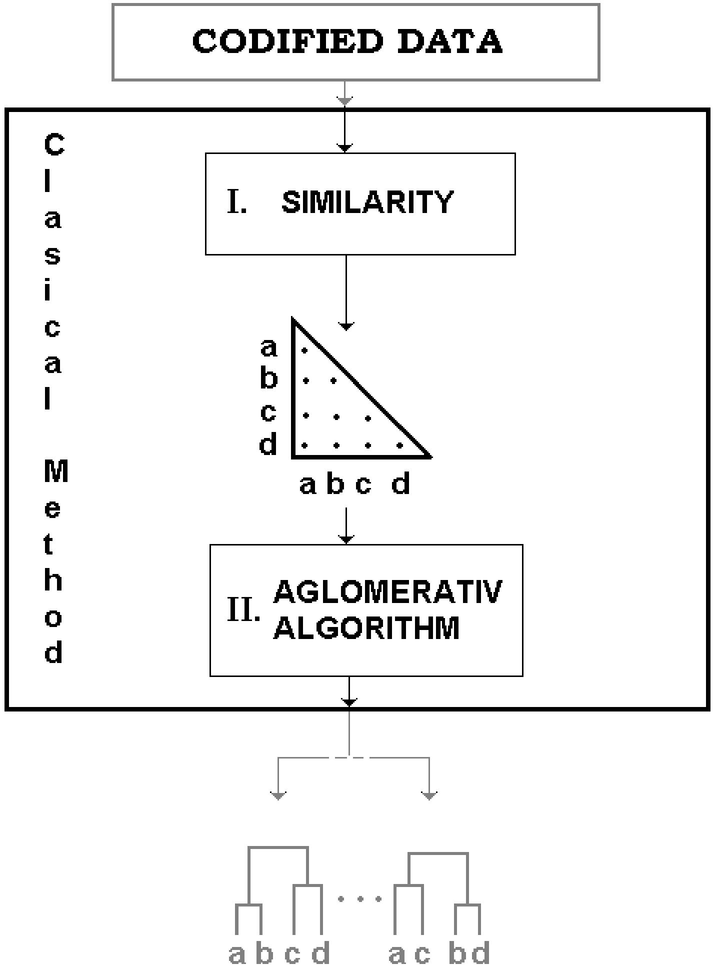


5. An NT method is VALID if it produces the same dendrogram as the one built by the biology specialist, by intuition and/or using other information than that used to build the object – characterisitic table (what we call VALIDATION BY SEMANTIC CONTROL).

1. NT should render TRANSPARENT the biologist's subjectivity.
2. This is expressed both in a:
   a. DATA STRUCTURE, and a
   b. PROGRAM STRUCTURE (IT).
3. **Data structure** includes:
   a. The preferred characteristics (cf. Watanabe),
   b. The objects put together to be classified and
   c. The data codification rules.
4. **Program structure** refers to the classification method.
   a. There cannot be a universally applicable method, as Watanabe claims.
   b. Nor do the tens of methods, each logically congruent, but in a logic independent from the data domain, which are operated at present, have any value.
   c. Specific methods should be found for the **mathematical shaping of the logic and extralogic characteristic to each specialist** in the respective biology sub-field.

The 5th Congress of Romanian Mathematicians, Romania, University of Pitesti, 22-28 iunie 2003

THE OPERATION PATTERN OF A CLASSICAL METHOD



I. A similarity *s,* defined for any pair of objects or OTU[s] (*Operational Taxonomic Unities*, acc. SNEATH and SOKAL, 1973) is computed:
$s: O\times O \to R_+$, $s(i, i) = 1$, $s(i, j) = s(j, i)$

- The results are retained in an inferior triangular matrix.

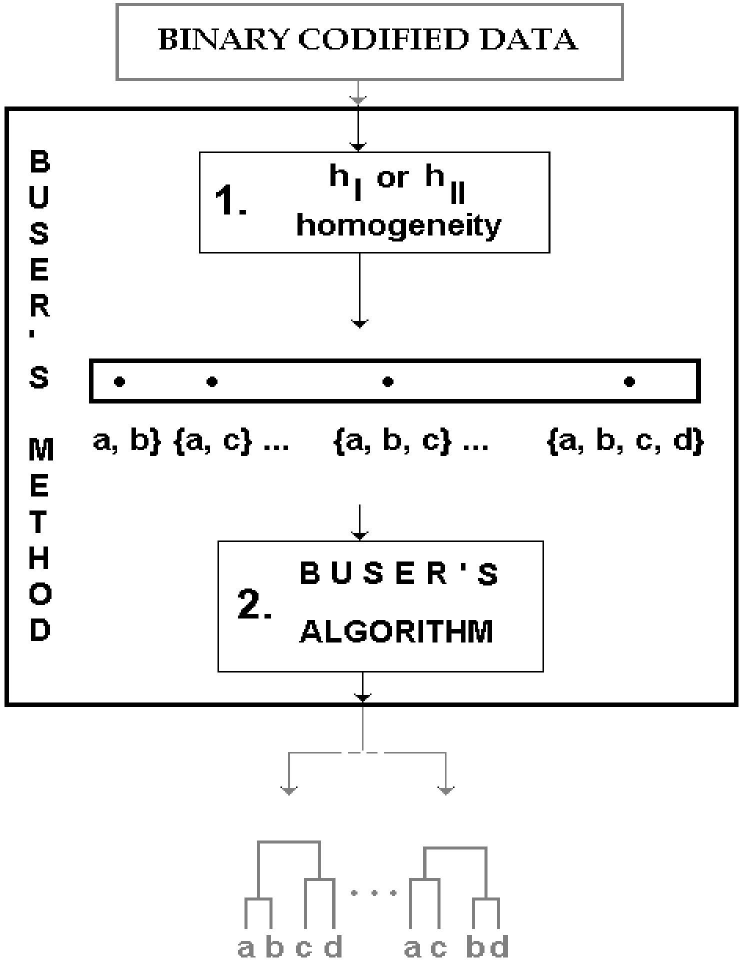
II. The matrix values are applied an agglomerative algorythm, such as: single link, complete link, group average link, weighted average link [GORDON] etc.

CRITICS:

- A classical method operates in two steps (I and II) „logically entierly different”.
- „Such a procedure necessarily implies a distortion of the original multidimensional similarity matrix, which could produce unpredictible mistakes and unnatural clusters.” (BUSER)

The 5th Congress of Romanian Mathematicians, Romania, University of Pitesti, 22-28 iunie 2003

THE OPERATION PATTERN OF BUSER'S METHOD (SPECIAL AGGLOMERATIVE METHOD)



Buser & Baroni – Urbani [1] have proposed:

1. replacing the similarity concept with that of „homogeneity within a set of two or **more** operational taxonomic unities" (objects) and two homogeneities (h_I, h_{II}) for BINARY data only and

2. replacing the various agglomerative algorythms with a sole agglomerative algorythm which operates „without loss or distortion of information as in the methods previously proposed."

In [2] we described this algorithm thus:

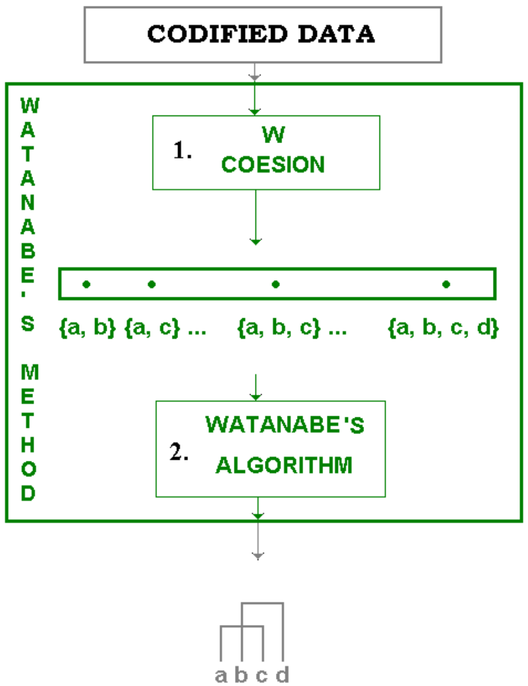
"Let a table which has L lines and N columns, representing a set of L OTU-s (*Operational Taxonomic Unities*, acc. SNEATH and SOKAL, 1973) describes in N characters.

(a) A LIST of all subsets of OTU-s is made up, calculating a homogeneity for each subset.

(b) A subset of the LIST which has the maximum homogeneity from among the other subsets of the LIST is considered a cluster.

(c) If the formed cluster is the whole set then the clustering is over, else the subsets which contain only strict parts of the already formed cluster(s) are eliminated from the LIST and the point (b) is applied to the new LIST."

THE OPERATION PATTERN OF WATANABE'S METHOD (SPECIAL DIVISIBLE METHOD)



1. Watanabe proposed a unique **COHESION**, **W**, named „interdependence" and defined for any subset $\bigcup_{k=1}^{r} A_k$ of OTUs which is divided into *r* parts:

$$W\left(\bigcup_{k=1}^{r} A_k \mid A_1, A_2, \ldots, A_r\right) = \sum_{k=1}^{r} H(A_k) - H\left(\bigcup_{k=1}^{r} A_k\right)$$

$(\forall)$ $\{A_k\}_{k=1, 2, \ldots, r}$ family of mutually disjoint parts, $A_k \subseteq I$, where

$H(A_k) = -\sum_{m} p(v_m) \cdot \log\, p(v_m)$ is the calculated entropy for the field of probability formed by the distinct column-vectors v_m from the OTU-character subtable which is made up only by the rows from A_k, where $p(v_m)$ are their relative frequencies.

2. Then *natural (divisive s. n.) classification strategies* (**WATANABE'S ALGORITHM – „NATURAL DIVISIVE CLASSIFICATION METHOD"**) are applied:

Ist Strategy: Strategies II and II are applied to the whole I set of OTUs. Then, they are applied to each A cluster thus obtained. This process is iterated until only one OTU-cluster is obtained.

IInd Strategy: Being given an A cluster, all the possible dichotomies $(A \mid A_1, A_2)$ (where $A=A_1 \cup A_2$; $A_1, A_2 \neq \emptyset$, $A_1 \cap A_2 = \emptyset$) are considered. If there is only one minimum branching cost dichotomy, it will be the prescribed "splitting". Otherwise, IIIrd strategy is applied.

IIIrd Strategy: The system of all minimum branching cost dichotomies is considered. The prescribed "splitting" will be the last sharp polychotomy, sharper than every dichotomy of the system. Intuitively, the clusters will stand for the "pieces" obtained by cutting the A subset according to all the minimum cost dichotomies.

CONTRIBUTIONS (within the framework of the proposed conception)

- **DEFINITION OF SEVERAL HOMOGENEITIES** (marked M_1*, M_2*, H*, h*) **WHICH PROVIDE A MATHEMATICAL MODEL TO THE BIOLOGIST'S SPECIFIC THINKING**

The biologist thinks, according to **Beckner** (acc. [14]), in „polithetic groups or natural classes”. The concept of “polithetic group” or “natural class”, is definable in terms of a set G of properties $f_1, f_2, \ldots, f_n$ so:

"(1) Each one (individual) possesses a large (but unspecified) number of properties in G.

(2) Each f in G is possessed by a large number of these individuals and

(3) no f in G is possessed every individual in the aggregate."

To these conditions one should add the condition formulated by Enescu: “A property is satisfied with more or less intensity” [10].

- ***M*₁*, respectively, *M*₂* models (nonbinary) the first condition, respectively Beckner's second condition, as well as Enescu's:**

[Let be A a set of L OTU-s described by N ordered (continuous or discrete) characters having the form of a matrix (a_{ij}) of real positive numbers.

$$M_1*(A)=\frac{\sum_{i=1}^{L}((\sum_{j=1}^{N}a_{ij})/S)}{L} \quad \text{were } S=\sum_{j=1}^{N}\max_{i=1,2,\ldots,L}(a_{ij})$$

$$M_2*(A)=\left[\sum_{j(\neq k\ \ for\ which\ \mu_k=0)=1}^{N}((\sum_{i=1}^{L}a_{ij})/(L\cdot\mu_j))\right]/N_1$$

where N_1 is the number of columns vanish non-identically.[5]]

- ***H**(A) = *M*₁*(A)· *M*₂*(A) models (nonbinary) the conditions of Beckner and Enescu [5].**

- ***h**(A) binary models any of Beckner's two conditions:**

$$h*(A)=\frac{1}{N*\cdot L}\sum_{j=1}^{N}\sum_{i=1}^{L}a_{ij}$$ were N* is the number of columns vanish non-identically and a_{ij} is the binary element which characterize the j-th attribute for the i-th OTU-s.

- **DEMONSTRATIONS WITH A PARTICULAR RELEVANCE IN APPLICATIONS**
  - **Binary homogeneity h^* is a generalization to sets A of L OTUs of Jaccard's index $J(A)$, defined for OTUs pairs and preferred by biologists and ecologists.** [If L=2 then $h^*(A) = (1/2)\cdot J(A)+1/2$]
  - **In binary logic, Beckner's first two conditions are equivalent, because homogeneities** M_1^*, respectively, M_2^* applied to a binary table become equal with h^*.
  - **THE BIOLOGIST'S THINKING IS NOT BINARY**, because in a nuanced logic Beckner's first two conditions are not equivalent (The implication $M_2^*(A) < M_2^*(B) \Rightarrow H^*(A) < H^*(B)$ is not true for any set A and any set B.)

- **THE IMPROVEMENT OF BUSER'S ALGORYTHM TO INCREASE THE PROBABILITY OF A UNIQUE RESULT**, by adding in step (b) condition "(b') which has a maximum number of OTU-s." [5]

- **THE AXIOMATIC DEFINITION OF THE CONCEPT OF HOMOGENEITY AND THE FORMULATION OF THE PRINCIPLES OF A NATURAL AGGLOMERATIVE CLASSIFICATION**

„Similar to WATANABE (1969) we formulate *principles of a Natural Agglomerative Classification*:

(A) By clustering, *homogeneity* (h) may be lost. Meaning that **the homogeneity must be a subadditive set function**:

$$h\left(\bigcup_{k=1}^{r} A_k\right) \leq \sum_{k=1}^{r} h(A_k) \quad (\forall) \quad \{A_k\}_{k=1,2,\ldots,r} \text{ family of mutually disjoint sub-sets.}$$

(B) We shall prefer maximum homogeneity clusters.

(C) In order that a classification should be natural, it must be unique (WATANABE, 1969).

(D) In consequence of the C principle the necessity of a unique decision at every stage of the method arises.

The B principle will grant priority to maximum homogeneity clusters while the C principle calls for forming a unique cluster (at the current) even though there are several maximum homogeneity clusters. The solution will be supported within the method by the following natural, in our opinion, decision:

Let A, B te two sets of OTUs which have the maximum homogeneity at the respective stage (h_{max}):

(i) if $A \cap B = \emptyset$, then both A and B subsets are considered to be clusters at the h_{max} level (within the same stage);

(ii) if $A \cap B \neq \emptyset$, then the $A \cup B$ subset is considered to be a cluster at the $h\,(A \cup B)(\leq h_{max})$ level. These principles bring forth the following agglomerative method, which yields a necessarily unique result." [6]

- **THE CONSTRUCTION OF THE „NATURAL AGGLOMERATIVE CLASSIFICATION METHOD" (NAC-MEROD), (which processes homogeneities without loss or distortion of information - as Buser's method – and produces a unique result – as Watanabe's method)**. **The method is based on the homogeneities defined above and the following algorythm** [6]:

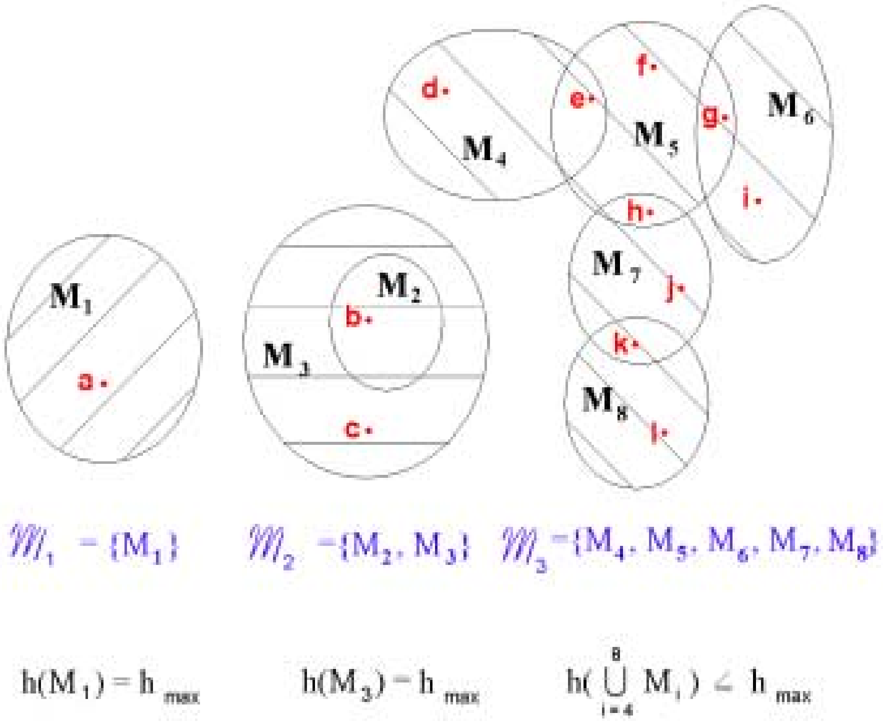
Ist Strategy: Strategies II and II are applied to the LIST of all OTUs subsets. The LIST is reduced by eliminating all the subsets which include only strict parts of clusters already formed by applying these strategies. The same process is undergone by *the new* LIST and so on, until LIST= Ø.

IInd Strategy: Being given a LIST of OTUs subsets, the $\mathcal{M}$ family of subsets is considered, ($\mathcal{M} \subseteq$ LIST), with the h_{max} maximum homogeneity within the LIST. If the subsets in $\mathcal{M}$ are mutually disjoint, these subsets are considered to be clusters at the respective level (h_{max}, within the current stage). Otherwise, IIIrd strategy is applied.

IIIrd Strategy: All the *complete families* (*in* $\mathcal{M}$)$^♣$ are considered (see $\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_3$ in the right figure for an intuitive understanding): $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_s$ and their unions $U_1, U_2, \ldots, U_s$ which are mutually disjoint and with homogeneities less than or equal to the maximum homogeneity (h_{max}) at the current stage. These unions are considered to be clusters in descending order of the homogeneity levels, substages within the current stage.

♣ **Definition:** *A complete (whole) family* (*in* $\mathcal{M}$) is called a conex family (in $\mathcal{M}$) $\mathcal{C}$ $^♦$, with the characteristic that $(\forall)\ A \in \mathcal{M}$ so as $(\exists) B \in \mathcal{C}$ with $A \cap B \neq \emptyset \Rightarrow A \in \mathcal{C}$.

♦ **Definition:** *A conex family* (in $\mathcal{M}$) is called a sub-family $\mathcal{C} \subseteq \mathcal{M}$ with the characteristic that $(\forall)\ A, B \in \mathcal{C}$ $(\exists)$ $p \in \mathbf{N}$ and $\{C_0, C_1, \ldots, C_p\} \subseteq \mathcal{C}$ so as $C_{i-1} \cap C_i \neq \emptyset$ $(\forall)$ $i=1, 2, \ldots, p$, where $A=C_0$ and $B=C_p$.



The 5th Congress of Romanian Mathematicians, Romania, University of Pitesti, 22-28 iunie 2003

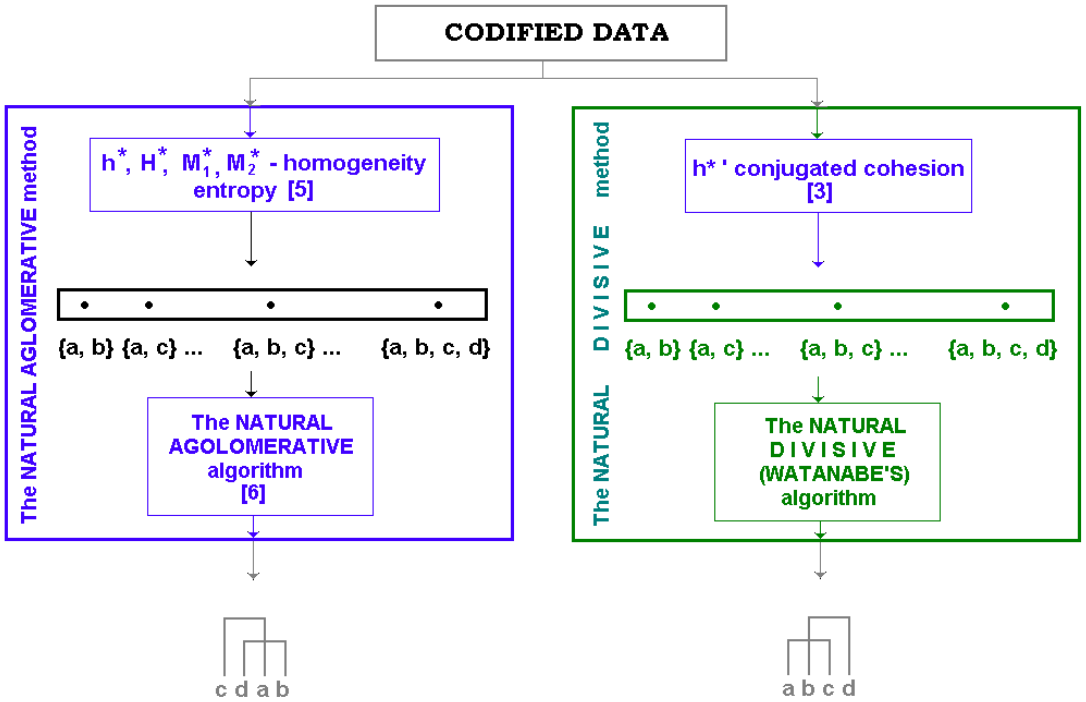
➢ THE CONSTRUCTION OF A „LINK” BETWEEN THE CONCEPT OF HOMOGENEITY AND THAT OF COHESION.

Starting from a theorem of the measure theory♣ one can build a linking bridge between the divisive and agglomerative algorithms: on the basis given cohesion, one can uniquely define a “conjugate” homogeneity and vice-versa [3, 4].

♣ If $\mu \geq 0$ is an additive measure (i.e. $\mu (A \cup B) = \mu(A) + \mu(B)\ \forall\ A, B$) that can be written as $\mu = \mu + \mu-$, the statement “μ+ is a superadditive measure” is equivalent to the statement “μ- is a subadditive measure” (μ+ and μ- are conjugate as againts μ) [15]

➢ PROOF OF h_I AND h^*’S SUBADDITIVITY [4] AND CONSTRUCTION OF CONJUGATED COHESIONS [3].

➢ CREATION OF A LINK BETWEEN THE NATURAL AGGLOMERATIVE METHOD AND THE NATURAL DIVISIVE METHOD:



These two algorythms are the best grounded ones, but they necessitate volumes of calculation that exponentially increase with the number of OTUs. Thus, a great practical advantage could be obtained provided a classification theory were built that may eventually lead to algorythms with less calculations, equivalent with these ones.

A „**theory of classification** is likely to become possible through the creation and depending of some **links between the theory of measure and topology** (knowing that a tree, the result of a cluster analysis, is equivalent to an ultrametric structure). One may possibly think of something analogous to the algebraic topology, that treats the links between the topological structures (of continuity) and the algebraic (discrete) ones.” [3].

Paralelly, as many applications as possible should be made to validate the superiority of these two algorythms.

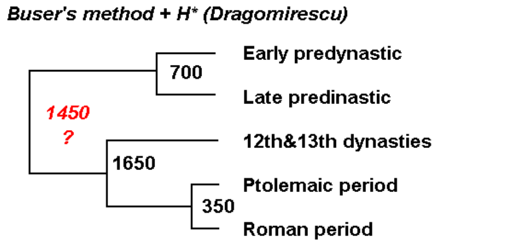
APLICATIONS

✓ ***h**** **was used by Scaalje & Beus and became the first measure of variability available to molecular biology (measuring the reliability of electrophoresis) [13].**

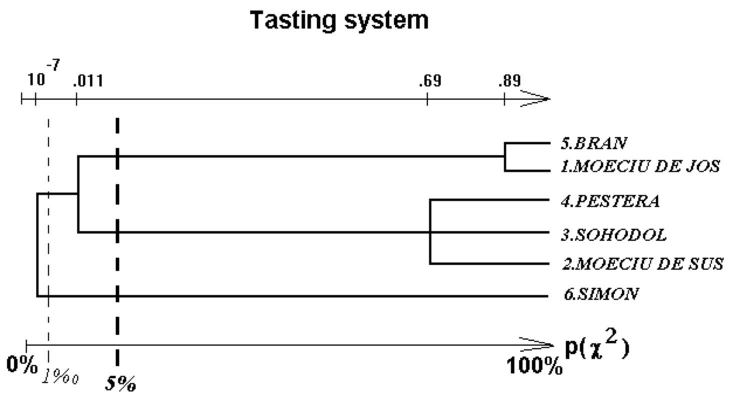
✓ ***The NATURAL AGLOMERATIV method applied to:***

✓ ***H**** **gave remarkable results in the classification of fish species: it produced the very classification presumed by the ichtiologist who built the object – characteristic table, which even classical algorythms did not lead to. [2]**

✓ ***H**** **produced the most accurate classification of five samples of 30 skulls. The anthropometric data comes from [12]. They were codified through an original procedure, described in [7].**

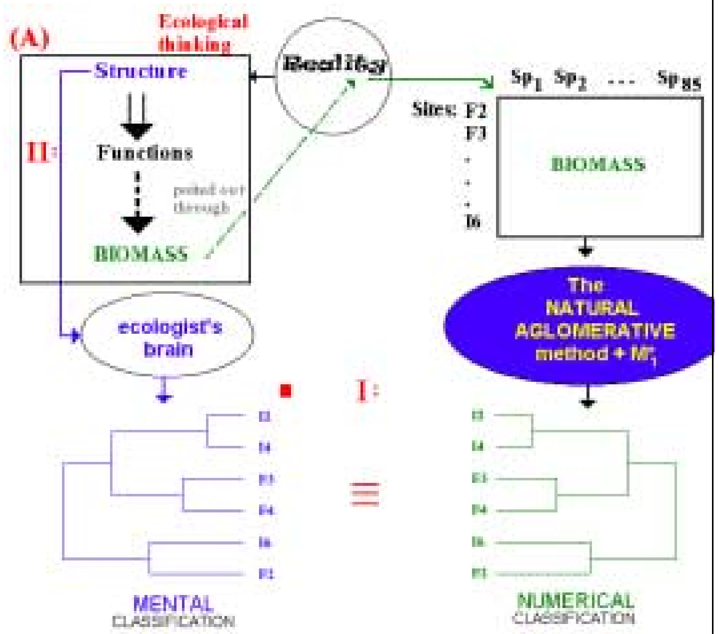


✓ **χ^2 probability in a contingency table leads classifications with statistical signification for each class. These classifications are useful in simultaneous comparison of several treatments efficiency [9] and in serological anthropology [7]:**



✓ ***M*₁* yielded the first "numerical demonstration" of correct ecological thinking, because**

✓ ***M*₁* proved to be suitable for measuring the homogeneity of a plant community in a given area [8].**



REFERENCES

[1] BUSER M., BARONI-URBANI C. 1982. A Direct Nondimensional Clustering Method for Binary Data. Biometrics, 38: 351 – 360.

[2] DRAGOMIRESCU L., CONSTANTINESCU V., BĂNĂRESCU P., 1985. A numerical taxonomy method adequate to the biological thinking. Applications Watanabe's example and the Achanthobrama Genus (*Pisces*, *Cyprinidae*). Traveaux du Muséum d'Histoire naturelle Grigore Antipa, 27 : 243-265.

[3] DRAGOMIRESCU L. 1986. Considerations on the structure and application of cluster analysis in biology (Note). Travaux du Museum d'Histoire Naturelle "Grigore Antipa", 28: 305-310.

[4] DRAGOMIRESCU (L.), 1986 – Contribuții privind aplicarea taxonomiei numerice în biologie. Teză de doctorat, Centrul de Statistică Matematică, Bucureşti.

[5] DRAGOMIRESCU L. 1987. Some extensions of Buser and Baroni-Urbani's clustering method. Biometrical Journal, 2: 345-353.

[6] DRAGOMIRESCU L., POSTELNICU T. 1991. A "natural" agglomerative clustering method for biology. Biometrical Journal, 7: 841-849.

[7] DRAGOMIRESCU L., POSTELNICU T. 1994. Specific numerical taxonomy methods in biological classification. pp. 31-45. In: Di Bacco M. *et al.* (eds), Statistical tools in human biology. Proceedings of the 17th Course of the International School of Mathemetics "G. Stampacchia". World Scientific Publishing CO. PTE. LTD., London.

[8] DRAGOMIRESCU L., GHEORGHE IULIANA. *Metodă de clasificare automată a tipurilor de comunități vegetale. Comunicare susținuta la a 6-a Conferință a Societății de Probabiliăți şi Statistică din România, ASE, Bucureşti, 21-22 februarie, 2003.*

[9] DRAGOMIRESCU L., STĂNESCU H. *A Method for Multiple Comparisons of the Efficiency of Several Treatments. Paper accepted for 3rd International Conference of Internaional Sociey for Clinical Biostatistics, 20th-24th July, 2003, London.*

[10] ENESCU, GH., 1980: Fundamentele logice ale gândirii. Editura ştiințifică şi enciclopedică. Bucureşti.

[11] GORDON A. D. 1999. Classification. 2nd Edition. Chapman & Hall/CRC. Boca Raton, London, New York, Washington, D. C.

[12] MANLY B. F. J., *Multivariate statistical methods. A primer.* (London, New York, Chapman and Hall, 1986).

[13] SCHAALJE G.B., BEUS B.D. 1997. A reliability index for presence-absence data. Communications in Statistical Theory and Methods, 26: 355-374.

[14] SNEATH, P.H.A., R.R. SOKAL, 1973: Numerical Taxonomy – the Principles and Practice of Numerical Classification. W.H.Freeman. San Francisco

[15] WATANABE (S.), 1969 – Knowing and Guessing: 299-448. New York.