Searching for order

*Andrzej DĄBROWSKI**

* Faculty of Mathematics and Computer Science, University of Wrocław

In 1950, one of the autumn meetings of the *Seminar on Applied Mathematics* in Wrocław was dedicated to the Ngandong skulls. Anthropologists expected mathematicians to order these skulls chronologically based on Czekanowski tables *graciously sent to us by Prof. T. Henzel* (quotation from the chronicle of the Seminar).

The full name is *Seminar on Applied Mathematics, General Section of the Department of Applied Mathematics of the State Mathematical Institute* [1].

The Ngandong skulls are a group of fossilized human skulls discovered in 1931 in the Solo River valley on the island of Java, Indonesia. These skulls are dated to approximately 100-200 thousands years ago and are considered to be remains of hominids of the species *Homo erectus*.

Jan Czekanowski (1882–1965) was an anthropologist and ethnographer. He studied anthropology, anatomy, ethnography, and mathematics at the University of Zurich. He participated in the *Seminar* in Wrocław (four times according to the Seminar's chronicle).

Order, understood as linear, is expected in many fields. Various rankings are created (for schools, universities, political support). It is hard to imagine sports events without rankings based on measured time, distance or points (pentathlon). Indexes are constructed to compare objects that are often difficult to compare. In economics, stock market states are assessed by stock market indices (e.g. Dow Jones, WIG), and the price level is assessed by the inflation index. The health status of a population is assessed by the average life expectancy or infant mortality rate. Comparison methods based on rankings and indices involve assigning a number to objects. And as we know, numbers can easily be ordered.

It is more difficult when an object can be described by a set of many numbers, and even more challenging when assignings numbers to it is hard. Such a problem arises when ordering shells with various patterns found on an ancient landfill (this is the so-called seriation problem in archaeology). Anthropologists turned to mathematicians with a similar problem, ordering the Ngandong skulls.

Czekanowski tables, which were available to Professor Steinhaus's team (Kazimierz Florek, Józef Łukaszewicz, Julian Perkal, and Stefan Zubrzycki), contained information on the degree of differentiation of each pair of skulls. This degree was expressed by the Euclidean distance in a seven-dimensional space of parameters representing the lengths of characteristic segments on the skull.

In general, such a table can be treated as a *discrepancy function*. This is any function $d(x,y)$ defined on pairs of elements from a given set X, which satisfies the conditions:

$$d(x,y) \geqslant 0, \quad d(x,y) = d(y,x) \text{ and } d(x,y) = 0 \Leftrightarrow x = y,$$

A special case of a discrepancy function is a distance functiona metric. A metric is a discrepancy function that also satisfies the triangle inequality: $d(x,z) \leqslant d(x,y) + d(y,z)$ for all x, y, z.

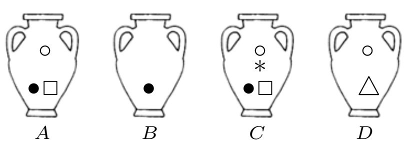


Fig. 1

Stanisław Kulczyński (1895-1975), zoologist, arachnologist, mountaineer. Rector of universities in Lviv and Wrocław.

Example 1. Vases. Four vases with decorations were discovered. Each vase is treated as a set of ornaments painted on it. The Kulczyński discrepancy function can be used as the discrepancy function, defined for two sets of ornaments R and S by the formula

$$d\,(R,S) = 1 - \frac{1}{2}\left(\frac{|R \cap S|}{|R|} + \frac{|R \cap S|}{|S|}\right).$$

It can be easily verified that such a function is indeed a discrepancy function. It is also quite intuitive, since $\frac{1}{2}\left(\frac{|R\cap S|}{|R|} + \frac{|R\cap S|}{|S|}\right)$ is the arithmetic mean of the fractions of common elements of R and S, contained in the set R and contained in the set S.

	A	B	C	D
A	0	1/3	1/8	7/12
B	1/3	0	3/8	1
C	1/8	3/8	0	5/8
D	7/12	1	5/8	0

Kulczyński discrepancy matrix.

The resulting discrepancy matrix is presented in the margin. Note that the Kulczynski function is not a metric, since $d(B,D) = 1 > \frac{11}{12} = d(B,A) + d(A,D)$.

Discrepancy function and linear order. How to introduce an order among objects for which we have a discrepancy matrix? Let us first consider a particular case. Assume that the objects are points distributed on the real line, and we take the discrepancy function to be the distance between them. Can we reconstruct the ordering of the points on the line from the discrepancy matrix alone? Yes, we can – we just need to choose an ordering that minimizes the sum of distances between consecutive points.

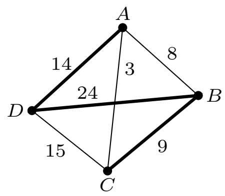


Fig. 2. Graph with Kulczyński discrepancies (multiplied by 24) and a path $ADBC$ of length 47.

It will be convenient for us to represent objects $O_1, O_2, \ldots, O_n$ with discrepancies $d_{ij} = d\,(O_i, O_j)$ by an undirected complete graph G with n vertices $O_1, O_2, \ldots, O_n$, whose edge O_iO_j has weight d_{ij}. An example of such graph is given in Fig. 2. We define the *length of the path* $O_{i_1}O_{i_2}\ldots O_{i_n}$ as the sum $d_{i_1i_2} + d_{i_2i_3} + \ldots + d_{i_{(n-1)}i_n}$.

Previous observations suggest that an order of objects $O_1, \ldots, O_n$ may be determined by the path in graph G of minimum length that passes through all its vertices (*an optimal path*).

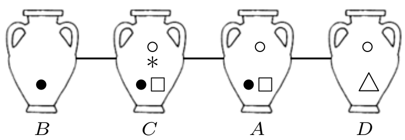


Fig. 3. Optimal path $BCAD$ in a graph with weights.

Example 2. Five capitals. The table below shows the road distances (in kilometers) between Amsterdam, Berlin, Paris, Rome, and Warsaw.

	A	B	P	R	W
A	0	650	510	1650	1140
B	650	0	1040	1460	570
P	510	1040	0	1430	1550
R	1650	1460	1430	0	1730
W	1140	570	1550	1730	0

Out of 60 possible paths, the shortest one goes from Warsaw through Berlin, Amsterdam, Paris, and all the way to Rome, and has the length of 3160 km.

Dendrites. Linear ordering is often insufficient and even inadequate. Let us quote here Julian Perkal's observation: *As I noticed, linear ordering is often unnatural in many cases, for example, a genealogical line often branches out.* [2]

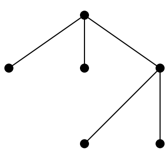


Fig. 4

The structure that allows a nearly linear ordering of the vertices of a graph is a dendrite, more commonly known as a tree – for historical reasons, we will use the former term here. A dendrite is a graph without cycles, which is also connected. These two conditions together mean that any two vertices are connected by a uniquely determined path (Fig. 4). Any path itself is also a special case of a dendrite.

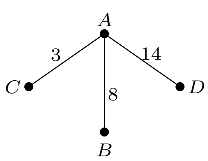


Fig. 5. Optimal dendrite for the problem of 4 weights. Its length (25) is smaller than the length of the optimal path $BCAD$ (26).

The length of a dendrite is defined as the sum of the weights of its edges. Inspired by previous observations, we assume that a dendrite of minimum length (*an optimal dendrite*) reproduces the order of n vertices of the graph in the best way (Fig. 5).

Wrocław taxonomy. In the example with weights, it was easy to identify the optimal dendrite. As the number of vertices in the graph increases, the question arises about computing it in an algorithmic way. In computer science, this problem is classical and known as the *minimum spanning tree problem.* According to [3], the earliest published solution to this problem (1926) comes from the Czech mathematician Otakar Borůvka, who was dealing with it in the context of developing an optimal electrical network in Moravia. The classical algorithms, known to participants of Olympiads in informatics, are Kruskal's and Prim's algorithms that were published in 1956 and 1957, respectively.
It seems that mathematicians from Wrocław were the first to address this problem in the context of methods of ordering and classifying (i.e. *taxonomy*) in anthropology, biology, or linguistics. Their method, published in 1951 ([4]), is basically Borůvka's algorithm (of which they were unaware) and is called the *Wrocław taxonomy* in Polish statistical literature.

As part of this work, the first step towards the general construction of an optimal dendrite was taken in 1949 by Kazimierz Florek. He noted that any optimal dendrite should contain segments connecting the nearest objects – that is, those connecting an object with its nearest neighbor. Such segments are called *connections of order I.*

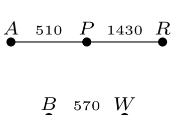
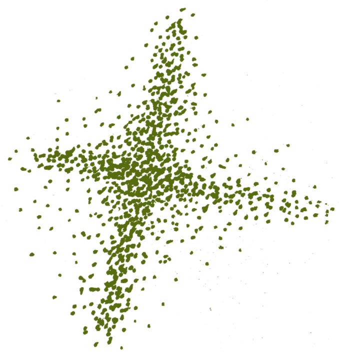


Fig. 6

In the example with the vases, this procedure solves the problem: connections of order I form a dendrite. However, this is not always the case. Connections of order I for 5 capitals form a disconnected subgraph, shown in the margin. Notice, however, that in the graph of connections of order I, cycles cannot occur – such a graph is called a *forest*, and it naturally breaks down into dendrites.



In 1950, the creators of taxonomy proposed an optimal dendrite construction, which is an iterative version of Florek's idea:

Method W

1. Build connections of order I for a given graph G. If they form a dendrite, the construction is finished. If not, proceed to step 2.

2. Create a new graph G_1. The vertices of graph G_1 are sub-dendrites formed from the connections of order I of graph G. The discrepancy between sub-dendrites A and B is the discrepancy of the nearest neighbors: $d(A, B) := \min\{d(P, Q) : P \in A, Q \in B\}$. Create connections of order I for graph G_1 (these are connections of order II for G). If they form a dendrite, the construction is finished. If not, repeat step 2 for graph G_1.

3. In this way, we create consecutive graphs $G_1, G_2, G_3, \ldots$ and for each of them, connections of the next order. These iterations must end because the number of vertices in subsequent graphs G_i decreases.

4. The construction of the final dendrite ends by connecting dendrites of successive orders with edges between objects that realize the nearest neighbor discrepancy.

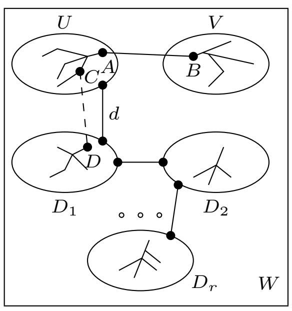


Fig. 7

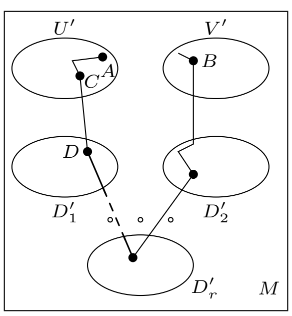


Fig. 8

Let us see how method W works for the example of 5 capitals. Connections of order I form sub-dendrites $O_1 = \{A, P, R\}$ and $O_2 = \{B, W\}$. Graph G_1 has vertices O_1 and O_2. The minimum distance from B to the set of points $\{A, P, R\}$ is 650, and the minimum distance from W to $\{A, P, R\}$ is 1140 km. Therefore, the discrepancy of the nearest neighbor between O_1 and O_2 is equal to 650 km, which is the distance between Berlin and Amsterdam. G_1 with the connection of order II between these capitals is already a dendrite. This ends the construction, creating the final optimal dendrite, which turned out to be the path $WBAPR$.

Proof of optimality of Method W. Without loss of generality, we can assume that all nonzero discrepancies in the graph G are different. If necessary, we can replace the discrepancies d_{ij} with $d'_{ij} = d_{ij} + \varepsilon_{ij}$ (for $i < j$), where $d'_{ji} = d'_{ij}$ and $\varepsilon_{ij} > 0$ are any chosen distinct numbers smaller than $\min d_{ij}$. After the construction is finished, we need to return to the discrepancies d_{ij}.

Let W be the dendrite constructed using method W. Suppose that W is not an optimal dendrite, so it is different from some optimal dendrite M. Therefore, there is an edge in W that does not appear in M – let it connect vertices A and B.

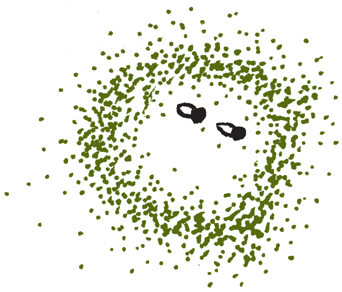
Assume that the edge AB has weight a and order k in the dendrite W. It connects dendrites U and V of order $k-1$ with vertex sets U' and V', respectively. The other dendrites of rank $k-1$, which we denote by $D_1, D_2, \ldots, D_r$, have vertex sets $D'_1, D'_2, \ldots, D'_r$. According to the construction of W, one of the following two cases holds:

1. a is smaller than each discrepancy between U and $D_1, \ldots, D_r$,
2. a is smaller than each discrepancy between V and $D_1, \ldots, D_r$.

Without loss of generality, assume that case 1 holds.

In dendrite M, vertices A and B are connected by a path that does not contain the edge AB, and it must contain an edge CD such that C belongs to U' and D does not. Assume that D belongs to D'_1. The weight of the edge CD is not smaller than the distance d between dendrites U and D_1 in dendrite W. Moreover, $a < d$ because case 1 holds. Therefore, we can create a new dendrite M' by replacing the edge CD in M with the edge AB. The length of M' is smaller than the length of M. This contradicts the assumption that M is an optimal dendrite.

Grouping. A set of objects might be non-homogenous: shells found in an ancient landfill or skulls found in a surveyed area may come from several distinct periods. How to divide the data so that dendrites, corresponding to the division, indicate significant differences in these groups?



A family of k dendrites $D_1, D_2, \ldots, D_k$ with sets of vertices $Z_1, Z_2, ..., Z_k$ is a *partition* of a given complete graph G with vertices Z when

$$Z = Z_1 \cup Z_2 \cup \ldots \cup Z_k \quad \text{and} \quad Z_i \cap Z_j = \emptyset \text{ for } i \neq j$$

The *length of the partition* is the sum of the lengths $l(D_i)$ of the component dendrites. The partition is *optimal* if the length of the partition is minimal. From this definition, it immediately follows that the components of an optimal partition must be optimal, so we can assume $D_i = W(Z_i)$.

Let's go back to the example of the four vases. They can be divided into two groups in seven ways:

Z_1	$\{A, B\}$	$\{A, C\}$	$\{A, D\}$	$\{A, B, C\}$	$\{A, B, D\}$	$\{A, C, D\}$	$\{B, C, D\}$
Z_2	$\{C, D\}$	$\{B, D\}$	$\{B, C\}$	$\{D\}$	$\{C\}$	$\{B\}$	$\{A\}$
$W(Z_1)$	A–B	A–C	A–D	B–A–C	B–A–D	C–A–D	B–C–D
$W(Z_2)$	C–D	B–D	B–C	D	C	B	A
$l(W(Z_1)) + l(W(Z_2))$	23	27	23	11	22	17	24

The optimal partition $D_1 = C$–A–B , $D_2 = D$ is a subgraph of the dendrite obtained by the W method (Fig. 5). It turns out that this is always the case.

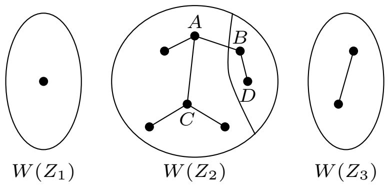


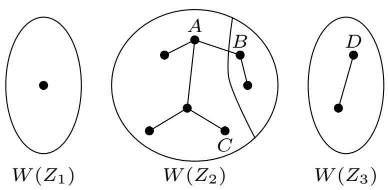
Fig. 9

Theorem. *If $\{W(Z_1), W(Z_2), ..., W(Z_k)\}$ is an optimal partition of the graph G with vertices Z, then*

$$W(Z_1) \cup W(Z_2) \cup \ldots \cup W(Z_k) \subset W(Z).$$

Proof. As before, without loss of generality, we assume that all non-zero discrepancies are distinct.

Suppose that there exists an edge AB in $W(Z_1) \cup W(Z_2) \cup \ldots \cup W(Z_k)$ that does not belong to $W(Z)$. There exists a path s_{AB} in $W(Z)$ connecting these vertices.



Rys. 10

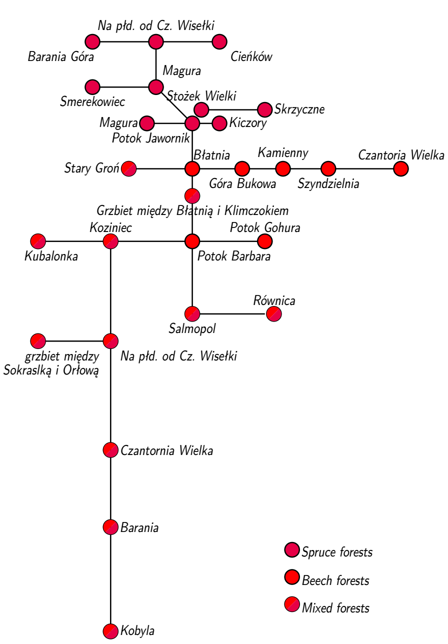
From the disjointness of Z_i, it follows that there exists an r such that the edge AB belongs to $W(Z_r)$. The elements of Z_r can be divided into subsets U and V as follows: U consists of the vertex A and all vertices in Z_r connected to A by a path in the dendrite $W(Z_r)$ that does not contain the edge AB; V consists of the vertex B and all vertices in Z_r connected to B by a path in the dendrite $W(Z_r)$ that does not contain the edge AB.

The path s_{AB} must contain an edge CD such that $C \in U$ and $D \notin U$. Let $d(A, B) = x$ and $d(C, D) = y$. The inequality $y < x$ holds; otherwise, replacing CD with AB in the dendrite $W(Z)$ would decrease its length, which would be a contradiction.

There are two cases: $D \in V$ and $D \in Z_s$ for $s \neq r$.

1. $D \in V$ (Fig. 9). In this case, replacing AB with CD in $W(Z_r)$ yields a dendrite that is not optimal, which is a contradiction.

2. $D \in Z_s$ for $s \neq r$ (Fig. 10). We replace the dendrite $W(Z_r)$ with the dendrite $W(U) \cup CD \cup W(Z_s)$ of length $l(W(U)) + y + l(W(Z_s))$, and the dendrite $W(Z_s)$ with $W(V)$ of length $l(W(V))$. The sum of the lengths of all k dendrites after the change is smaller than the sum of their lengths before the change, which contradicts the assumption that the partition is optimal.



Rys. 11

From the above theorem, we obtain a useful result in the context of finding an optimal grouping:

Corollary. *The optimal partition of a graph G into k subdendrites involves removing from the dendrite $W(Z)$ the $k-1$ edges with the largest discrepancies.*

A bit about applications. The *Seminar on Applied Mathematics* actively promoted the idea of taxonomy, applying it in various fields. In Steinhaus's *Mathematical Snapshots*, one can read about the taxonomy of liverworts (*Hepaticae*) in the Beskidy Mountains. The dendrite corresponding to the frequency of occurrence of different liverwort species was found to be related to the type of forest – which was an interesting discovery (Fig. 11).

A characteristic feature of the work of the *Seminar* was tackling of every problem, even the most unusual ones. At a meeting in January 1952, *Julian Perkal announced that he had made a dendrite of folk songs for Professor Czekanowski's daughter* (quote from the carefully kept minutes of the *Seminar*). The classification of folk songs using the method of Wrocław taxonomy became one of the important research tools for Anna Czekanowska-Kuklińska, a professor at the University of Warsaw (d. 2021) and the head of the Ethnomusicology Department she established.

In 1953, Stefan Zubrzycki published a work [5] using the Wrocław taxonomy, which answered astronomer Włodzimierz Zonn's question of whether stars form non-random constellations (referred to by the authors as "chains") or are randomly distributed on the celestial sphere. He showed that they are randomly arranged, confirming that constellations are only a mnemonic method of remembering the position of stars.

Julian Perkal ends his work on taxonomy (op. cit.) with a warning that *"...one can construct a machine for making dendrites. This creates a danger of a mechanical approach to natural objects and of gyrating false sometimes natural bills with mathematical methods."* It is worth remembering this.

Bibliography

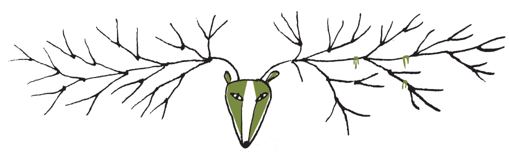
[1] Szczotka, W., *Fenomen Steinhausowskiego Seminarium Matematyki Stosowanej*, Antiquitates Mathematicae, (2018) vol. 12(1), 197-228, (2019) vol. 13(1), 177-233

[2] Perkal, J. (1953). *Taksonomia wrocławska.* Przegląd Antropologiczny

[3] Graham, R. L., Hell, P. (1985). *On the history of the minimum spanning tree problem.* Annals of the History of Computing, 7(1), 43-57.

[4] Florek K., Łukaszewicz J., Perkal J., Steinhaus H., Zubrzycki S., *Sur la liaison et la division des points d'un ensemble fini*, Colloquium Mathematicum (1951), vol.2, nr 3-4 str, 282-285

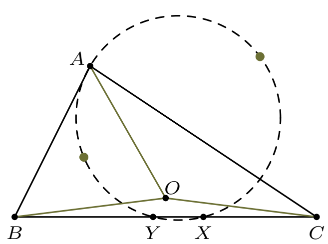
[5] Zubrzycki, F. (1953). *O łańcuszkach gwiezdnych.* Applicationes Mathematicae, 1(3), 197-205.



Problems

Edited by Dominik BUREK

M 1750. Can the numbers from 1 to 2023^2 be placed in the squares of a 2023×2023 board in such a way that for any choice of a row and a column, we can find three numbers on them, where one of the numbers is the product of the other two?
Solution on page 2



M 1751. Let O be the circumcenter of triangle ABC. Points X and Y on side BC are such that $AX = BX$ and $AY = CY$. Prove that the circumcircle of triangle AXY passes through the circumcenters of triangles AOB and AOC.
Solution on page 8

M 1752. Let $x_1, \ldots, x_n \in [0,1]$. Prove that
$$(1 - x_1x_2 + x_1^2) \cdot (1 - x_2x_3 + x_2^2) \cdot \ldots \cdot (1 - x_{n-1}x_n + x_{n-1}^2) \cdot (1 - x_nx_1 + x_n^2) \geqslant 1.$$
Solution on page 4

Edited by Andrzej MAJHOFER

F 1075. An eclipsing binary star system with radii r_1 and r_2 is observed from Earth at an angle α to the plane of the stars' mutual orbit. What is the relation between the angle α, radii r_1 and r_2, and the diameter d of the orbit? We assume that the orbit is circular.
Solution on page 7

F 1076. On one of the plates of a flat capacitor with capacitance C, a charge Q_1 is placed, and on the other plate, a charge Q_2 is placed. What is the potential difference between the plates?
Hint: As usual in problems of this type, we neglect boundary effects.
Solution on page 7