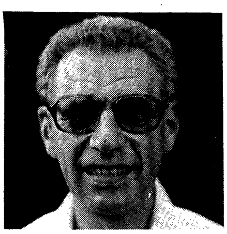
The Construction of Venn Diagrams*

Branko Grünbaum



Branko Grünbaum is Professor of Mathematics at the University of Washington in Seattle. He was born in Yugoslavia in 1929, and educated in Yugoslavia and in Israel. After receiving his Ph.D. from the Hebrew University in Jerusalem in 1958, he spent two years at the Institute for Advanced Study in Princeton, NJ. Since then he has taught at Hebrew University, Michigan State University, and the University of Washington. His main fields of interest are geometry (in particular, convex sets and polytopes, tilings and patterns, arrangements of lines, and other areas of intuitive geometry) and combinatorics (in particular, graph theory). He has published two books and some 120 research and survey papers on these topics. He received the Lester R. Ford award for 1975 for the article [2] related to the topic of the present paper and the Carl B. Allendoerfer award for 1978 from the Mathematical Association of America. He is a member of the editorial boards of several mathematical journals. Since 1974, he has served as a lecturer in the Program of Visiting Lecturers of the Mathematical Association of America.

1. Introduction. Most people have heard, on some occasion or another, about "Venn diagrams" and seen drawings of two or three circles which illustrate the idea in a simple case. But Venn diagrams can lead to many interesting problems of a geometric, topological, or combinatorial character. Several aspects of such questions were examined in recent articles [2, 4]. The present note deals with two other aspects.

To make this paper self-contained, we shall first define the usual Venn diagrams, and then formulate the two new results. The proofs will be given in Sections 2 and 3, while a number of comments are collected in Section 4.

A family $\mathcal{T} = \{A_1, A_2, \ldots, A_n\}$ of n simple closed curves A_j in the plane is called *independent* if every intersection of the type

$$X_1 \cap X_2 \cap \cdots \cap X_m \qquad (*)$$

is a nonempty set, where each X_j is chosen to be either the *interior* A_j^i or the *exterior* A_j^e of the curve A_j. Each set of the form (*) is called a *region* of the independent family. An independent family $\mathcal{T} = \{A_1, A_2, \ldots, A_n\}$ is called a *Venn diagram* (for n classes, or of n curves) provided each region of $\mathcal{T}$ is *connected*—that is, it is not the union of two disjoint, relatively open sets. In Figure 1, we illustrate these concepts, which are discussed in more detail in [2].

*Research supported by National Science Foundation Grant MCS-8001570.

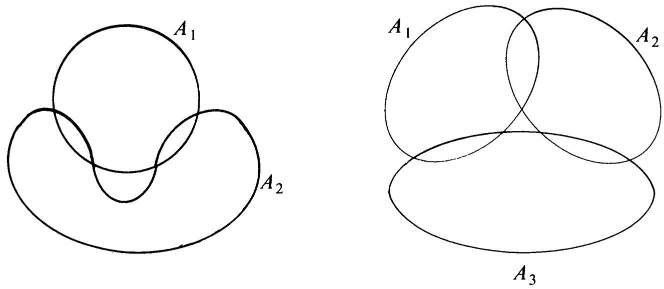


Figure 1. (a) The family $\mathscr{T} = \{A_1, A_2\}$ is independent, but is not a Venn diagram since the region $A_1^i \cap A_2^i$ is not a connected set. (b) The family $\mathscr{T} = \{A_1, A_2, A_3\}$ is not independent since $A_1^i \cap A_2^i \cap A_3^i = \emptyset$.

Venn diagrams were introduced by J. Venn in 1880 (see [6]) and popularized in his book [7]. Venn did consider the question of existence of Venn diagrams for an arbitrary number n of classes, and provided in [6] an inductive construction of such diagrams. His construction is illustrated in Figure 2. However, in his better known book [7], Venn did not mention the construction of diagrams with many classes; this was often mistakenly interpreted as meaning that Venn could not find such diagrams, and over the past century many papers were published in which the existence of Venn diagrams for n classes is proved.

One such paper is by Nowicki [5] who seems unaware of the fact that the "systems of curves" he is constructing are what other people call "Venn diagrams". But in this same paper Nowicki raises an interesting question: do there exist *irreducible* Venn diagrams for n classes? Here a Venn diagram for n classes $\mathscr{V}$ is called irreducible if each of the n independent families of $n-1$ curves, obtained from $\mathscr{V}$ by deleting in turn one of the n curves, fails to be a Venn diagram. In other words, a Venn diagram is irreducible if it cannot be obtained from a smaller Venn diagram by the addition of one curve.

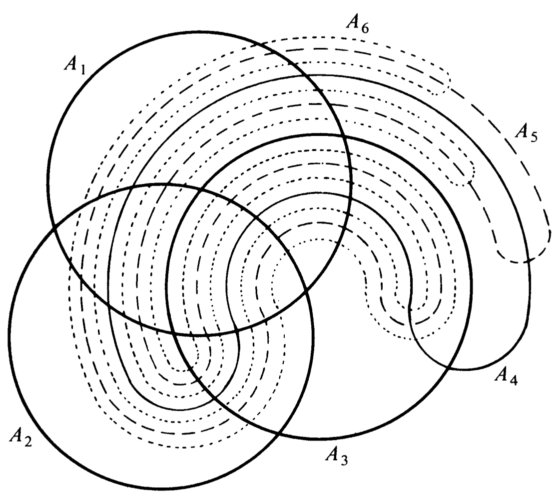


Figure 2. The inductive construction of Venn diagrams described by Venn (see [6], [2]). The curves marked A_1 to A_n form a Venn diagram for n classes, for each $n = 1, 2, 3, 4, 5, 6$, and it is clear how to continue for larger values of n.

A Venn diagram $\mathscr{V}$ is called *simple* provided no three of the curves of $\mathscr{V}$ pass through the same point. One of the aims of the present paper is to answer affirmatively a stronger form of Nowicki's question, by proving the following result:

Theorem 1. *For every $n \geqslant 5$, there exist simple and irreducible Venn diagrams for n classes.*

There is some irony in the fact that the proof of Theorem 1, given in Section 2, employs an inductive construction.

Another aim of this note is to consider Venn diagrams for situations with several possible outcomes (that is, for multivalued logic). In Section 3, we shall define "Venn diagrams for n classes with k outcomes each", and then indicate how to prove the following analogue of Venn's result:

Theorem 2. *For any positive integers n and k, with $k \geqslant 2$, there exist Venn diagrams for n classes with k outcomes each.*

2. Proof of Theorem 1. In Figures 3 and 4, we show Venn diagrams for five and six classes which have the properties required by Theorem 1. The example in Figure 3 is particularly interesting since it consists of five congruent ellipses, and contradicts the assertion made by Venn [6] (and repeated often since then) that with ellipses one cannot form Venn diagrams for five classes. Due to its symmetry, it is easy to check that the family of five ellipses in Figure 3 is a simple and irreducible Venn diagram. (It is not hard to show that any independent family of ellipses contains at most five members [2, Theorem 1].)

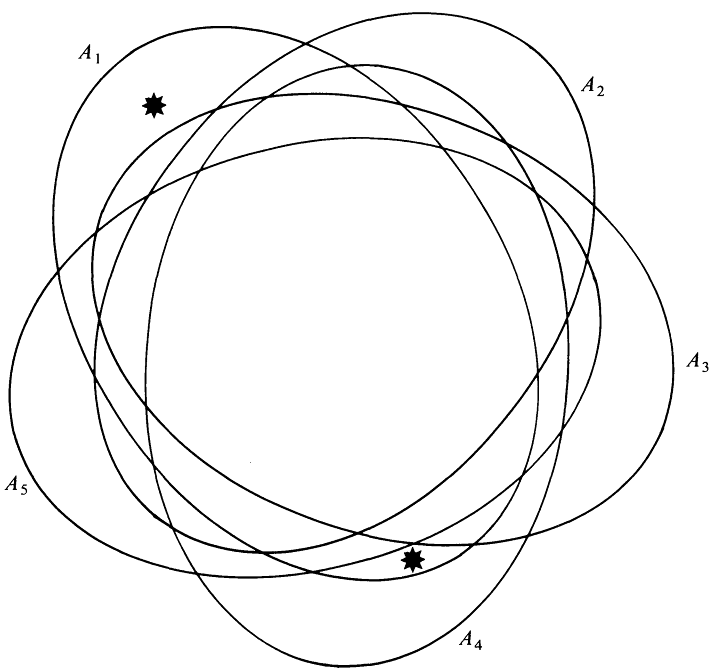


Figure 3. A simple and irreducible Venn diagram $\mathscr{V}$ which consists of five congruent ellipses. The stars indicate the two disjoint closed sets which form one of the regions in the independent family obtained by deleting from $\mathscr{V}$ the ellipse A_4.

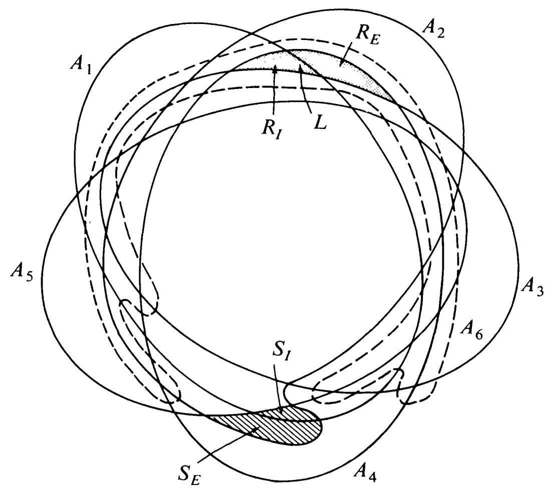


Figure 4. A simple and irreducible Venn diagram for six classes.

The family $\mathscr{V}$ of six curves in Figure 4 is derived from the family in Figure 3 by an obvious modification. It requires only a little patience to verify that $\mathscr{V}$ is an irreducible Venn diagram. If the curve deleted is A_j, where $1 \leqslant j \leqslant 5$, the fact that the remaining five curves form an independent family which is not a Venn diagram follows as for the family in Figure 3. But if A_6 is deleted, then $A_1^i \cap A_2^i \cap A_3^e \cap A_4^i \cap A_5^e$ consists of the disconnected region $R_I \cup S_I$, where $R_I(S_I)$ is the dotted (shaded) region of $\mathscr{V}$ interior to A_1. A similar result holds for $R_E \cup S_E$, where $R_E(S_E)$ is the dotted (shaded) region of $\mathscr{V}$ exterior to A_1. The curve A_6 (indicated by a dashed line) will play a special role in the general construction to be described next, as will the four regions R_I, R_E, S_I, S_E. For now it should be clear (letting A_1 and A_6 play the roles of B and C, respectively) that the family $\mathscr{V}$ of six curves in Figure 4 has property (**) below.

A Venn diagram $\mathscr{V}$ is said to have *property* (**), with respect to two of its curves B and C, if among the regions of $\mathscr{V}$ there are four—say R_I, R_E, S_I and S_E—so that the following holds:

(i) The regions R_I and S_I are interior to B, while R_E and S_E are exterior to B.

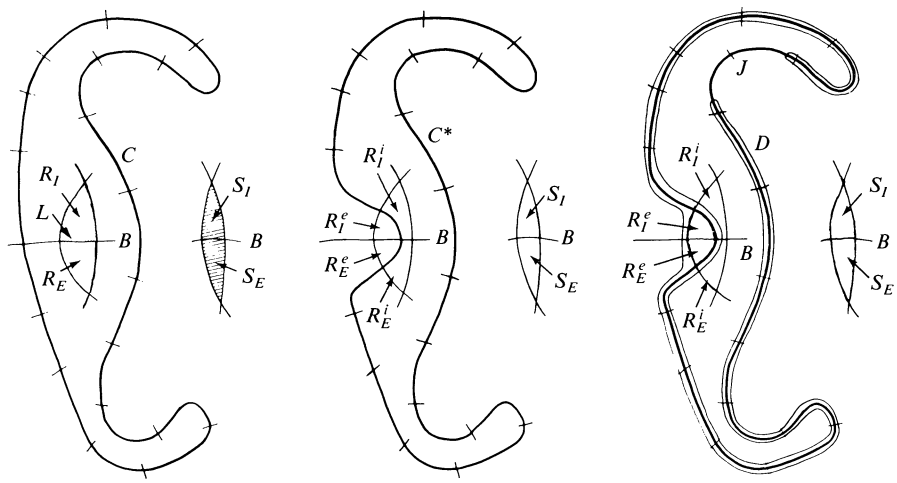
(ii) The common part of the boundaries of R_I and R_E is an arc of B, and so is the common part of the boundaries of S_I and S_E.

(iii) The regions R_I and R_E are not separated by any curve of $\mathscr{V}$ other than B. In other words, for any curve A of $\mathscr{V}$ other than B, the region R_I is interior to A if and only if R_E is interior to A. Similarly, S_I and S_E are separated by no curve of $\mathscr{V}$ other than B.

(iv) The boundary of every region of $\mathscr{V}$ except R_I, R_E, S_I and S_E contains an arc of the curve C. The regions R_I and R_E are interior to C, the regions S_I and S_E are exterior to C.

(v) The extension along B (in one direction) of the arc L of B, which separates R_I from R_E, meets C before it meets any other curve of $\mathscr{V}$ (not counting the one it meets at the endpoint of L).

To construct families of $n \geqslant 7$ curves satisfying the requirements of Theorem 1, we proceed by induction: assuming that we have a family of $n-1$ curves which has property (**) and satisfies Theorem 1, we shall obtain a family of n curves, with property (**), that satisfies Theorem 1.



(a) (b) (c)

Figure 5. Schematic illustration of the inductive step in the proof of Theorem 1.

Figure 5a illustrates the general situation, where a Venn diagram $\mathscr{V}$ of $n-1$ curves has property (**) with respect to curves B and C. In Figure 5b, we show the next step in the construction of the nth curve D. We first modify C to C^*, so that C^* crosses R_I and R_E. Then $R_I = R_I^i \cup R_I^e$ and $R_E = R_E^i \cup R_E^e$, where the superscripts "i" and "e" denote the parts interior and exterior to C^*. This modified family $\mathscr{V}^*$ (that is, $\mathscr{V}$ with C replaced by C^*) is not a Venn diagram, since the region $R_I^e \cup S_I$ is not connected. Figure 5c shows how to draw a curve D so that adding D to $\mathscr{V}^*$ will result in a Venn diagram $\mathscr{W}$:

(a) Along arcs where C and C^* coincide, D surrounds C^* sufficiently closely on both sides so as not to completely contain any of the regions of $\mathscr{V}^*$; the only exception to that happens near one of the points at which one of the other curves crosses C^* (such as the point J in Figure 5c), where each part of D inside C^* is connected to the part of D outside C^* to yield one closed curve.

(b) Along the arc where C and C^* differ, D follows C^* closely inside C^*, but outside C^*—where C^* cuts across R_I and R_E—the curve D does not follow C^* closely; instead, D completely encloses R_I^e and R_E^e.

Now it is easy to verify that the family of curves $\mathscr{W}$ is a simple Venn diagram, that it has property (**) with respect to the curves B and D, and that the deletion of any curve from $\mathscr{W}$ leads to an independent family which is not a Venn diagram. Indeed, for the curves other than C^* and D this follows from the analogous property of $\mathscr{V}$. Deletion of C^* yields an independent family which is not a Venn diagram because its region $R_I^i \cup S_I$ is a disjoint union. Similarly, the deletion of D yields the nonconnected region $R_E^e \cup S_E$.

This completes the proof of Theorem 1.

3. Venn Diagrams for k Outcomes and the Proof of Theorem 2.

Venn diagrams are based on the fact that a closed Jordan curve in the plane (or on the sphere) determines precisely two regions. Moreover, on the sphere these two regions are topologically equivalent. In the usual interpretation of Venn diagrams, these regions are made to correspond with the two truth values possible for the statement represented by the curve. Therefore, in order to obtain a natural generalization of Venn diagrams to the case in which k outcomes are possible (instead of just two), we should look for partitions of the sphere (that is, the surface of the sphere) into k sets which are topologically equivalent to each other. Then these k-partitions can be used to represent multivalued logic in the same way that usual logic is presented by the traditional Venn diagrams.

For certain values of k there are several ways of partitioning the sphere into k topologically equivalent regions; hence for such k a number of "kinds" of Venn diagrams could be considered. However, we shall concentrate on one method which works for all $k \geqslant 2$: the partition of the sphere into k regions by k "meridians" connecting the "north pole" to the "south pole", or any topologically equivalent partition. In other words, and switching for convenience of drawing from the sphere to the plane, we shall represent the k possible truth-value outcomes of one class A_j by a k-partition of the plane into the regions A_j^m $(m = 1, 2, \ldots, k)$ formed by k arcs which are simple and pairwise disjoint except for the two endpoints common to all k arcs (see Figures 6a and 7a for $k = 3$ and 4).

Taking such k-partitions to represent the various classes, a *Venn diagram* for the n classes A_j $(j = 1, 2, \ldots, n)$ is a family of k-partitions $\{A_j : j = 1, 2, \ldots, n\}$ of the plane such that:

(i) for every selection of $m_1, m_2, \ldots, m_n$, with each m_j satisfying $1 \leqslant m_j \leqslant k$, the intersection $\bigcap_{1 \leqslant j \leqslant n} A_j^{m_j}$ is a nonempty connected set; and

(ii) no point of the plane is on the boundary of regions of three or more different partitions A_j, nor is a "pole" of any partition on the boundary of any region of another partition.

For $k = 2$ this is clearly equivalent to the original definition in Section 1.

Examples of Venn diagrams are shown for $n = 1, 2, 3, 4$ with $k = 3$ in Figure 6 and for $n = 1, 2, 3$ with $k = 4$ in Figure 7.

We shall now briefly sketch an inductive proof of Theorem 2; first we deal with the case $k = 3$. The construction is very similar to the construction of ordinary Venn diagrams as given in [6], [2], [5], or many other places (see Figure 2). The main idea is to have the arcs, which define the regions of one partition, form a "circuit" that passes through all the regions determined by the other partitions. In Figures 6c and 6d we indicate how this can be done in a manner which obviously can be continued indefinitely, yielding the desired diagrams. The critical point of the construction is to make one of the three arcs of the partition which is being added very short, cross it by the analogous portion of the previous partition, and then follow—with the narrow "strand" formed by the two long arcs of the new partition—along one long arc of the previous partition while "weaving" across its other long arc.

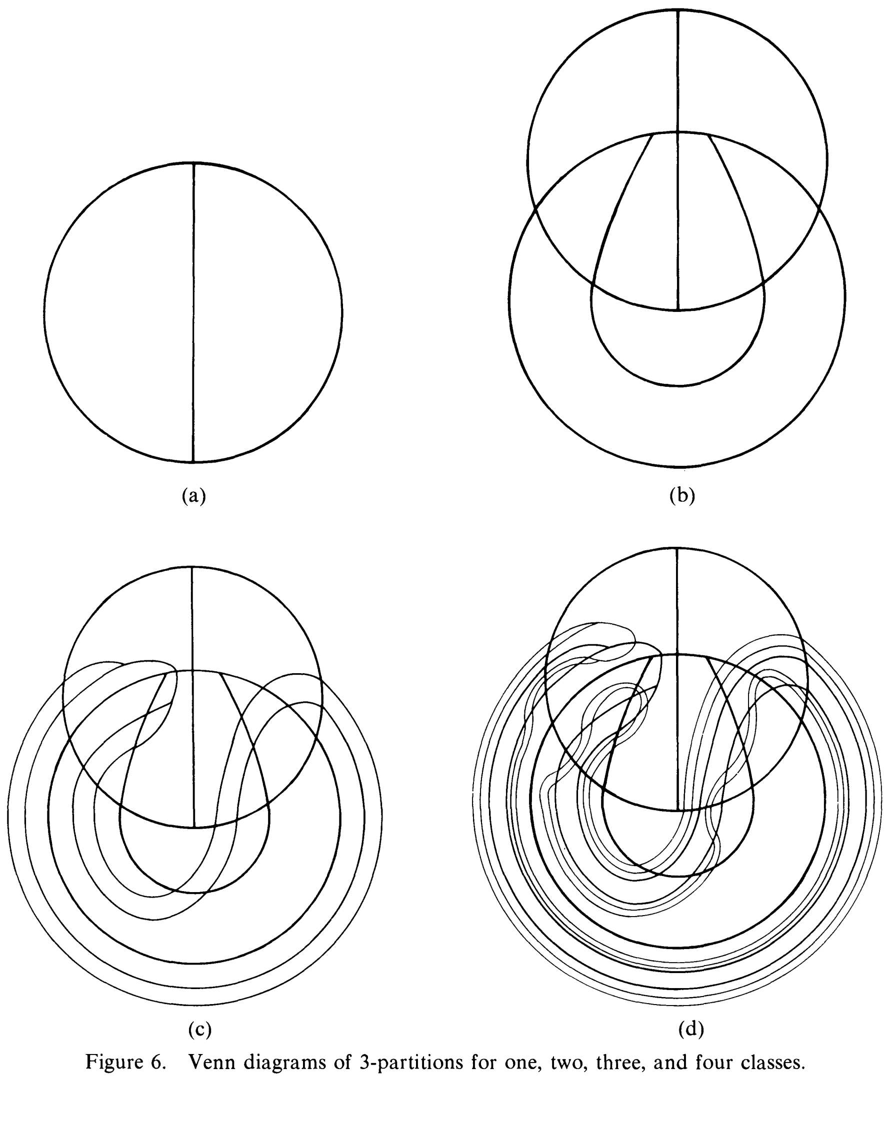
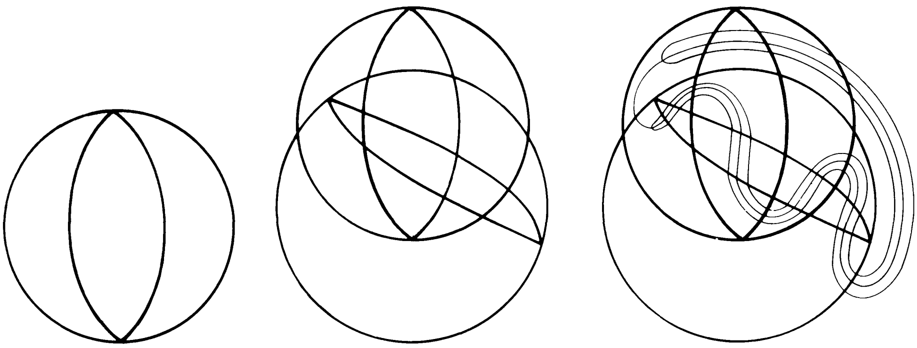


Figure 6. Venn diagrams of 3-partitions for one, two, three, and four classes.

The construction is very similar for $k \geqslant 4$, except that the "weaving" is across $k - 2$ arcs of the previous partition. The illustration in Figure 7c should suffice to convey the general idea.

This completes our sketch of the proof of Theorem 2.



(a) (b) (c)

Figure 7. Venn diagrams of 4-partitions for one, two, and three classes.

4. Remarks and problems. Several additional comments can be made.

(i) The example in Figure 3 is a minimal irreducible Venn diagram: each Venn diagram formed by 2, 3 or 4 curves contains at least one curve whose omission leads to a Venn diagram with fewer classes. While this is easily seen for two and three curves, no reasonable proof (except for an exhaustive case examination) is known for four curves.

(ii) It is not hard to verify that the Venn diagram for six classes shown in Figure 4 has the following property for $r = 2$: *Property* (N_r) Whenever any j curves $(1 \leqslant j \leqslant r)$ are deleted, the resulting independent family is not a Venn diagram.

Generalizing Nowicki's question solved by Theorem 1, one can ask whether for every $r \geqslant 2$ and for every sufficiently large n there exists Venn diagrams with n classes which have property (N_r). The answer seems to be affirmative, with the Venn diagrams constructed in the proof of Theorem 1 serving as examples whenever $n \geqslant r + 4$.

(iii) Still unsolved is the following question posed by P. Winkler (see [4]): Can every simple Venn diagram for n classes be extended to a Venn diagram for $n + 1$ classes by the addition of a suitable curve?

(iv) The only previous mention of Venn diagrams for multichoice classes seems to be a preprint [1], kindly supplied to the author by Professor S. W. Golomb. In [1] there is no insistence that the several regions of each partition be topologically equivalent to each other, and the approach also differs in other respects from the one followed here.

(v) In case $k = 4$, a reasonable representation of the four outcomes for one partition is by a tetrahedral partition of the sphere (or any partition homeomorphically equivalent to it); see Figures 8a and 8b for a representation in the plane. It is not hard to see how the proof of Theorem 2 can be modified to cover, for all n, this choice of partitions. It is only necessary (as shown in Figure 8c for $n = 2$) to slightly alter the previous partition near the branching points. Similarly, since the octahedral partition of the sphere (or the plane, see Figure 9a) can be homeomorphically distorted as shown in Figure 9b, this partition can also be used (for $k = 8$) to form Venn diagrams for arbitrarily large n, by slight changes in the above construction.

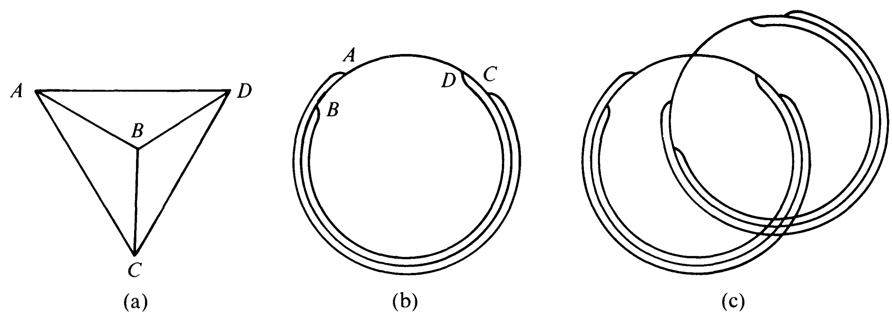


Figure 8. Construction of Venn diagrams using the tetrahedral 4-partition.

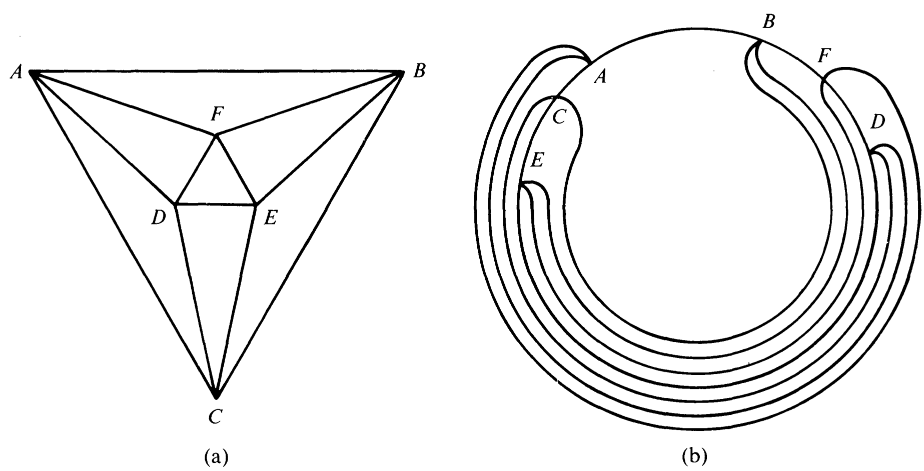


Figure 9. The octahedral 8-partition of the sphere, and the homeomorphic image of it which can be used to construct Venn diagrams.

In general, it is easy to verify that all k-partitions of the sphere by topological images of other isohedral polyhedra (that is, polyhedra in which all k faces are mutually equivalent under symmetries of the polyhedron—for example, the cube, regular dodecahedron, rhombic dodecahedron, icosahedron, etc.; see [3] for a description of all possible types) also lead to Venn diagrams for the appropriate values of k. One approach is to note that the above method depends essentially only on the fact that the branching points (vertices) can be assigned to two classes so that among the edges of each of the regions (faces) precisely two have endpoints in different classes. In Figure 10 this is illustrated for the cube. The reader may wish to furnish the detailed proof in this case, and also to show that appropriate "colorings" of the vertices are possible for all isohedral polyhedra.

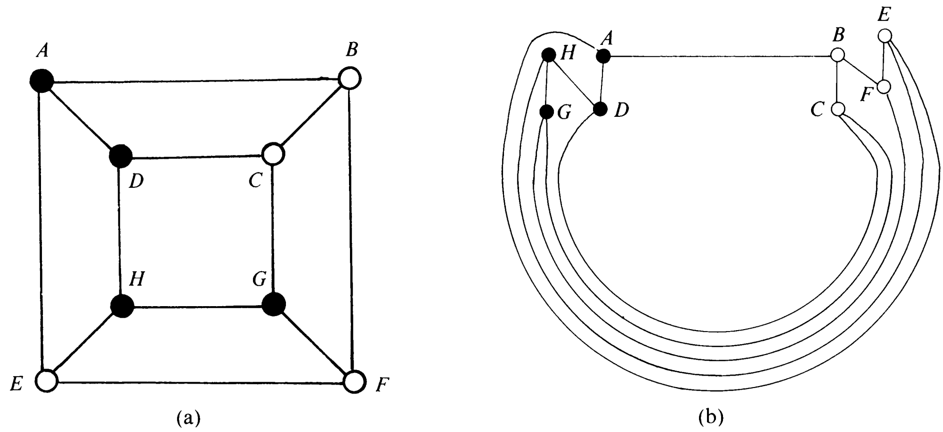


Figure 10. The cubical 6-partition of the sphere, and its homeomorphic image which can be used to construct Venn diagrams. Note that all vertices of the same "color" are "bunched together", and correspond to the two "poles" in the construction described in Section 3.

REFERENCES

1. S. W. Golomb, Connected logic diagrams, unpublished note, received by the author in February 1974.
2. B. Grünbaum, Venn diagrams and independent families of sets, Math. Mag., vol. 48 (1975), 12–22.
3. B. Grünbaum and G. C. Shephard, Spherical tilings with transitivity properties, In "The Geometric Vein: The Coxeter Festschrift," C. Davis et al., eds. Springer-Verlag, New York 1982, pp. 65–98.
4. B. Grünbaum and P. Winkler, A Venn diagram of 5 triangles, Math. Mag., vol. 55 (1982), 311.
5. P. Nowicki, Koniczynka n-listna, [In Polish] Wiadom. Mat., vol. 19 (1975), 11–18.
6. J. Venn, On the diagrammatic and mechanical representation of propositions and reasonings, The London, Edinburgh, and Dublin Philos. Mag. and J. Sci., vol. 9 (1880), 1–18.
7. J. Venn, Symbolic Logic, Macmillan, London 1881, second ed., 1894.

At its January 1984 annual meeting, in Louisville, the Mathematical Association of America announced a special citation honoring those who have furthered the progress of mathematics by significantly enhancing the status of women in mathematics. The full text of this citation appears in the March–April issue of FOCUS, the MAA newsletter.