Graph Theory in Venn Diagrams

Jacky567

June 3, 2025

Abstract

Venn diagram is a widely used diagram style that shows the logical relation between sets, popularized by John Venn in the 1880s. In this work, we will discuss about how to create a Venn diagram for n sets using graph theory, and some of its applications.

Contents

1 Introduction

Def 1.1. n-Venn diagram $\mathbf{C}$ is a collection of n simple closed curves $\{C_i\}$ in the plane that intersect in only finitely many points and create exactly 2^n regions, one for every possible combination of being inside or outside of each of the n curves

Exam 1.1. *the most common Venn diagram is the* 3*-Venn diagram with three unit circles:*

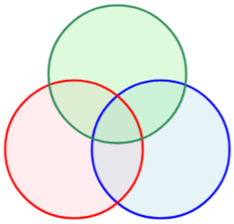


Figure 1: 3-Venn diagram

Naturally, we would wonder whether n-Venn diagram can be given by n circles. Unfortunately, this is wrong for $n \geqslant 4$.

Prop 1.1. *n circles in the plane cannot divide the plane into 2^n regions when $n \geqslant 4$.*

Proof. Notice that every two circles have at most 2 intersections.

Let m_i be the number of intersections of i-th circle and all other circles.

We first consider the case that at most two of the n circles intersect in any point.

Then i-th circle is divided into m_i edges and there are totally $\frac{1}{2}\sum\limits_{i=1}^{n} m_i$ points.

So by Euler formula, the number of regions is $\leqslant 2 + \sum\limits_{i=1}^{n} m_i - \frac{1}{2}\sum\limits_{i=1}^{n} m_i = \frac{1}{2}\sum\limits_{i=1}^{n} m_i + 2$.

And since $m_i \leqslant 2(n-1)$.

Therefore there are at most $n(n-1)+2$ regions.

Now for the general case, consider a point that is intersections of k circles.

Then under a perturbation of one of these k circles, the number of points increase by 2 and the number of edges increase by 3.

So the number of regions increase by 1 and every two curves still have at most 2 intersections.

After some perturbations, we can reduce to the first case and the number of regions increase.

Hence there are at most $n(n-1)+2$ regions, which is less than 2^n when $n \geqslant 4$. $\square$

Def 1.2. A Venn diagram is simple if at most two of the n curves intersect in any point.

Exam 1.2.

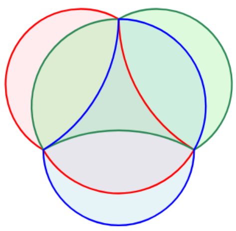


Figure 2: Non-simple 3-Venn diagram

By the above proposition, we have:

Prop 1.2. *a simple n-Venn diagram has exactly $2^n - 2$ intersections and $2^{n+1} - 4$ edges.*

While we cannot draw an n-Venn diagram by n circles, we can give the $4, 5$-Venn diagram by ellipses since two ellipses can have 4 intersections:

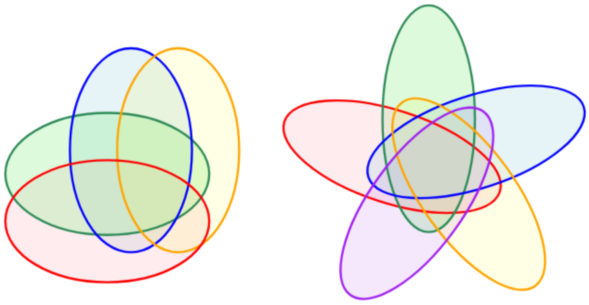


Figure 3: Venn diagram with $n = 4, 5$ epllipses

2 Dual of Venn diagram

Def 2.1. For a Venn diagram $\mathbf{C}$, the Venn dual $D(\mathbf{C})$ is the planar dual of the Venn diagram: its vertices are the region of $\mathbf{C}$ and two vertices are connected by an edge if they are adjoin.

Prop 2.1. *$\mathbf{C}$ is simple $D(\mathbf{C})$ has the following properties:*

(1) It is a subgraph of n-dimensional hypercube Q_n with 2^n vertices.

(2) It is a planar graph and every face is a 4-cycle.

(3) The subgraph $D(\mathbf{C}) \cap H$ is connected, where H is any hypersurface of Q_n.

Proof. (1) For every vertex in $D(\mathbf{C})$, we can give it an length-n bitstring, where 0 and 1 represent that the corresponding region is outside and inside the curve resp.

And for every edge, the corresponding regions of the vertices are adjoin.

So they are differ in exactly one bit position.

Hence $D(\mathbf{C})$ is a subgraph of Q_n, and has 2^n vertices.

(2) Since every intersection(vertex) in $\mathbf{C}$ is given by exactly two curves.

So its degree is 4, *i.e.* every face in $D(\mathbf{C})$ is a 4-cycle.

(3) WLOG, assume $H = \{(x_1, \cdots, x_n) | x_1 = 0\}$.

Then $C(\mathbf{C}) \cap H$ corresponds to all region that are not contained in C_1, *i.e.* the outside region of C_1.

So it is connected since C_1 is a simple closed curve.

□

Prop 2.2. *Conversely, if a graph G satisfying the above three properties, then its dual is a simple n-Venn diagram.*

Proof. Let $\mathbf{C}$ be its dual.

Define $C_i = \partial H_i$ where H_i is the union of regions corresponding to points in $\{(x_1, \cdots, x_n) | x_i = 1\} \subset V(G)$.

By property (3), H_i is connected, *i.e.* C_i is a closed simple curve.

And since for any edge e in $G \subset Q_n$, its vertices are differ in some x_i.

So one of their corresponding faces in $\mathbf{C}$ is contained in H_i and the other is not.

Therefore their common edge, which is the corresponding edge of e, is in C_i.

Hence $\mathbf{C}$ is a Venn diagram.

Moreover, degree of every vertex is 4 since every face in G is a 4-cycle, *i.e.* $\mathbf{C}$ is simple. □

Def 2.2. We called the graph that satisfying the above the above three properties n-Venn quadrangulation.

Exam 2.1. *Here is an example of a Venn diagram and its dual graph:*

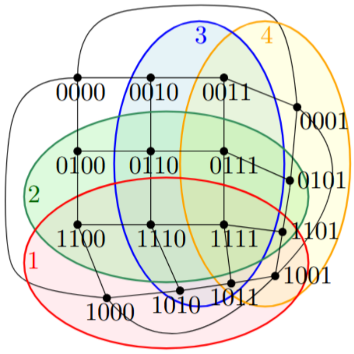


Figure 4: A 4-Venn diagram and its dual graph

3 Construction of Venn diagram

Obviously, we consider this question inductively. Suppose we now have an n-Venn diagram $\mathbf{C}$, how can we add a curve C_{n+1} on it so that it becomes an $(n+1)$-Venn diagram? Notice that C_{n+1} must enter and leave every region of $\mathbf{C}$ exactly once and divided every region into two parts. So finding such curve is the same as finding a Hamilton cycle in the dual graph $D(\mathbf{C})$.

Exam 3.1. *For* 4*-Venn quadrangulation in fig. 4, we can find a Hamilton cycle and extend the corresponding* 4*-Venn diagram to* 5*-Venn diagram.*

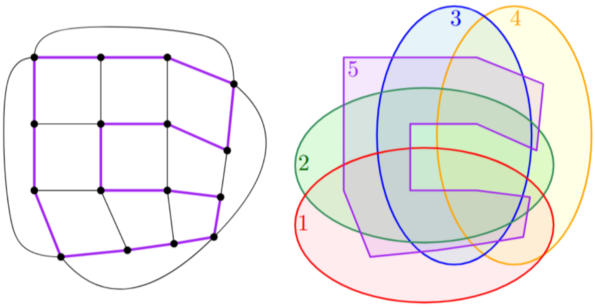
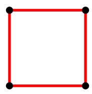


Figure 5: a Hamilton cycle and the 5-Venn diagram obtained from fig. 4

To find a Hamilton cycle in $D(\mathbf{C})$, we first consider a Hamilton cycle in Q_n.

For $n = 2$, let C_2 be a Hamilton cycle given by:

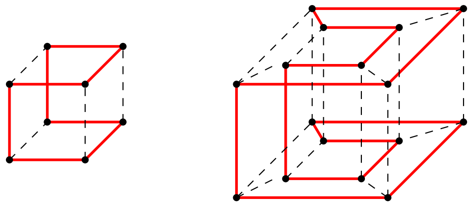


Now assume we have constructed a Hamilton cycle C_n of Q_n, let $C_n = [v_1\ v_2\ \cdots\ v_{2^n}]$ with $v_1 = (0, \cdots, 0)$.

Define $C_{n+1} = [(v_1, 0)\ \cdots\ (v_{2^n}, 0)\ (v_{2^n}, 1)\ (v_{2^n-1}, 1)\ \cdots\ (v_1, 1)]$.

Then C_{n+1} is obviously a Hamilton cycle in Q_{n+1} and the front half vertices in C_{n+1} is 0 on the last component, the back half vertices in C_{n+1} is 1 on the last component.

Exam 3.2. *Here is the figure of C_3, C_4:*



Prop 3.1. *For every n, there exists an n-Venn quadrangulation containing C_n such that every edges $(x_1, \cdots, x_{n-1}, 0)$ to $(x_1, \cdots, x_{n-1}, 1)$ is contained in G_n.*

Proof. When $n = 2$, $G_2 = Q_2 = C_2$.

Assume G_n is an n-Venn quadrangulation containing C_n such that every edges $(x_1, \cdots, x_{n-1}, 0)$ to $(x_1, \cdots, x_{n-1}, 1)$ is contained in G_n.

Let $e_1, \cdots, e_p, f_1, \cdots, f_q$ be the edges in $G_n \backslash C_n$, where the n-th component of vertices of e_i are the same and those of f_i are different.

Then $q = 2^{n-1} - 2$ and $p + q = 2^{n+1} - 4 - 2^n = 2^n - 4$ by proposition 1.2.

So $p = q = 2^{n-1} - 2$.

Construct $G_{n+1} = (G_n \backslash \{e_i\}) \times \{0\} \cup (G_n \backslash \{f_i\}) \times \{1\} \cup \{(x_1, \cdots, x_n, 0) \leftrightarrow (x_1, \cdots, x_n, 1)\}$.

Then $C_n \times \{0\}, C_n \times \{1\} \subset G_{n+1}$ and $|V(G_{n+1})| = 2^{n+1}, |E(G_{n+1})| = 2|E(G_n)| - p - q + 2^n = 2^{n+2} - 4$.

It remains to prove the properties $(2), (3)$ in proposition 2.1.

We first assume that we have proven G_{n+1} is planar.

Since Q_{n+1} is a bipartite graph(divided by the parity of $x_1 + \cdots + x_{n+1}$).

So G_{n+1} has no 3-cycle, *i.e.* every faces has at least 4 edges.

And by Euler formula, the faces of G_{n+1} is $2 - 2^{n+1} + 2^{n+2} - 4 = 2^{n+1} - 2 = \frac{|E(G_{n+1})|}{2}$.

Therefore every face must be 4-cycle since every edge is contained in at most 2 faces.

Now consider a hypersurface $H = \{(x_1, \cdots, x_{n+1}) | x_i = \varepsilon\}$ with a constant $\varepsilon \in \{0, 1\}$.

If $i = n + 1$, then $G_n \times \{\varepsilon\} \subset G_{n+1} \cap H$ is connected.

If $i \leqslant n$, let $H' = \{(x_1, \cdots, x_n) | x_i = \varepsilon\}$.

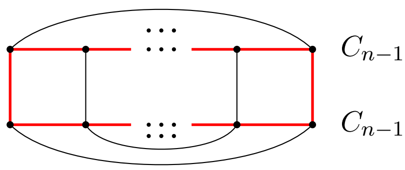
Since there are edges between $(x_1, \cdots, x_n, 0)$ and $(x_1, \cdots, x_n, 1)$.

So we can pairwisely identify them together, and then the graph we get is $((G_n \backslash \{e_i\}) \cap H') \cup ((G_n \backslash \{f_i\}) \cap H') = G_n \cap H'$, which is connected by the induction assumption.

Therefore $G_{n+1} \cap H$ is always connected.

Finally, it remains to prove that G_{n+1} is planar.

Moreover, we inductively prove that G_n looks like:



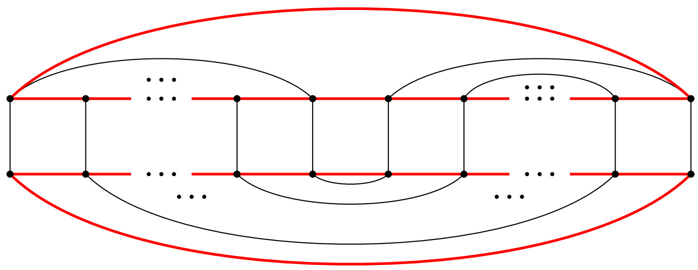
Such that the half above part corresponds to $G_n \cap \{(x_1, \cdots, x_n) | x_n = 0\}$, the half below part corresponds to $G_n \cap \{(x_1, \cdots, x_n) | x_n = 1\}$ and the large loops above and below are both C_{n-1}.

Then C_n is the red loop in the diagram.

Edges $\{e_i\}$ are all extra edges above or below the "ladder".

And edges $\{f_i\}$ are all vertical lines in the middle except two red lines.

So G_{n+1} looks like:



$\square$