Drawing Area-Proportional Venn-3 Diagrams with Convex Polygons

Peter Rodgers[1], Jean Flower[2], Gem Stapleton[3], and John Howse[3]

[1] University of Kent, Canterbury, UK
p.j.rodgers@kent.ac.uk
[2] Autodesk
[3] Visual Modelling Group, University of Brighton, UK
{g.e.stapleton,john.howse}@brighton.ac.uk

Abstract. Area-proportional Venn diagrams are a popular way of visualizing the relationships between data sets, where the set intersections have a specified numerical value. In these diagrams, the areas of the regions are in proportion to the given values. Venn-3, the Venn diagram consisting of three intersecting curves, has been used in many applications, including marketing, ecology and medicine. Whilst circles are widely used to draw such diagrams, most area specifications cannot be drawn in this way and, so, should only be used where an approximate solution is acceptable. However, placing different restrictions on the shape of curves may result in usable diagrams that have an exact solution, that is, where the areas of the regions are exactly in proportion to the represented data. In this paper, we explore the use of convex shapes for drawing exact area proportional Venn-3 diagrams. Convex curves reduce the visual complexity of the diagram and, as most desirable shapes (such as circles, ovals and rectangles) are convex, the work described here may lead to further drawing methods with these shapes. We describe methods for constructing convex diagrams with polygons that have four or five sides and derive results concerning which area specifications can be drawn with them. This work improves the state-of-the-art by extending the set of area specifications that can be drawn in a convex manner. We also show how, when a specification cannot be drawn in a convex manner, a non-convex drawing can be generated.

1 Introduction

Area-proportional Venn diagrams, where the areas of the regions formed from curves are equal to an area specification, are widely used when visualizing numerical data [6, 8, 7]. An example of such a diagram is shown in Figure 1, adapted from [10]. It shows the intersections of physician-diagnosed asthma, chronic bronchitis, and emphysema within patients who have obstructive lung disease. Here, the areas only approximate the required values, which are given in the regions of the diagram, as an exact solution does not exist. In fact, when using three circles to represent an area specification there is almost certainly no exact area-proportional Venn-3 diagram [3].

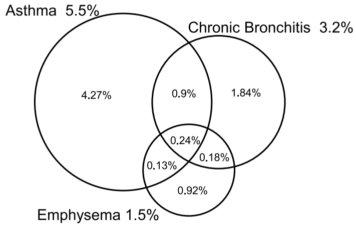

Fig. 1. An approximate area-proportional Venn-3 diagram.

There have been various efforts towards developing drawing methods for area-proportional Venn-3 diagrams but their limitations mean that the vast majority of area specifications cannot be drawn automatically, even for this seemingly simple case. Chow and Rodgers devised a method for producing approximate Venn-3 diagrams using only circles [3], and the Google Charts API includes facilities for drawing approximate area-proportional Euler diagrams with at most three circles, including Venn-3 [1]. The approximations are often insufficiently accurate to be helpful and can be misleading. Recent work, which is limited and not practically applicable, establishes when an area specification can be drawn by a restricted class of symmetric, convex Venn-3 diagrams [9]. Extending to non-symmetric cases is essential to ensure practical applicability.

More generally, other automated area-proportional drawing methods include that by Chow, which draws so-called monotone diagrams (including Venn-3) [5]; usability problems arise because the curves are typically non-convex. It is unknown which area specifications are drawn with convex curves by Chow's method, although we note that attaining convexity was not a primary concern. Other work, by Chow and Ruskey [4], produced rectilinear diagrams, further studied in [2]. These are diagrams drawn with curves whose segments can only be horizontal or vertical, hence have a series of 90 degree bends. However, in general, rectilinear layouts of Venn-3 require non-convex curves. Diagrams drawn with convex curves are more likely to possess desirable aesthetic qualities, with reduced visual complexity, and in addition, the most desirable shapes (such as circles, ovals and rectangles) are convex.

This paper analyzes area-proportional Venn-3 diagrams drawn with convex polygons, with four main contributions: (a) a classification of some area specifications that can be drawn using convex polygons, (b) construction methods to draw a convex diagram, where the area specification has been identified as drawable in this manner, (c) a method to draw the area specification when these methods fail, thus ensuring that every area specification can be drawn, and (d) a freely available software tool for drawing these diagrams; see *www.cs.kent.ac.uk/people/staff/pjr/ConvexVenn3/diagrams2010.html*. Section 2 gives preliminary definitions, of Venn-3 diagrams and related concepts, as well as presenting some required linear algebra concepts. In Section 3, we define several classes of Venn-3 diagrams that allow us to investigate which area specifications can be drawn using convex curves. We also demonstrate (analytical) methods to draw such diagrams. Finally, Section 4 describes our software implementation, including details of drawing methods that rely on numerical approaches when our analytical drawing methods cannot be applied.

2 Venn Diagrams and Shears of the Plane

The labels in Venn diagrams are drawn from a set, $\mathcal{L}$. Given a closed curve, c, the set of points interior to c will be denoted $int(c)$. Similarly, the set of those points exterior to c is denoted $ext(c)$.

Definition 1. *A* ***Venn diagram****, $d = (Curve, l)$, is a pair where*

1. *$Curve$ is a finite set of simple, closed curves in $\mathbb{R}^2$,*
2. *$l\colon Curve \to \mathcal{L}$ is an injective labelling function that assigns a label to each curve, and*
3. *for each subset, C, of $Curve$, the set*

$$\bigcap_{c \in C} int(c) \cap \bigcap_{c \in Curve - C} ext(c)$$

   is non-empty and forms a simply connected component in $\mathbb{R}^2$ called a ***minimal region*** *and is* ***described*** *by $\{l(c) : c \in C\}$.*

If $|Curve| = n$ then d is a ***Venn-n diagram****. If at most two curves intersect at any given point then d is* ***simple****. If each curve in d is convex then d is* ***convex****.*

We focus on the construction of simple, convex Venn-3 using only polygons (recall, a polygon is a closed curve). Therefore, we assume, without loss of generality, that $\mathcal{L} = \{A, B, C\}$. Thus, the set $R = \mathbb{P}\{A, B, C\}$ is the set of all minimal region descriptions for Venn-3. Given a Venn-3 diagram, we will identify minimal regions in the diagram by their descriptions. For instance, the minimal region inside all 3 curves is properly described by $\{A, B, C\}$, but we will abuse notation and write this as ABC. A minimal region inside curves labelled A and B but outside C will, therefore, be identified by AB. Furthermore, we will often blur the distinction between the minimal region and its description, simply writing ABC to mean the region inside A, B and C. We call the minimal region ABC the **triple intersection**, AB, AC and BC are **double intersections**, and A, B, and C are **single intersections**. A **region** in d is a (not necessarily connected) component of $\mathbb{R}^2$ that is a union of minimal regions. The region formed by the union of ABC and AB is, therefore, described by $\{ABC, AB\}$. The region $\{AB, AC, BC, ABC\}$ is called the **core** of d [2].

In order to provide a construction of an area-proportional Venn-3 diagram, we need to start with a specification of the required areas. Our definition of an area specification allows areas to be zero or negative; allowing for non-positive areas is for convenience later in the paper when we present methods for drawing Venn-3 diagrams.

Definition 2. *An* ***area specification*** *is a function, area*: $R \to \mathbb{R}$, *that assigns an area to each minimal region description. Given a Venn-3 diagram,* $d = (Curve, l)$, *if, for each* $r \in R$, *the area of the minimal region, mr, described by* r, *is* $area(r)$ *(that is,* $area(r) = area(mr)$*) then* d ***represents*** *area*: $R \to \mathbb{R}$.

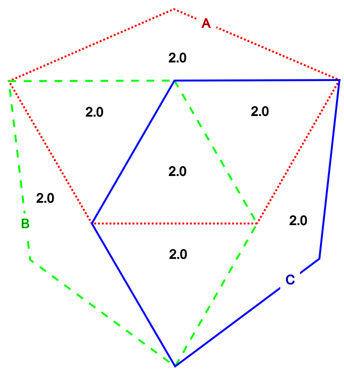

Fig. 2. An area-proportional Venn-3 diagram.

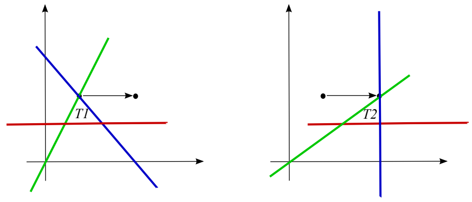

Fig. 3. A shear of the plane.

For example, the diagram in Figure 2 represents the area specification given by the areas written inside the minimal regions and was produced using our software. For convenience and readability, we will further blur the distinction between regions, their descriptions, and their areas. For instance, we will take the region description ABC to also mean either the minimal region it is describing or the area of that region; context will make the meaning clear.

Given a Venn-3 diagram, we might be able to apply a transformation to it that converts it to a form that is more easily analyzable, whilst maintaining the areas. Since we are constructing diagrams with polygons, we can apply a type of linear transformation, called a shear, to a diagram that alters its appearance but maintains both convexity and the areas of the regions. A shear of the plane can be seen in Figure 3, which keeps the x-axis fixed and moves the point as indicated by the arrow. The effect of the transformation can be seen in the righthand diagram. The two triangles $T1$ and $T2$ have the same area.

Definition 3. *A shear of the plane,* $\mathbb{R}^2$, *is a linear transformation defined by fixing a line,* $ax + by = c$, *and moving some other point,* (p, q), *some distance,* d, *parallel to the line.*

For example, if we keep the x-axis fixed and the point $(0, 1)$ moves to $(1, 1)$ (here, $d = 1$) then each point, (x, y), maps to $(x + y, y)$.

Lemma 1. *Let* l *be a shear of the plane. Then* l *preserves both lines and areas.*

As a consequence of Lemma 1, we know that, under a shear of the plane, any triangle maps to another triangle and that any convex polygon remains convex.

3 Drawing Convex Venn-3 Diagrams

To present our analysis of area specifications that are drawable by convex, Venn-3 diagrams, we introduce several classes of Venn-3 diagram. These classes are characterized by the shapes of (some of) the regions in the diagrams. The first two classes allow us to identify some area specifications as drawable by a convex Venn-3 diagram and, moreover, we demonstrate how to draw such a diagram. The remaining two classes allow more area specifications to be drawn.

For all the diagrams in this section we can choose to use an equilateral triangle for the triple intersection as any triangle can be transformed into an equilateral triangle by applying two shears.

3.1 Core-Triangular Diagrams

To define our first class of diagram, called the core-triangular class, we require the notion of an inscribed triangle: a triangle, T_1, is **inscribed** inside triangle T_2 if the corners of T_1 lie on the edges of T_2.

Definition 4. *A Venn-3 diagram is **core-triangular** if it is convex and*

1. *the core is triangular,*
2. *ABC is a triangle inscribed inside the core, and*
3. *the regions A, B and C are also triangles.*

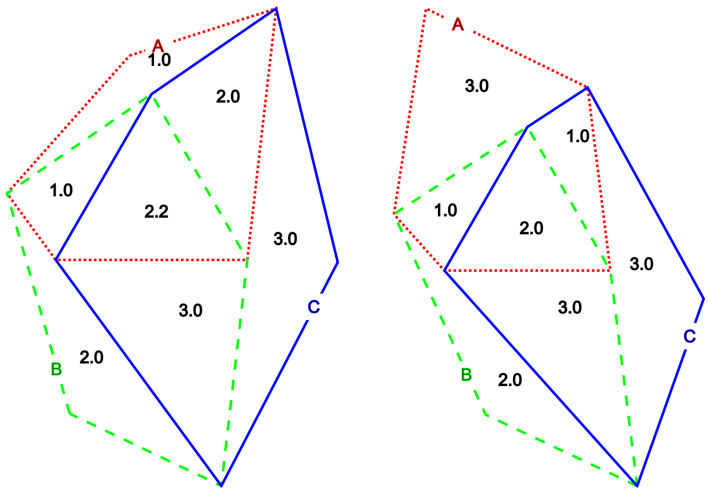

Fig. 4. Core-triangular diagrams.

For example, the diagrams in Figure 4 are core-triangular. We observe that, in any core-triangular diagram, the regions AB, AC, and BC are also triangles.

We can derive a relationship between the sum of the double and triple intersection areas and a product involving these areas that establishes whether an area specification is drawable by a diagram in this class. The derivation relies on an analysis of the geometry of these diagrams, but for space reasons we omit the details.

Theorem 1 (Representability Constraint: Core-Triangular). *An area specification, $area \colon R \rightarrow \mathbb{R}^+ - \{0\}$, is representable by a core-triangular diagram if and only if*

$$AB + AC + BC + ABC \geq 4 \times \left(\frac{AB}{ABC} \times \frac{AC}{ABC} \times \frac{BC}{ABC} \right) \times ABC.$$

To illustrate, the area specifications as illustrated in Figure 4 satisfy the inequality in Theorem 1. However, any area specification with $AB = AC = BC = 2$ and $ABC = 1$ is not representable by a core-triangular diagram, since

$$2 + 2 + 2 + 1 = 7 \geq 4 \times 2 \times 2 \times 2 \times 1 = 32$$

is false.

Core-triangular diagrams form a sub-class of our next diagram type, triangular diagrams, in which all minimal regions are triangles but the core need not be triangular. We provide many more details for the derivation of a representability constraint for triangular diagrams, with that for core-triangular diagrams being a special case. Moreover, we provide a method for drawing triangular diagrams which can be used to draw core-triangular diagrams.

3.2 Triangular Diagrams

The diagrams in Figure 5 are triangular. As with core-triangular diagrams, we also identify exactly which area specifications can be drawn with triangular diagrams.

Definition 5. *A Venn-3 diagram is **triangular** if it is convex and all of the minimal regions are triangles.*

If an area specification can be represented by a diagram in the triangular class then it can be represented by such a diagram where the inner-most triangle, ABC, is rightangled (Theorem 2 below). We use this insight to establish exactly which area specifications are representable by triangular diagrams.

Theorem 2. *Let d_1 be a triangular diagram. Then there exists a triangular diagram, d_2, such that the region ABC in d_2 is an rightangled triangle with the two edges next to the rightangled corner having the same length and d_1 and d_2 both represent the same area specification.*

Proof. We can apply two shears in order to obtain d_2. Assume, without loss of generality, that A is located at $(0, 0)$ and that B does not lie on either axis.

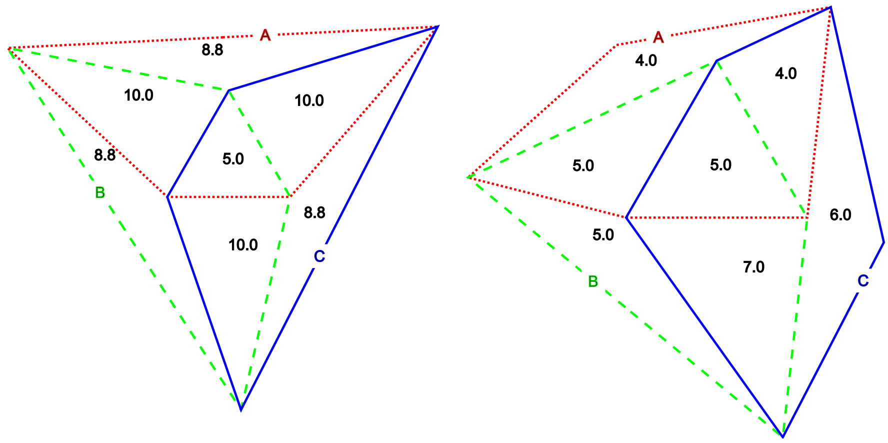

Fig. 5. Triangular diagrams.

Apply a shear such that A is fixed and B maps to $B' = (\sqrt{2 \times area}, 0)$ where $area$ is the area of ABC. Define C' to be the point to which C maps under this shear. The triangle $AB'C'$ has the same area as the triangle ABC, so C' is at $(\lambda, \sqrt{2 \times area})$ for some λ. The second shear fixes the line AB' and maps C' to $C'' = (0, \sqrt{2 \times area})$. The final triangle $AB'C''$ has a rightangle at A and the adjacent sides both of length $\sqrt{2 \times area}$. Since we have applied shears, we remain in the triangular class, by Lemma 1, as required.

We can, therefore, assume that the region ABC is a rightangled triangle, with the 90-degree corner at $(0, 0)$, as indicated in Figure 6 which shows a partially drawn Venn-3 diagram. Given such a drawing of ABC, we can determine three lines, each parallel with one of the edges of ABC, the distance from which is determined by the required area of the double intersections. The triangles AB, AC and BC each have a vertex on one of these lines. The location of the vertex is constrained, since we must ensure convexity. For example, AB in Figure 6 must have a vertex lying between the points $\gamma = 0$ (if negative, the B curve becomes non-convex) and $\gamma = 1$ (if bigger than 1, the A curve becomes non-convex). When attempting to construct a triangular diagram for a given area specification, our task is to find a suitable α, β and γ.

Now, the area of a triangle can be computed via the determinant of a matrix: a triangle, T, with vertices at (x_1, y_1), (x_2, y_2), and (x_3, y_3) has area

$$area(T) = \frac{\det}{2} \begin{pmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}.$$

In what follows, we assume that two sides of each of the triangles A, B and C are formed from edges of the double intersections, discussing the more general

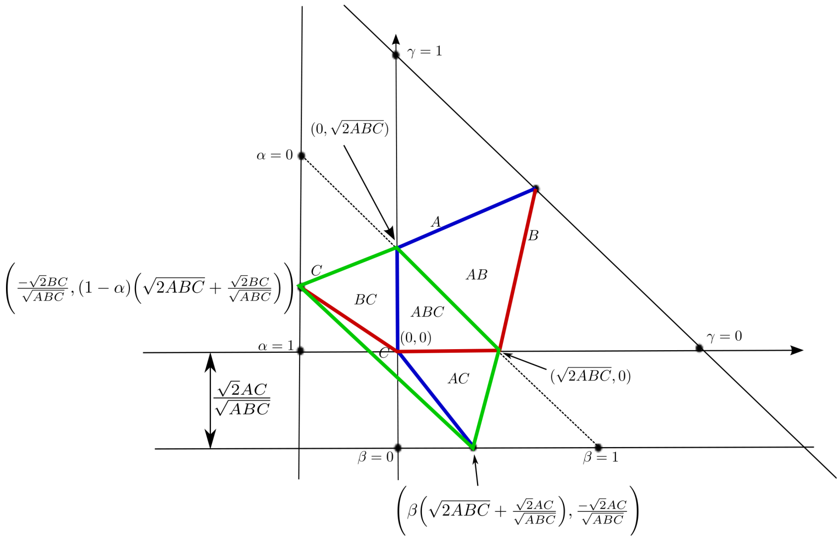

Fig. 6. Deriving area constraints.

case later. The triangle C has area:

$$C = \frac{\det}{2}\begin{pmatrix} 1 & 1 & 1 \\ 0 & \frac{-\sqrt{2}BC}{\sqrt{ABC}} & \beta\left(\sqrt{2ABC} + \frac{\sqrt{2}AC}{\sqrt{ABC}}\right) \\ 0 & (1-\alpha)\left(\sqrt{2ABC} + \frac{\sqrt{2}BC}{\sqrt{ABC}}\right) & \frac{-\sqrt{2}AC}{\sqrt{ABC}} \end{pmatrix}$$
$$= \frac{AC \times BC}{ABC} - \frac{(1-\alpha)\beta}{ABC}\Big(ABC + BC\Big)\Big(ABC + AC\Big)$$

Rearranging the above, setting $X = (1-\alpha)\beta$, we have:

$$X = (1-\alpha)\beta = \frac{AC \times BC - C \times ABC}{(AC + ABC)(BC + ABC)} \tag{1}$$

Similarly, using the triangles A and B, we can deduce

$$Y = (1-\beta)\gamma = \frac{AB \times AC - A \times ABC}{(AB + ABC)(AC + ABC)} \tag{2}$$

and

$$Z = (1-\gamma)\alpha = \frac{BC \times AB - B \times ABC}{(BC + ABC)(AB + ABC)}. \tag{3}$$

Thus, we have three equations, (1), (2) and (3), with three unknowns, α, β and γ, from which we can derive a quadratic in α:

$$(1-Y)\alpha^2 + (X + Y - Z - 1)\alpha + Z(1-X) = 0.$$

This has real solutions provided the discriminant, $(1-X-Y-Z)^2 - 4XYZ$, is not negative. Once we have found α, we can then compute β and γ. If solvable, with all of α, β and γ between 0 and 1 (for convexity; see Figure 6), then the area specification is representable by a triangular, convex diagram. Such a solution is called **valid**.

To illustrate, consider the area specification for the lefthand diagram in Figure 5. We have

$$\begin{aligned} X = (1-\alpha)\beta &= \frac{AC \times BC - C \times ABC}{(BC+ABC)(AC+ABC)} \\ &= \frac{10 \times 10 - 8.8 \times 5}{(10+5)(10+5)} \\ &= \frac{56}{225} \approx 0.25 \end{aligned}$$

Similarly, $Y = (1-\beta)\gamma = (1-\gamma)\alpha \approx 0.25$ which has solutions $\alpha = \beta = \gamma \approx 0.5$.

As previously stated, the algebra above assumes that two sides of each of the triangles A, B and C are formed from edges of the double intersection areas. For the general case, an area specification, $area$, is representable by a triangular diagram if there is a valid solution to (1), (2) and (3) with any of the single intersection areas reduced to zero. If we can determine drawability using the above method when some of these single intersection areas are reduced to 0 then we can enlarge those areas to produce a diagram with the required specification. For example, in Figure 5, the area specification for the righthand diagram would be deemed undrawable unless we reduce the A and C areas to zero before following the above process; without altering the area specification, there is no solution with all of α, β and γ between 0 and 1. Taking the area specification, $area'$, equal to area but with $area'(A) = area'(C) = 0$, we find a valid solution with $\alpha = 0.65$, $\beta = 0.75$ and $\gamma = 0.87$. Using this solution, we can draw all of the diagram except A and C. As a post process, we add triangles for A and C to give a diagram as shown. The arguments we have presented establish the following result:

Theorem 3 (Representability Constraint: Triangular). *An area specification, $area\colon R \to \mathbb{R}^+ - \{0\}$, is representable by a triangular diagram if and only if there exists another area specification, $area'\colon R \to \mathbb{R}^+$ which is the same as area, except that some of the single intersections may map to 0, and there exists valid solution to (1), (2), and (3) for $area'$.*

3.3 DT-Triangular Diagrams

A third class of diagram restricts the regions representing the double and triple set intersections to being triangular (hence the name DT-triangular):

Definition 6. *A Venn-3 diagram is **DT-triangular** if AB, AC, BC and ABC are all triangles and A, B, and C are all polygons.*

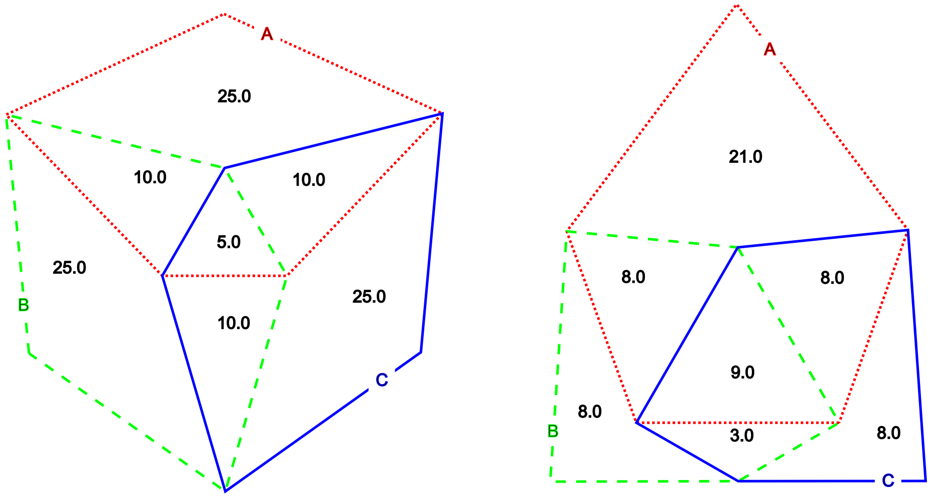

Fig. 7. DT-triangular diagrams.

The diagrams in Figure 7 are DT-triangular; the lefthand diagram has an area specification that cannot be drawn by a triangular diagram. Clearly, any triangular diagram is also DT-triangular and so, therefore, is any core-triangular diagram. We will use our drawing method for triangular diagrams to enable the construction of DT-triangular diagrams. The method relies on a numerical search (which we have implemented) to find a 'reduced' area specification (reducing the single set areas) that can be drawn by a triangular diagram, d. We then enlarge the single set regions, to give a diagram with the required area specification. It can be shown that every area specification can be represented by a DT-triangular diagram, but not necessarily in a convex manner:

Theorem 4. *Any area specification, area*: $R \rightarrow \mathbb{R}^+ - \{0\}$, *can be represented by a DT-triangular diagram.*

To justify this result, draw an equilateral triangle for the triple intersection and appropriate triangles for the double intersections. Complete the polygons to produce the correct areas for the single intersections.

3.4 CH-Triangular Diagrams

Our final diagram class is illustrated in Figure 8. Here, the convex hull (CH) of the core is a triangle and the triple intersection is also a triangle. CH-triangular diagrams allow some area specifications to be drawn using only convex curves that cannot be drawn in this manner by diagrams from any of the other classes.

Definition 7. *Let d be a Venn-3 diagram. Then d is **CH-triangular** if*

1. *the convex hull of the core is a triangle, T,*

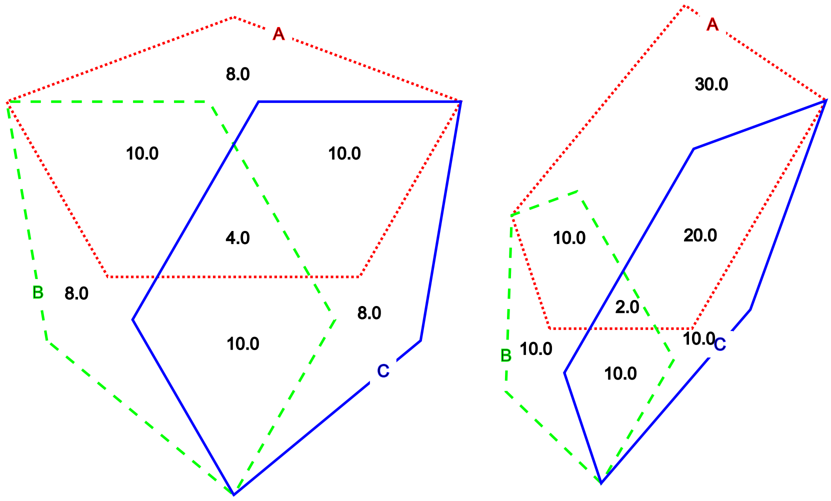

Fig. 8. CH-triangular diagrams.

2. *ABC is a triangle,*
3. *each of AB, AC, and BC is a polygon with at most five sides,*
4. *the convex hull T less the core consists of connected components that form triangles, of which there are at most three, each of which has two edges colinear with two edges of ABC, and*
5. *the remaining minimal regions, A, B, and C, are all polygons.*

In Figure 8, the area specification for the lefthand diagram cannot be represented by a convex DT-triangular diagram or, therefore, a triangular or a core-triangular diagram. We demonstrate a construction method for CH-triangular diagrams in the implementation section.

4 Implementation

In this section we discuss details of the implemented software, for a Java applet see *www.cs.kent.ac.uk/people/staff/pjr/ConvexVenn3/diagrams2010.html*. This software allows the user to enter an area specification and diagram type, and show the resultant diagram if a drawing is possible. The diagrams in all figures in this paper except 1, 3, 6 and 9 were drawn entirely with the software. In the case of DT-triangular diagrams, the software can produce diagrams with non-convex curves (when the double intersection areas are proportionally very small). As in the examples given previously, we use an equilateral triangle for the triple intersection of all diagram types because this tends to improve the usability of the final diagram, and also reduces the number of variables that need to be optimized during the search process described below.

In the case of core-triangular diagrams and triangular diagrams, the implementation uses the construction methods previously outlined. In the case of

DT-triangular and CH-triangular diagrams, we have no analytical approach for drawing an appropriate diagram, instead we use search mechanisms, outlined in the following two sections.

4.1 Constructing DT-Triangular Diagrams

In order to draw a convex, DT-triangular diagram, we reduce the areas of the single intersections until we obtain an area specification that is drawable as a triangular diagram. That is, we seek A', B' and C' where $A' \leq A$, $B' \leq B$, and $C' \leq C$ where the new area specification has a valid solution, as described in Section 3; note that A', B' and C' could be negative. Moreover, we seek a solution where the discriminant for the quadratic arising from equations (1), (2) and (3) is zero. We note that, when the discriminant is zero, the to-be-drawn diagram will be more symmetric, as the outer points of AB, AC and BC will be closer to the centre of inner equilateral triangle. If this solution is valid, we can then proceed to draw a triangular diagram for the reduced specification. Once we have this drawing, we can enlarge the single intersection minimal regions, until they have the areas as required in the original area specification. If no valid solution can be found, it is possible that the original area specification can be drawn as a convex CH-triangular diagram, for which we discuss our drawing method in the next subsection. However, a non-valid solution can still yield a DT-triangular diagram, but it will not be convex.

To illustrate the drawing method for DT-triangular diagrams, we start with the area specification for the lefthand diagram in Figure 7. Here, the single set areas are all 25. Reducing them to 8.8 yields an area specification that is representable by a triangular diagram; see Figure 5. We can then enlarge the single intersections, pulling the polygons outwards, to produce the shown DT-triangular diagram (lefthand side of Figure 7). Drawing the righthand diagram of Figure 7 requires both B' and C' to be negative, however this is not a barrier to the drawing of the diagram, and the same process can be applied.

4.2 Constructing CH-Triangular Diagrams

If an area specification cannot be represented by a DT-triangular, convex diagram then it might be representable by a CH-triangular, convex diagram, as illustrated previously. In fact, we have not found any area specifications that can be represented by a DT-triangular, convex diagram that cannot also be represented by a CH-triangular, convex diagram. A further advantage of using CH-triangular diagrams is that they can have fewer points where two curves intersect and that is also a vertex of a polygon, which can increase the usability of the diagram. The process for drawing a CH-Triangular diagram is as follows:

1. The method first finds a drawing (convex or non-convex) for the specification with the DT-Triangular method, as given in the previous subsection. The convex hull of the core of this diagram forms a starting point for the search that finds the core of the CH-triangular diagram.

2. The DT-Triangular diagram is converted to a CH-Triangular diagram by extending the vertices of the curves beyond the middle triangle as shown in Figure 9: the solid lines show the DT-Triangular core, the dotted lines show how the curve line segments change when the diagram is converted to a CH-Triangular core.
3. An attempt is made to produce the correct double intersection areas using a search mechanism. The search ensures that if the diagram can be drawn using only convex curves with the given single intersection areas, then it will be. Each point on the convex hull of the core is tested for a number of possible moves close to the current location, to see whether there is a location that improves a heuristic. The heuristic is based on minimizing the variance of difference of the current double intersection areas against required areas. However, a move is only made as long as it results in a convex diagram. The process is repeated until the heuristic gives a zero result or no more improvement can be made. If the search finishes and the heuristic is zero, the diagram can be drawn in a convex manner. If it is not zero, then an additional search is made, this time relaxing the convex requirement until the heuristic becomes zero. In this case the diagram will be non-convex.
4. The outer vertices of the single intersections are placed so that the corresponding areas are enlarged until they match the area specification. This results in diagrams of the form shown in Figure 8. If the diagram can only be drawn with a CH-Triangular diagram type in a non-convex manner (in the case the extra search is required in the previous step), then a non-convex drawing is be generated by indenting the one set areas into the cut outs.

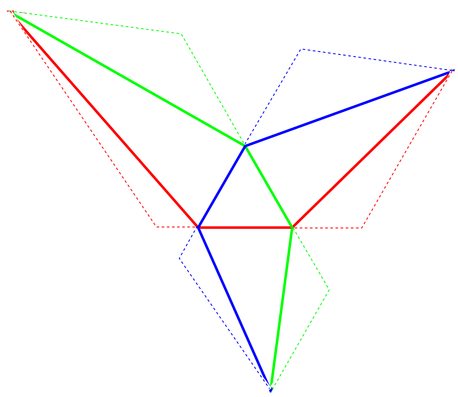

Fig. 9. Converting between DT and CH diagram types.

4.3 Layout Improvements

One significant usability issue is that, in these diagram types, the vertices of the curves often coincide with intersection points. This can make following the correct curve through the intersection difficult, particularly when colour cannot be used to distinguish the curves. As a result, we demonstrate a layout improvement mechanism that could be applied to all diagram types, but is only implemented for CH-triangular diagrams. The layout improvement method moves the vertices away from the intersection points by elongating the relevant curve segment out from the centre of the diagram. The vertices of the affected polygons are then moved to compensate for this change. An example for the CH-triangular diagram shown in Figure 10, where the lefthand diagram is modified to give the (improved) righthand diagram. Here, only the outer intersection points are affected, as the diagram type naturally separates the ABC triangle intersection points from the curve vertices.

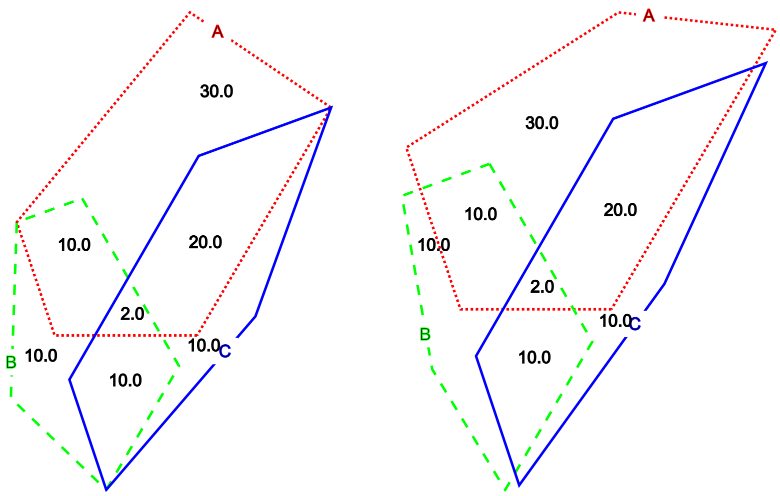

Fig. 10. Improving the layout of a CH-Triangular diagram.

In these cases, for a given polygon, the line segments that border the two set areas point 'outwards' when the two set areas are of sufficient size, allowing the one set area enclosed by the curve to be as large as required. However, as the two set areas bordered by the curve reduce in size, they become closer to the three set border to the point where they become parallel. If the two set areas reduce further in size, these line segments start to point 'inwards'. At this stage the one set area is restricted in maximum size, if the diagram is still to remain convex. However, a more sophisticated implementation would avoid this by a more exact measurement of the elongation.

5 Conclusion

We have provided several classes of Venn-3 diagram that have allowed us to identify some area specifications as drawable with a convex diagram. Given an area specification, we have provided construction methods that draw a diagram representing it. In order to enhance the practical applicability of our results, we have provided a software implementation that draws an appropriate diagram, given an area specification.

Future work will involve a further analysis of which area specifications can be represented by a convex diagram. Ultimately, we would like to know exactly which area specifications can be represented in this way and how to construct such drawings of them. By considering convex polygons, we restricted the kinds of diagrams that could be drawn. The general case, where arbitrary convex curves can be used, is likely to be extremely challenging. In addition, natural extensions of the research are to consider Euler-3 diagrams, where not all of the minimal regions need to be present, and to examine diagrams with more than three curves.

Acknowledgements We thank Shena Flower for helpful discussions on aspects of this paper. This research forms part of the Visualization with Euler Diagrams project, funded by EPSRC grants EP/E011160/1 and EP/E010393/1. The work has also been supported by EPSRC grants EP/H012311/1 and EP/H048480/1.

References

1. Google Charts API. http://code.google.com/apis/chart/, accessed August 2009.
2. S. Chow. *Generating and Drawing Area-Proportional Euler and Venn Diagrams*. PhD thesis, University of Victoria, 2007. hdl.handle.net/1828/128.
3. S. Chow and P. Rodgers. Constructing area-proportional Venn and Euler diagrams with three circles. In *Proceedings of Euler Diagrams 2005*, 2005.
4. S. Chow and F. Ruskey. Drawing area-proportional Venn and Euler diagrams. In *Proceedings of Graph Drawing 2003, Perugia, Italy*, volume 2912 of *LNCS*, pages 466–477. Springer-Verlag, 2003.
5. S. Chow and F. Ruskey. Towards a general solution to drawing area-proportional Euler diagrams. In *Euler Diagrams 2004*, volume 134 of *ENTCS*, pages 3–18. ENTCS, 2005.
6. M. Farfel et al. An overview of 9/11 experiences and respiratory and mental health conditions among world trade center health registry enrollees. *Journal of Urban Health*, 85(6):880–909, 2008.
7. E. Ip. Visualizing multiple regression. *Journal of Statistics Education*, 9(1), 2001.
8. H. Kestler, A. Muller, J. Kraus, M. Buchholz, T. Gress, H. Liu, D. Kane, B. Zeeberg, and J. Weinstein. Vennmaster: Area-proportional Euler diagrams for functional GO analysis of microarrays. *BMC Bioinformatics*, 9(67), 2008.
9. P. Rodgers, J. Flower, G. Stapleton, and J. Howse. Some results for drawing area proportional Venn3 with convex curves. In *Information Visualization*, pages 667–672. IEEE, 2009.
10. J. Soriano, K. Davis B. Coleman, G. Visick, D. Mannino, and N. Pride. The proportional Venn diagram of obstructive lung disease. *Chest*, 124:474–481, 2003.