Kent Academic Repository

The Impact of Topological and Graphical Choices on the Perception of Euler Diagrams

Andrew Blake, Gem Stapleton[a], Peter Rodgers[b], John Howse[a]

[a]University of Brighton, England
{a.l.blake,g.e.stapleton,john.howse}@brighton.ac.uk
[b]University of Kent, England
p.j.rodgers@kent.ac.uk

Abstract

This paper establishes the impact of topological and graphical properties on the comprehension of Euler diagrams. To-date, various studies have examined the impact of individual properties of Euler diagrams, such as curve shape and orientation. This has allowed us to establish guides for using these properties such as 'draw Euler diagrams with circles' and 'draw Euler diagrams without regard to orientation'. However, until the work described here, questions still remain, for example 'do these guides, when combined, make a significant difference to real-world Euler diagrams?', and if so, 'should they be used by those visualizing set data with Euler diagrams?' To answer these questions an empirical study was conducted to compare Euler diagrams that have been drawn by others for their real-world data, against versions that adhere to all of the guides in combination. The study establishes that both the accuracy and the speed with which information is derived from Euler diagrams is significantly improved when Euler diagrams adhere, where possible, to all the guides. The improvement is considerable when using the guided diagrams, with on average, the error rate being more than halved from 21.4% to 10.3%, and a 9 second improvement in the average time taken, from 34.2 seconds to 24.9 seconds. As Euler diagrams are regularly used to visualize information in a multitude of areas, ranging from crime control to social network analysis, our results indicate that applying the guides to these diagrams will improve the ability of users to accurately and quickly extract information.

Keywords: Euler diagram, visual communication, empirical study
PACS: 01.20.+x

1. Introduction

There has been a rapid rise in the volume of data where the data items lie in overlapping sets. Reflecting the importance of understanding such data there have been a large number of techniques proposed for visualizing it such as [1, 10, 16, 25, 28, 29, 32, 33, 39]. Most of these techniques represent the sets using Euler diagrams.

Euler diagrams are often regarded a natural [30] and intuitive [27] way to depict sets. They represent sets using graphical elements called closed curves. The interior of each curve represents items that are in the set [30, 26]. Figure 1 shows a simple classification. The curve labelled 'Mammals' intersects with the curve labelled 'Aquatic' meaning that there are some mammals that are aquatic. 'Cats' is contained within 'Mammals' meaning that all cats are mammals. 'Cats' is disjoint from 'Aquatic' so there are no cats that are aquatic.
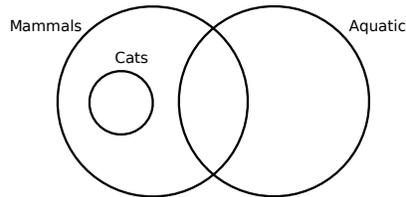

Figure 1: Simple Euler diagram.

We can observe Euler diagrams deployed in many situations. In criminal investigations, sets represent organizations to which people (the data items) belong or locations they frequent [12]. Similar complex data occur in biological settings where data items are genes, whilst sets represent shared features of the genes [16]. We also find Euler diagrams being used in both art and architecture [2], education [14], computer file organisation [11] and classification systems [35]. Given such abundant use, there is a strong motivation to draw Euler diagrams that aid users' comprehension of the information being visualized.

Here we present a new study which focuses on all guides that are known to aid the comprehension of Euler diagrams. Section 2 presents a list of these guides and in doing so identifies the research from where they have been derived. To establish if these guides, in combination, aid users' comprehension

we compare real-world Euler diagrams with guided diagrams, explained in detail in section 3. The study considered 12 real-world diagrams identified, in part, from the application domains mentioned earlier. We establish that Euler diagrams adhering to all of the guides significantly outperform real-world diagrams both in terms of accuracy and speed. We conclude that drawing Euler diagrams adhering to all of the guides best facilitates performance. All of the diagrams used in our study, and the data collected, are available from `http://www.cem.brighton.ac.uk/staff/alb14/experimental_resources/combined_guides/combined_guides.html`.

2. Related Work

At present, there exists an array of guidance that has been established, either theoretically or empirically, to aid users' comprehension of Euler diagrams. In this section, we define this guidance in terms of fundamental choices that must be made when laying out Euler diagrams. In doing so we identify a series of guides 1 to 10. For each guide we explain how the efficacy of an Euler diagram is improved with reference to the guide's theoretical or empirical underpinning.

2.1. Euler Diagram Layout Choices

There are numerous choices that need to be made when drawing Euler diagrams for a given data set. We categorize these choices into three types: descriptional (the abstract syntax level), topological and graphical (both at the concrete syntax level). The first choice that must be made is descriptional, which determines the *zones* that must be present in the Euler diagram [32]. The zones are the smallest regions in Euler diagrams which represent overlaps between the sets; Figure 1 has four zones inside its three curves. A descriptional choice determines which zones are present, in addition to those that represent non-empty overlaps between sets. An example is given in Figure 2, where the left hand diagram contains no superfluous zones whereas the right hand diagram contains additional (shaded) zones.

This choice is embodied in Gurr's theory of well-matchedness [13]:

> "The transitive, irreflexive and asymmetric relation of set inclusion is expressed via the similarly transitive, irreflexive and asymmetric visual of proper spatial inclusion in the plane."
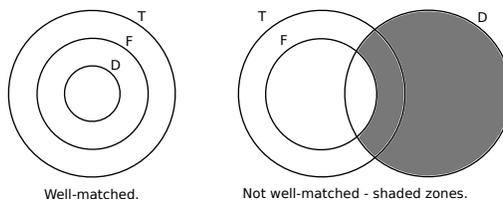
3

Figure 2: Choices of abstract syntax.

Euler diagrams are well-matched precisely when they do not include any 'extra' zones. To summarise, we define guide 1 for Euler diagram drawing:

Guide 1 (Well-matched). *Draw well-matched Euler diagrams (i.e. no extra zones).*

Being able to draw a well-matched diagram only solves part of the problem producing an effective Layout. At the concrete syntax (i.e. drawn diagram) level, a variety of topological properties are known to impact comprehension [26] and are illustrated in figures 3 (b) to (g). Diagrams that possess such properties are said to be not well-formed. Figure 3 (a) illustrates both a well-matched and well-formed diagram while figures 3 (b) to (g) each illustrate a similar well-matched diagram but are not well-formed due to the following properties. Figure 3(b) illustrates brushing points, where two or more curves meet but do not cross at a point i.e. curve 'a' is brushing with curve 'b'. Figure 3(c) illustrates concurrency, where segments of two or more curves are concurrent i.e. the top segment of the curve labelled 'b' being concurrent with the top segment of the curve labelled 'c'. Figure 3(d) illustrates duplicate curve labels, where two or more curves represent the same set of data i.e. curve labels 'a' are duplicated. Figure 3(e) illustrates a disconnected zone, where a zone consists of one or more minimal regions i.e. the zone labelled 'c' consisting of two minimal regions being 'divided' by the curve labelled 'b'. Figure 3(f) illustrates a triple point, where three curves meet at a point i.e. curve 'a' is brushing with curve 'b' and, in turn, curves 'a' and 'b' are brushing with curve 'c'. Finally, figure 3(g) illustrates a non-simple curve, where a curve self-intersects i.e. the curve labelled 'b' is a non-simple, self-intersecting, curve. The results of [26] are summarised here:

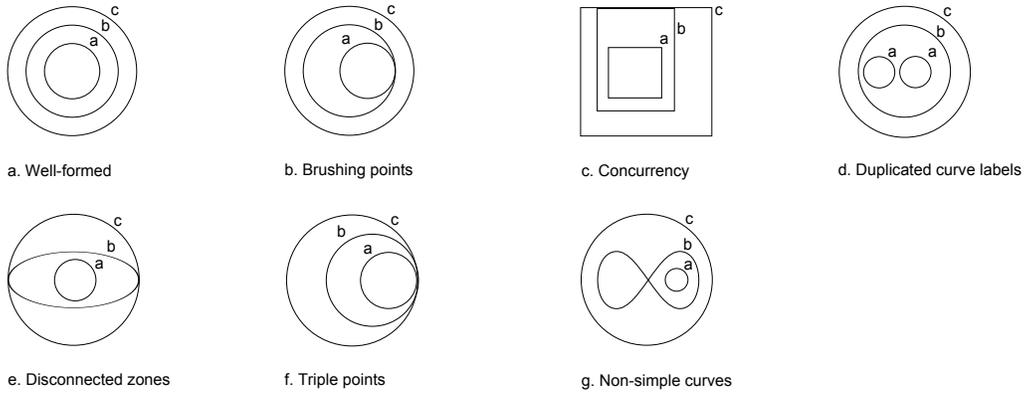**Guide 2 (Well-formed).** *Draw well-formed Euler diagrams.*

Figure 3: Non well-formed Euler diagrams.

Irrespective of laying out well-matched and well-formed Euler diagrams there still exist numerous graphical choices to be made when ascertaining an effective layout of an Euler diagram. Benoy and Rodgers [3], with their work on aesthetics, also acknowledge the importance of making correct graphical choices when laying out Euler diagrams. They conducted a study that focused on the jaggedness of curves, zone area equality and the closeness of one closed curve to another. To summarise their results, we define three further guides:

Guide 3 (Smooth curves). *Draw Euler diagrams with smooth curves.*

Guide 4 (Zone area equality). *Draw Euler diagrams with zone area equality.*

Guide 5 (Diverging lines). *Draw Euler diagrams with diverging lines.*

There are many other graphical choices that might be considered. Bertin [4] identifies both planar and retinal variables, which constitute a variety of graphical choices, to which we are known to be perceptually sensitive. With respect to planar variables, Blake et al. established that the effect of an Euler diagram's orientation does not impact on users' comprehension [5], leading to:

Guide 6 (Orientation). *Draw Euler diagrams without regard to orientation.*

5

Retinal variables include shape and colour, both of which are fundamental choices that must be made when drawing Euler diagrams. Recently Blake et al. established that shape significantly impacts users' comprehension [6], leading to three further guides:

Guide 7 (Shape). *Draw Euler diagrams with circles.*

Guide 8 (Symmetry). *Draw Euler diagrams with highly symmetrical curves.*

Guide 9 (Shape Discrimination). *Draw Euler diagrams so that the zones are discernable from the curves via their shape, but not at the expense of symmetry.*

With respect to shape, guide 7 encapsulates the significant impact a retinal variable can impose on users' comprehension of Euler diagrams. Consequently, this observation afforded the motivation to address the question "how does the use of colour in Euler diagrams affect users' comprehension?". To this end, Blake et al. established that how we use colour does indeed significantly impact users' comprehension [7], leading to a further guide:

Guide 10 (Colour). *Draw Euler diagrams with curves that have no fill and different colours for each set represented.*

3. Evaluating Guidance on How to Draw Euler Diagrams

We now have a comprehensive set of guides for how best to draw Euler diagrams. These guides were independently produced, but nobody has yet established whether their combined effect is positive. We hypothesise that guides 1 to 10, when combined, will improve the comprehension of real-world Euler diagrams. To test this hypothesis we executed an empirical study and, congruent with previous studies [1, 25, 5, 18, 15, 23, 24], we view comprehension in terms of task performance: one diagram is more comprehensible than another diagram if users can interpret it, on average, more accurately or more quickly.

3.1. Choosing Diagrams

Twelve diagrams were chosen to form the basis of the study. Four scaled examples of these diagrams are illustrated in figures 4 to 7. These diagrams manifest from a variety of application domains. Together, these diagrams strive to visualize information for academic, arts, business and social communities alike. Most importantly, these diagrams break at least one of our guides, for example by employing colour fill. The matrix in table 1 shows which diagrams adhere to which guides. If a diagram adheres to a guide a ✓ is present in the corresponding cell. If a diagram *does not* adhere to a guide an ✗ is present in a corresponding cell. The end column provides a count of the number of times a diagram does not adhere to a guide. The bottom row provides a count of how many times a guide is not adhered to. Scaled versions of all 12 diagrams chosen for the study are presented in appendix 1 along with references to the corresponding websites from which they were sourced.

Diagram \ Guides	1	2	3	4	5	6	7	8	9	10	Guides broken
1	✓	✓	✗	✓	✓	✓	✗	✓	✓	✗	*3*
2	✓	✗	✓	✓	✓	✓	✗	✓	✓	✗	*3*
3	✗	✓	✓	✗	✓	✓	✓	✓	✓	✗	*3*
4	✗	✗	✓	✗	✓	✓	✗	✓	✗	✗	*6*
5	✗	✓	✗	✗	✗	✓	✗	✗	✗	✗	*8*
6	✗	✗	✓	✗	✓	✓	✗	✓	✓	✗	*5*
7	✓	✗	✓	✓	✓	✓	✗	✓	✗	✗	*4*
8	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	*1*
9	✗	✓	✓	✗	✓	✓	✗	✗	✗	✓	*5*
10	✗	✗	✓	✗	✓	✓	✗	✓	✓	✗	*5*
11	✗	✗	✓	✗	✗	✓	✗	✗	✗	✗	*8*
12	✗	✗	✓	✗	✓	✓	✓	✓	✓	✗	*4*
Guides broken	8	7	*2*	8	*2*	*0*	9	*3*	5	11	

Table 1: Real-world diagrams and the guides they break.

To summarise the information in table 1, we consider how frequently a guide was not adhered to. Guide 10, which says we should use of curves that have no fill and different colours for each set represented, was most frequently not adhered to, with a count of *11*. Guide 7 , which tells us to use of circles, was the second most frequent guide not to be adhered to, with a count of *9*.
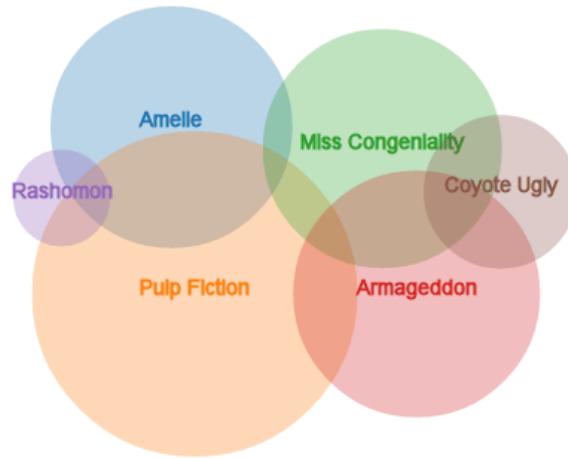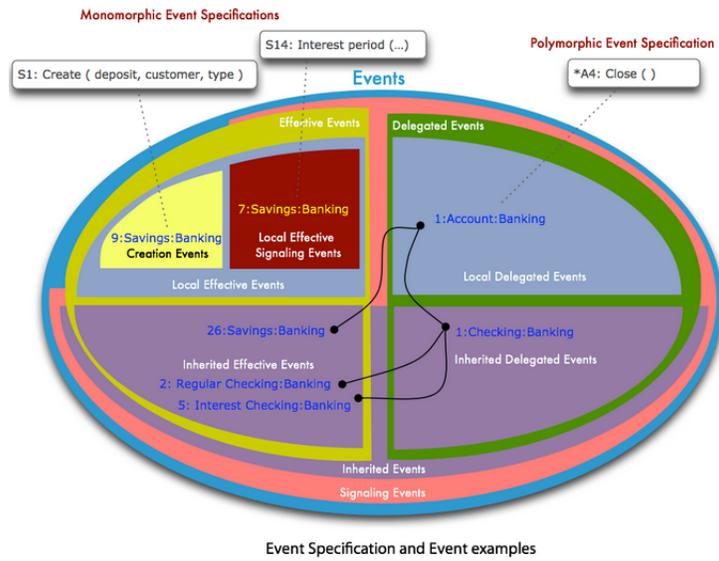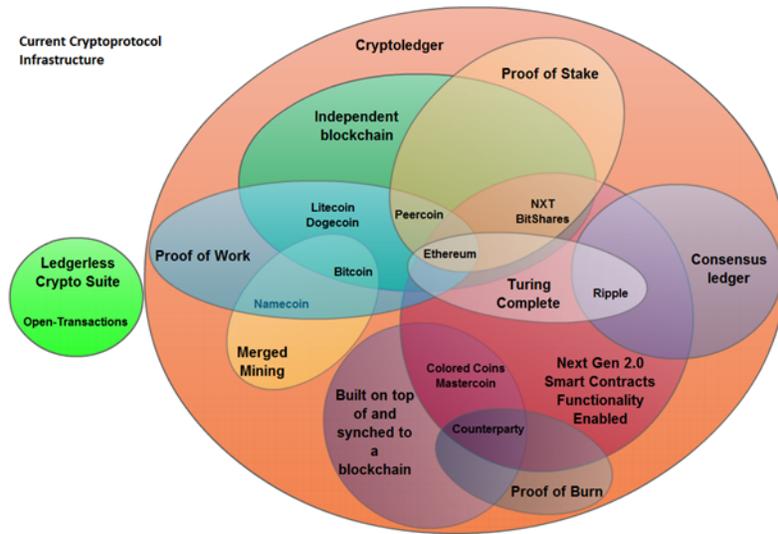
Figure 4: Diagram 3.

Figure 5: Diagram 5.
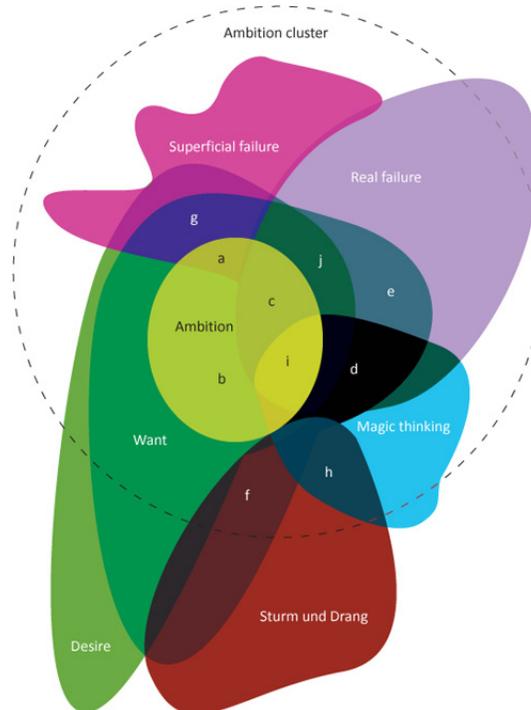
8

Figure 6: Diagram 10.

Figure 7: Diagram 11.

9

Guides 1 and 4, which tell us to avoid extra zones i.e. well-matched diagrams and poor zone area equality respectively, were the third most frequently guide not to be adhered to, with a count of *8*. Guide 6, which tells us we can draw Euler diagrams without regard to their orientation, was by definition applied.

3.2. Creating Diagrams for the Study

Two sets of diagrams were drawn for the study and were derived from the 12 diagrams identified in section 3.1. The first set of diagrams are called real-world diagrams. These diagrams were drawn maintaining the same topological and graphical properties as those in section 3.1. However, they were also drawn omitting information deemed to be too technical or ambiguous for the participants during the study. Instead, information was replaced with that specified in section 3.3.1. When replacing the information it was important that each real-world diagram exhibited consisted layout features, for example the size and style of font used for each curve label. This was an essential design consideration in order to help minimise unwanted variations between each of the real-world diagrams. Figures 8 to 11 illustrate scaled versions of four real-world diagrams derived from those diagrams presented in figures 4 to 7. To further minimise unwanted variations between diagrams, it was necessary that any unfamiliar or extraneous layout features were removed. An example of an extraneous feature is the supplementary syntax in the form of dots and lines connecting data items present in the diagram illustrated in figure 5. The diagram illustrated in figure 8 does not exhibit the aforementioned extraneous feature. Scaled versions of all 12 real-world diagrams are presented in appendix 2. Each of the real-world diagrams adhere to the following layout features:

1. the curve labels were written using upper case letters in Times New Roman, 14 point size, font in bold,
2. each curve label was positioned on the outside, and closest to, its corresponding curve,
3. each curve label was oriented horizontal in the plane,
4. fonts inherited the colour treatment from the original diagrams,
5. data items were written using lower-case letters, except that the first letter was capitalised, and with Arial 12 point size font,
6. data items were evenly distributed, where possible, within each zone, and
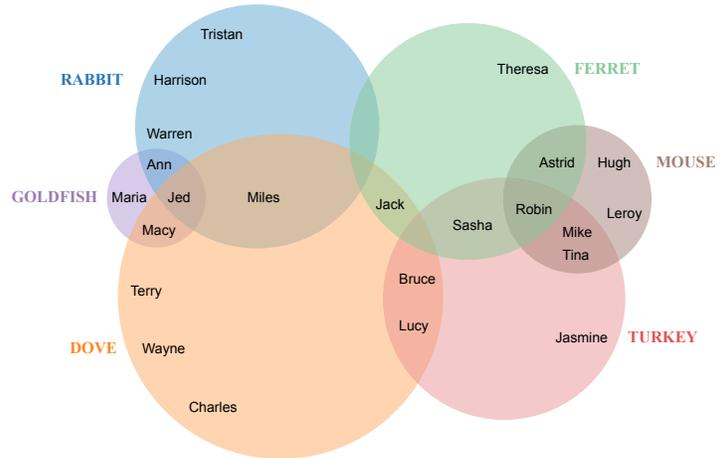7. all diagrams were drawn to fit in an area of 810×765 pixels.
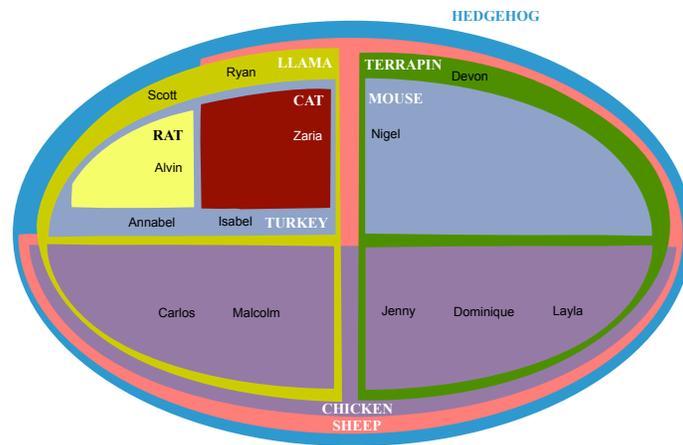
Figure 8: Real-world diagram 3.
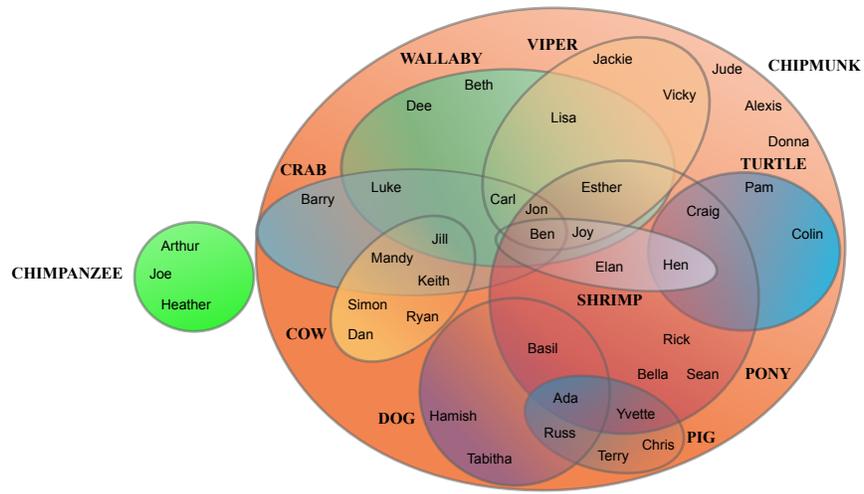

Figure 9: Real-world diagram 5.

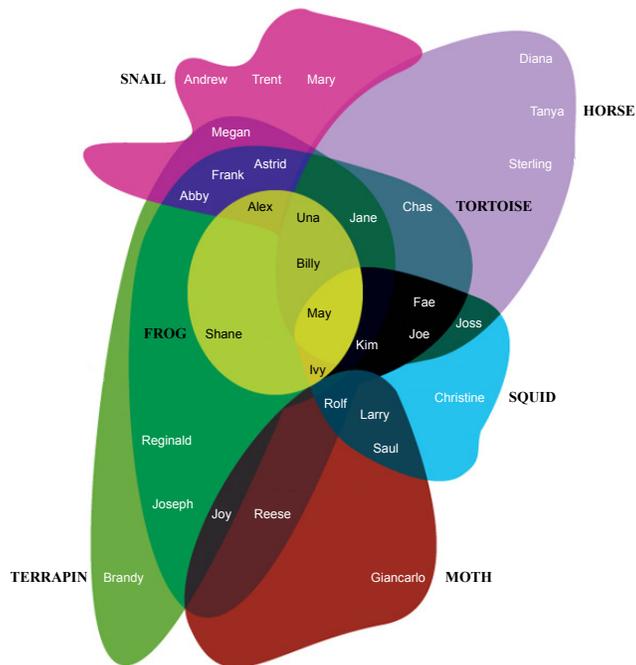Figure 10: Real-world diagram 10.

Figure 11: Real-world diagram 11.

The second set of diagrams that were drawn for the study were referred to as guided diagrams. The guided diagrams were derived from the real-world diagrams. As such, they inherited the information being conveyed by the real-world diagrams. However, the guided diagrams were further constrained by a layout feature specified in previous studies [5, 6, 7], which was that all closed curves had a 3 pixel stroke width.

Most importantly, the guided diagrams were drawn to adhere to guides 1 to 10 as illustrated by the scaled diagrams in figures 12 to 15. However, diagrams 6, 10 and 11 could not be drawn to meet all of the guides due to their underlying abstract syntax. In each case they were drawn not well-formed, breaking guide 2, in the same way as their real-world counterpart. Diagram 10 also broke guide 7 as one set, labelled SHRIMP, had to be drawn using an ellipse as illustrated in figure 14. Scaled versions of all 12 guided diagrams are presented in appendix 3.

3.3. Experimental Design

The study uses a between group design consisting of two participant groups, A and B. Table 2 illustrates how the real-world and guided diagrams were mixed between participants. Mixing diagrams between groups A and B in this way helped mitigate possible carry-over effects, explained in section 5. While participants attempted to elicit information from the Euler diagrams two dependent variables were recorded: error and time. An error was recorded when a question had been answered incorrectly and time was measured on how long it took a participant to answer a question. The independent variable was diagram. It was assumed that there would be no statistical significance between the real-world and the guided diagrams with respect to error counts and mean times. We will now present examples of the information that was added to the diagrams and go onto describe the experimental execution.

Group	Real-world diagrams	Guided diagrams
A	1 to 6	7 to 12
B	7 to 12	1 to 6

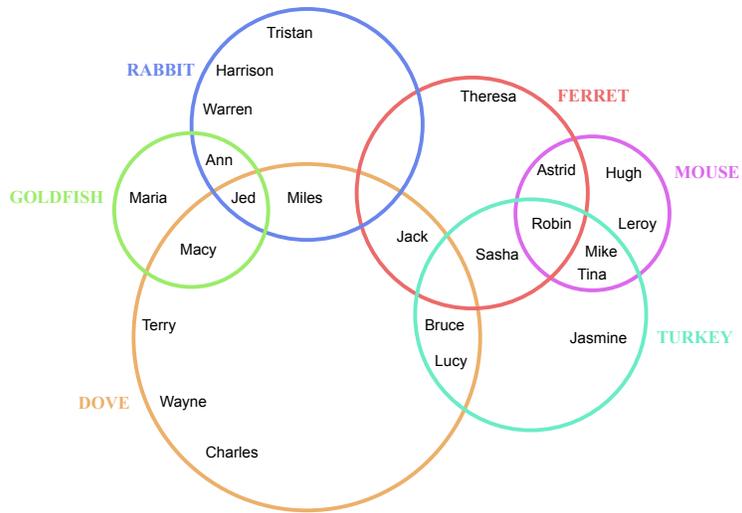Table 2: Diagrams mixed within and between participant groups.
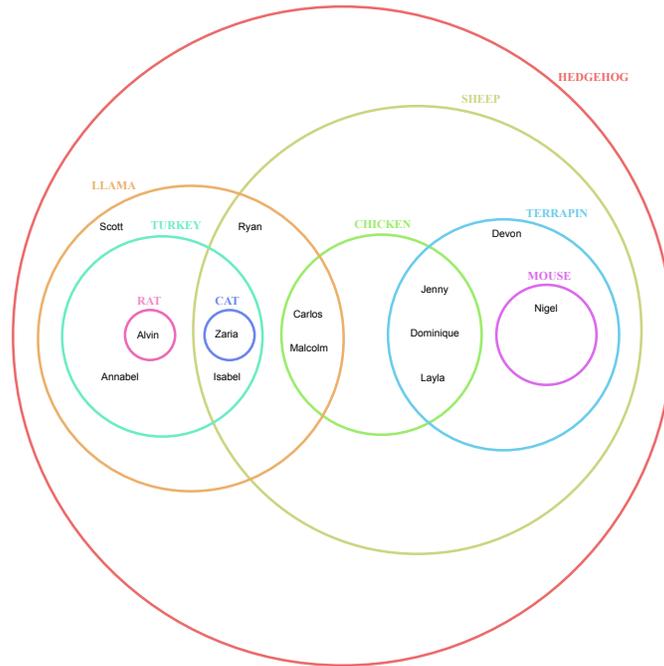
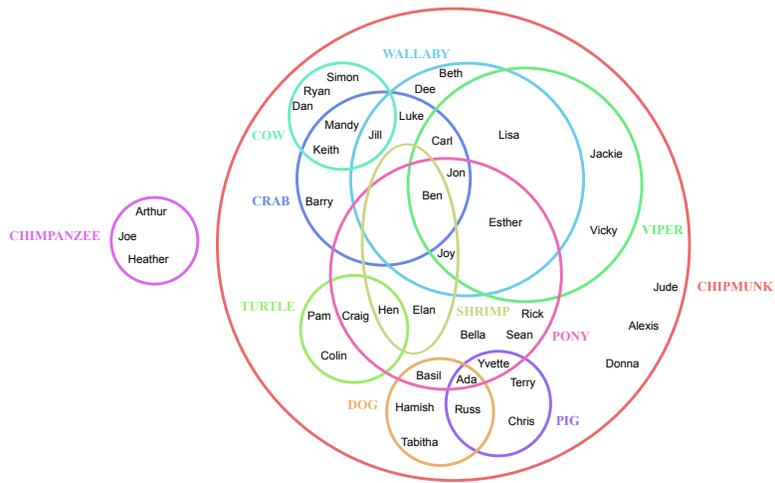Figure 12: Guided diagram 3.

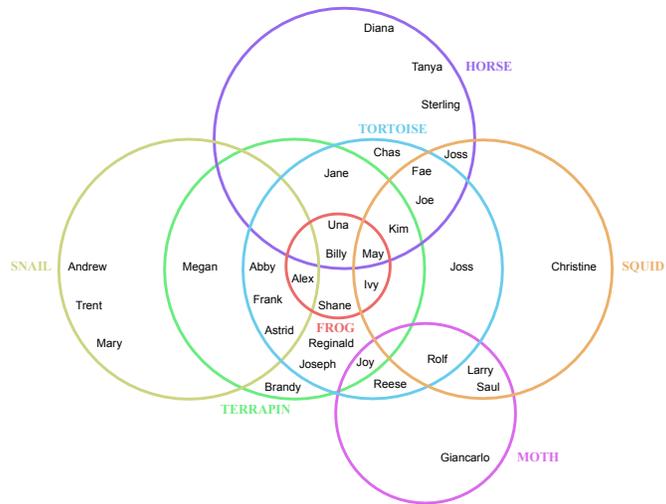Figure 13: Guided diagram 5.

Figure 14: Guided diagram 10.

Figure 15: Guided diagram 11.

3.3.1. Information and Question Styles

All the diagrams were drawn to convey information that was regarded accessible to the reader. That is to say, the terminology was neither technical, specialist or culturally dependant. To this end, we chose to visualize information about people and the pets that they own. This information was regarded as generic and it was anticipated that participants would all be familiar with this context. Pet names reflected a variety of common, and not so common, domesticated animals. People names were first names only and included a mixture of both male and female names, reflecting a variety of ethnicities.

Three styles of question were specified, 'Who', 'Which' and 'How many'. In order to answer these questions participants had to consider a variety of set theoretic relationships, namely intersection, subset and disjointness. These are the same set theoretic considerations required to elicit information from the original diagrams upon which the study is based. Examples of the three question styles are listed below:

1. How many people own a GOLDFISH and a DOVE but not a RABBIT?
2. Which pet is owned by 10 people?
3. Who owns a CHIPMUNK, VIPER and a WALLABY but neither a PONY nor a CRAB?

The above questions were asked of the real-world diagrams illustrated in figures 8 to 10 and the guided diagrams illustrated in figures 12 to 14 respectively. The answer to the above questions is 1 (Macy), SHEEP and Lisa. During the study, questions had a choice of five possible answers. When answering 'Who' style questions participants were required to identify an individual's name. When answering 'Which' style questions participants were required to identify a pet name. When answering 'How many' style questions participants were required to count the number of individuals in a region which was always a value between 0 and 4. If the number of individuals counted was greater than 0 then each individual name was also captured. This was an attempt to avoid false positives. In each instance, only one answer out of the five choices presented was correct.

3.4. Experiment Execution

Initially, a pilot study was conducted consisting of six participants, three per group. It was established that a number of minor modifications were

required before the main study could begin. These modifications involved adjusting the phrasing of two questions and correcting the spelling of several curve labels. Consequently, the pilot data was discarded. Having made the modifications, 60 participants were identified for the study. Participants filled out an electronic form capturing, among other attributes, their gender and age (46 M – 1 of whom was colour blind, 14 F, ages 18 to 36) . All participants were randomly allocated to equal sized groups. They were all students from the University of Brighton's School of Computing, Engineering and Mathematics and they spanned both undergraduate and postgraduate levels.

For the collection of performance data, we used a software tool (called the research vehicle) to display the diagrams and questions, to gather answers and the time taken. The diagrams were presented in a random order. The time taken to answer a question was determined from the instant a question was presented until the instant a participant had selected an answer to the question. Each time the participant answered a question, the research vehicle would ask them to indicate when they were ready to proceed to the next question, thus allowing a pause between questions. There was a maximum time limit of two minutes for each question to ensure that the experiment did not continue indefinitely. If a participant took longer than two minutes then no answer was captured and the time logged as a 'time out'. In this event the software tool moved on to the next question.

The experiment was performed within a usability laboratory which affords a quiet environment free from noise and interruption. The same computer and monitor was used by each participant. To ensure colours were optimally displayed the monitor provided a 1920 x 1200 resolution, 0.270 mm pixel pitch, 300 cd/m^2 brightness and a contrast ratio of 1000:1 (static). All participants were alone during the experiment, in order to avoid distractions, with the exception of an experimental facilitator who was present throughout. All participants that took part in the study successfully completed the experiment. The experiment took approximately 40 minutes per participant and they were given a canteen voucher worth £6 for their contribution to the research.

The first phase of the experiment was initial training. All participants were asked whether they were familiar with Euler diagrams. While no participant acknowledged familiarity with Euler diagrams several stated they had seen or heard of Venn diagrams. Consequently, all participants were treated as having no previous experience of Euler diagrams and were given the same

training. Training began by introducing participants to the notion of Euler diagrams and the styles of questions to be asked. This was achieved using hard copy printouts of the diagrams, one for each style of question. Participants were given a few minutes to study the diagrams and questions, after which the experimental facilitator explained how to answer the questions.

The second phase of the experiment provided participants with further training on the notion of Euler diagrams as well as how to use the research vehicle. Participants were presented with six questions, one at a time. If a question was answered incorrectly the facilitator went through the question with the participant. The third phase of the experiment is where we collected performance (error and time) data.

4. Statistical Analysis

The 60 participants in the main study each answered 12 questions giving a total of 720 observations. We consider error data to be a more important indicator of performance in our evaluation.

4.0.1. Error Analysis

Of the 720 observations there were a total of 114 errors giving an error rate of $\frac{114}{720} = 15.8\%$. The real-world diagrams yielded 77 errors giving an error rate of $\frac{77}{360} = 21.4\%$ while the guided diagrams yielded 37 errors giving an error rate of $\frac{37}{360} = 10.3\%$. The comparative error counts for diagrams 1 to 12 are illustrated in figure 16. To establish if there was an effect, we performed a χ^2 goodness-of-fit test. If $p \leq 0.05$ then we would reject the statistical assumption stated in section 3.3. There was a significant difference between the real-world and guided diagrams, with $p < 0.001$. Hence, we conclude that the guided diagrams significantly increase a users' accuracy when interpreting Euler diagrams.

4.1. Time Analysis

Of the 720 observations the mean time was 25.55 seconds (sd:22.08). From these data, the mean time for the real-world diagrams was 34.18 seconds (sd:24.52) while the mean time for the guided diagrams was 24.91 seconds (sd:18.23). The comparative mean times for diagrams 1 to 12 are illustrated in figure 17. Here we observe that all guided diagrams, with the exception of diagram 10, have smaller mean times than the real-world diagrams.
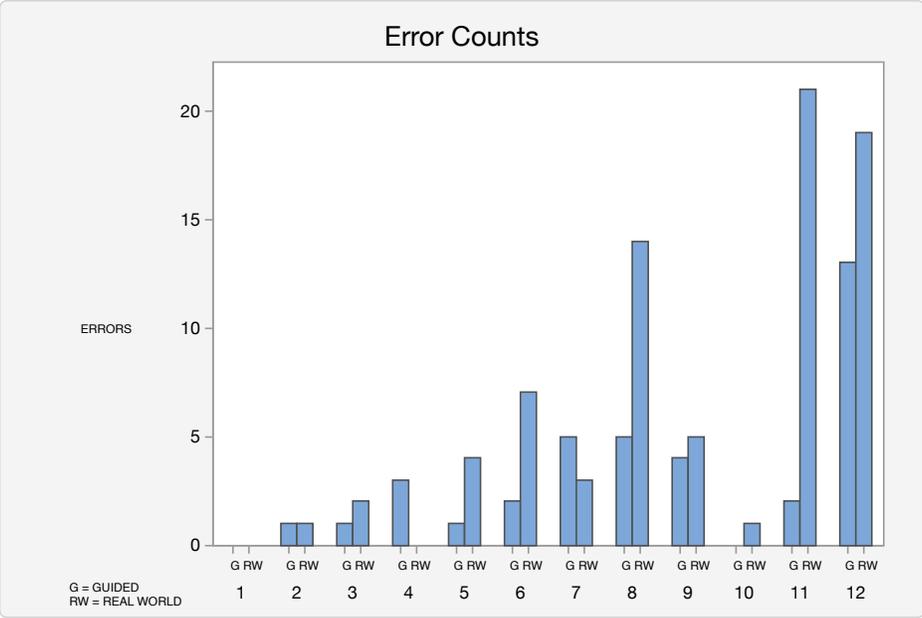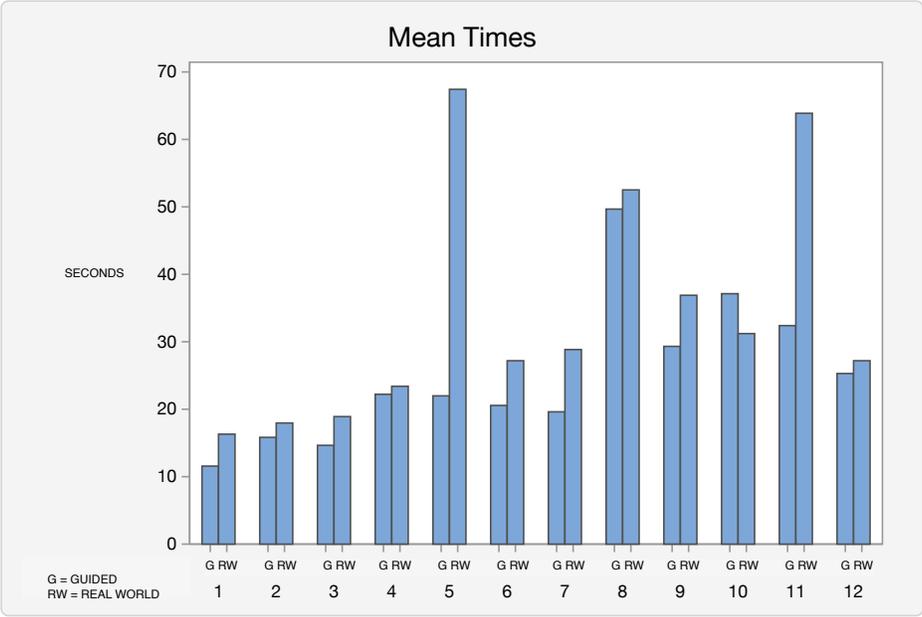
Figure 16: Comparitive error counts.

Figure 17: Comparitive mean times.

19

To determine if there existed significant variation between these means we performed a Repeated Measures Analysis of Variance (RM-ANOVA). Participants were regarded to be a random factor in the analytical model. This is a conservative approach to the analysis as the RM-ANOVA did not estimate the effect of each of the participants in the sample, but instead estimated the variability attributable to the participants. The RM-ANOVA was performed on logged data; whilst normality is not achieved, the skewness is 0.44 meaning our analysis is robust. If $p \leq 0.05$ then we would reject the statistical assumption stated in section 3.3. There was a significant difference between the mean times for real-world and guided diagrams ($F(1, 11) = 7.04$, $p =< 0.001$). This difference persisted when both the error data and the 'time out' data was removed from the analysis. There were 7 time out measurements and how they manifest is detailed in section 3.4. The effect size is 66% and tells us approximately $\frac{2}{3}$ of participants were, on average, faster interpreting guided diagrams than the real-world diagrams. Hence, we conclude that guided diagrams significantly reduce the time taken to answer questions.

4.2. Summary of Performance Data

To summarise the analysis of the performance data, we have presented evidence that suggests guided diagrams significantly improve upon users' comprehension of Euler diagrams. Specifically, guided diagrams significantly reduce the number of errors accrued when answering questions while simultaneously reducing the time taken to elicit an answer. It is important to remind the reader that as with all empirical studies, including those cited in section 2, the results are valid within the constraints of the study.

4.3. Discussion of Results

With respect to errors, guided diagrams performed significantly better. Indeed, eight of the 12 guided diagrams accrued fewer errors. In the most extreme case, real-world diagram 11, which broke eight out of ten guides, accrued 21 errors while its guided counterpart accrued two. If we consider this diagram an outlier and perform a further χ^2 goodness-of-fit test, with the data for diagram 11 removed, we still observe a significant difference with $p = 0.019$. Therefore, we posit that real-world diagram 11, illustrated in figure 11, to be a good example of how *not* to draw an Euler diagram if the information there in is to be interpreted accurately. On the other hand, guided diagram four accrued three more errors than its real-world

counterpart, which did not accrue any. Real-world diagram four broke six out of the ten guides.

With respect to time, guided diagrams performed significantly better. Eleven of the 12 guided diagrams yielded quicker mean times. In the most extreme case, the biggest difference in mean times is between diagram 5. Real-world diagram 5, which broke eight out of ten guides, yielded a mean time of 67.4 seconds. Its guided counterpart took only 21.7 seconds, a difference of 45.7 seconds. Therefore, we posit that real-world diagram five, illustrated in figure 8, to be a good example of how *not* to draw an Euler diagram if the information there in is to be interpreted quickly. On the other hand, real-world diagram ten had a mean time of 31.1 seconds will its guided counterpart had a mean time of 40 seconds, a difference of 8.9 seconds. Real-word diagram ten broke five out of the ten guides.

We observe that there is no direct correlation between the number of guides that a real-world diagram breaks and it performance, either in terms of error or time. Indeed, this is to be expected. For example, we have yet to ascertain if the use of colour, as defined by guide 10, will have a greater or smaller impact on performance compared with the use of shape, as defined by guide 7. However, the results of this study do illustrate that, in general, when real-world diagrams break multiple guides it is likely to significantly impact the accuracy and speed with which information is interpreted.

5. Threats to Validity

Threats to validity are categorized as internal, construct and external [24]. Internal validity considers whether confounding factors, such as carry-over effect, affects the results and, if so, to what extent. Construct validity examines whether the independent and dependent variables yield an accurate measure to test our hypotheses. External validity considers the extent to which we can generalise the results. The following discusses the primary threats to validity that were considered and addressed to ensure the study is robust and fit for purpose. With regard to internal validity, the following two factors were among a number that were considered in an attempt to manage potential disadvantages of our study design:

Carry-over effect: in a repeated measure experiment this threat occurs when the measure of a treatment is affected by the previous treatment. To manage this effect a between group design was employed. However, the two sets of diagrams were mixed within each group participant group, i.e. real-word

diagrams 1 to 6 and guided diagrams 7 to 12 were shown to participant group A. Guided diagrams 1 to 6 and real-world diagrams 7 to 12 were shown to participant group B.

Learning effect: the learning effect was considered a threat if questions were delivered in a fixed order. Therefore, questions were delivered to each participant in a random order. Further, to reduce the learning curve during the main study, participants were given appropriate training prior to the data collection phase.

Next we consider construct validity by focusing on our dependent variables (error rate and time) and the independent variable (diagram) respectively, and examine their rigour for measuring comprehension:

Error rate: all diagrams were drawn with consistent layout features reducing the possibility of confounding variables creeping into each diagram, for example inconsistent use of fonts. Further, only the known topological and graphical properties were present in the real-world diagrams. All other properties such as lines connecting curve labels to their corresponding curves were removed.

False negatives: to reduce false negatives, i.e. a participant selecting the wrong answer while reading it to be the correct answer, the similarity of pet and people names was minimised during all phases of the experiment.

False positives: to reduce false positives when answering 'How many' style questions i.e. a participant selecting the correct number of names but from the wrong region, the research vehicle also required participants to enter the peoples names having identified to be in the correct region.

Time: to ensure the rigour of time measurements, consideration was paid to the precise duration elapsed to answer each question. To this end, the same PC, located in a particular laboratory, with no applications running in the background was used by all participants.

Diagram: It was considered a threat if real-world and guided diagrams differed other than by their topological and graphical choices. This was addressed by drawing both the real-world and guided diagrams with consistent layout features. It was also considered a threat if participants did not spend time reading and understanding the diagrams. To manage this threat *diversity* was introduced in the diagrams so that participants had to read and understand each diagram before being able to answer the posed question. Further, it was also considered a threat if the diagrams were regarded as trivial; having only a few curves, zones, or data items was deemed insufficient to yield noticeable differences in response times, should they exist.

To manage this, diagrams were chosen to exhibit an appropriate level of complexity in order to demand cognitive effort.

The following factors consider the limitations of the results and the extent to which the they can be generalised, thus examining their external validity: *Diagrams*: The sample population consisted of 12 diagrams.
Set theoretic concepts: Euler diagrams conveyed set disjointness, subset and intersecting relationships.
Question styles: three styles of questions were asked: 'Who', 'Which' and 'How many'.
Participant: participants were representative of a wider student population. Thus, the results should be taken to be valid within these constraints.

6. Conclusion

This paper establishes the impact of topological and graphical properties for Euler diagrams that improve their comprehension. Our study gives a clear indication that a combination of Euler diagram guides should be used when drawing these diagrams because the guides significantly improve user comprehension of real-world data. Testing diagrams drawn by others against diagrams produced using our guides, our empirical study indicates significant improvements in accuracy with a halved error rate, down from 21.4% to 10.3%. Our timing results are also significant, with a 9 second improvement in the average time taken to solve problems, from 34.2 seconds to 24.9 seconds.

The implication of the results is far reaching. Indeed, not only do they provide a robust foundation for guidance when drawing Euler diagrams in general but they provide a profitable direction for researchers and developers creating tools that automatically generate Euler diagrams. Many of the Euler diagrams generated by the aforementioned automated layout tools fall considerably short of meeting the guides. Our empirical study has demonstrated that future layout tools should strive to adhere to guides 1 to 10. In doing so, the resulting Euler diagrams will be of far greater value to the target end users.

There are still a notable number of important research questions to be addressed. Perhaps the most immediate is that identified in section 4.3. We have yet to understand the extent one guide impacts upon performance when compared with another. That is to say, does the colour of an Euler diagram's closed curve have a greater impact on performance than the shape

of the closed curve? By establishing differences in performance between pairs of guides, we can then prioritise the application of these guides. This, again, will assist with the development of future layout tools in helping decide which guides need to be considered.

References

[1] Alper, B., Riche, N., Ramos, G., Czerwinski, M., 2011. Design study of linesets, a novel set visualisation technique. IEEE Transactions on Visualization and Computer Graphics 17, 2259–2267.

[2] Architectural Association, London, accessed April 2013. http://www.aadip9.net/shenfei/.

[3] Benoy, F., Rodgers, P., 2007. Evaluating the comprehension of Euler diagrams, in: 11th International Conference on Information Visualization, IEEE Computer Society. pp. 771–778.

[4] Bertin, J., 1983. Semiology of Graphics: Diagrams, Networks, Maps. University of Wisconsin Press.

[5] Blake, A., Stapleton, G., Rodgers, P., Cheek, L., Howse, J., 2012. Does the orientation of an Euler diagram affect user comprehension?, in: 18th International Conference on Distributed Multimedia Systems, Knowledge Systems Institute. pp. 185–190.

[6] Blake, A., Stapleton, G., Rodgers, P., Cheek, L., Howse, J., 2014a. Does the shape of an Euler diagram's closed curve affect user comprehension?, in: International Conference on the Theory and Application of Diagrams, Springer. pp. 124–138.

[7] Blake, A., Stapleton, G., Rodgers, P., Howse, J., 2014b. How should we use colour in Euler diagrams?, in: International Symposium on Visual Information Communication and Interaction, ACM. pp. 149–155.

[8] Blog.ffctn.com, accessed November 2014. http://blog.ffctn.com/what-is-data-visualization.

[9] Cargocollective.com, accessed November 2014. http://cargocollective.com/lizmeyer/Fast-Company-Venn.

[10] Collins, C., Penn, G., Carpendale, M.S.T., 2009. Bubble sets: Revealing set relations with isocontours over existing visualizations. IEEE Transactions on Visualization and Computer Graphics 15, 1009–1016.

[11] DeChiara, R., Erra, U., Scarano, V., 2005. A system for virtual directories using Euler diagrams, in: Proceedings of Euler Diagrams, pp. 33–53.

[12] Farrell, G., Sousa, W., 2001. Repeat victimization and hot spots: The overlap and its implication for crime control and problem-oriented policing. Crime Prevention Studies 12, 221–240.

[13] Gurr, C., 1999. Effective diagrammatic communication: Syntactic, semantic and pragmatic issues. Journal of Visual Languages and Computing 10, 317–342.

[14] Ip, E., 2001. Visualizing multiple regression. Journal of Statistics Education 9.

[15] Isenberg, P., Bezerianos, A., Dragicevic, P., Fekete, J., 2011. A study on dual-scale data charts, in: IEEE Transactions on Visualization and Computer Graphics, IEEE Computer Society. pp. 2469 – 2478.

[16] Kestler, H., Muller, A., Kraus, J., Buchholz., M., 2008. Vennmaster: Area-proportional Euler diagrams for functional GO analysis of microarrays. BMC Bioinformatics 9(1).

[17] Lukelab, accessed April 2013. http://www.lukelab.com/graphs/portrait-of-the-artists-interior-as-venn-diagram/.

[18] Meulemans, W., Henry Riche, N., Speckmann, B., Alper, B., Dwyer, T., 2013. Kelpfusion: A hybrid set visualization technique. IEEE Transactions on Visualization and Computer Graphics 19, 1846–1858.

[19] Miuml.org, accessed November 2014. http://www.miuml.org/news/2012/1/22/update-state-and-polymorphism-subsystems.html.

[20] Nature.com, accessed November 2014. http://www.nature.com/ncomms/journal/v3/n2/-fig_tab/ncomms1688_F3.html.

25

[21] Ofnumbers.com, accessed November 2014. http://www.ofnumbers.com/2014/03/04/chapter-3-next-generation-platforms/.

[22] Puppetlabs.com, accessed November 2014. https://docs.puppetlabs.com/puppet/latest/reference/lang_scope.html.

[23] Purchase, H., 1997. Which aesthetic has the greatest effect on human understanding?, in: 5th International Symposium on Graph Drawing, Springer. pp. 248–261.

[24] Purchase, H., 2012. Experimental Human Computer Interaction: A Practical Guide with Visual Examples. Cambridge University Press.

[25] Riche, N., Dwyer, T., 2010. Untangling Euler diagrams. IEEE Transactions on Visualisation and Computer Graphics 16, 1090–1097.

[26] Rodgers, P., , Zhang, L., Purchase, H., 2012. Wellformedness properties in Euler diagrams: Which should be used? IEEE Transactions on Visualization and Computer Graphics 18, 1089–1100.

[27] Rodgers, P., Stapleton, G., Howse, J., Zhang, L., 2010. Euler graph transformations for Euler diagram layout, in: Visual Languages and Human-Centric Computing, IEEE. pp. 111–118.

[28] Rodgers, P., Zhang, L., Fish, A., 2008. General Euler diagram generation, in: International Conference on the Theory and Application of Diagrams, Springer. pp. 13–27.

[29] Simonetto, P., 2012. Visualisation of Overlapping Sets and Clusters with Euler Diagrams. Ph.D. thesis. Université Bordeaux.

[30] Simonetto, P., Auber, D., 2008. Visualise undrawable Euler diagrams, in: 12th International Conference on Information Visualization, IEEE. pp. 594–599.

[31] Stackoverflow.com, accessed November 2014. http://stackoverflow.com/questions/12545027/venn-diagram-layout-with-d3-js.

[32] Stapleton, G., Flower, J., Rodgers, P., Howse, J., 2012. Automatically drawing Euler diagrams with circles. Journal of Visual Languages and Computing 12, 163–193.

[33] Stapleton, G., Rodgers, P., Howse, J., Zhang, L., 2009. Inductively generating Euler diagrams. IEEE Transactions of Visualization and Computer Graphics 17, 88–100.

[34] The American Physiology Society, accessed April 2013. http://www.physiology.org/.

[35] Thièvre, J., Viaud, M., Verroust-Blondet, A., 2005. Using Euler diagrams in traditional library environments, in: Euler Diagrams, ENTCS. pp. 189–202.

[36] Vannevar.blogspot.co.uk.co.uk, accessed November 2014. http://vannevar.blogspot.co.uk/2009/06/venn-diagrams-chart-geek-favorite.html.

[37] Wikipedia.org, accessed November 2014a. http://en.wikipedia.org/wiki/Homophone.

[38] Wikipedia.org, accessed November 2014b. http://en.wikipedia.org/wiki/File:SolarSystemBodies.png.

[39] Wilkinson, L., 2012. Exact and approximate area-proportional circular Venn and Euler diagrams. IEEE Transactions on Visualisation and Computer Graphics 18, 321–330.

7. Appendix 1 - Original Diagrams
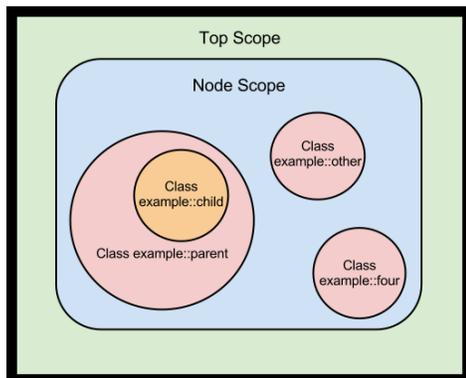
Figure 18: Original diagram 1 [22].
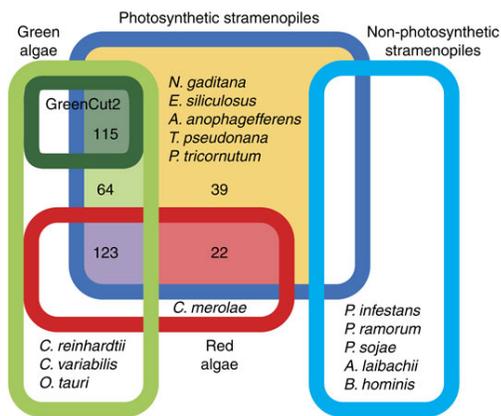
Figure 19: Original diagram 2 [20].
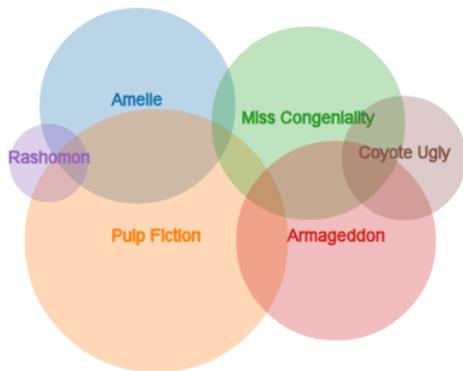
Figure 20: Original diagram 3 [31].
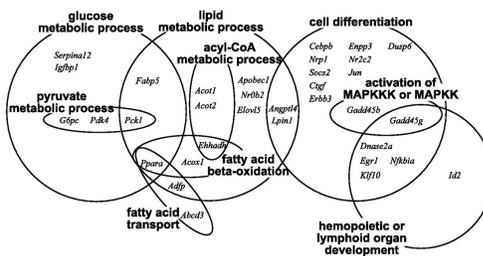
Figure 21: Original diagram 4 [34].
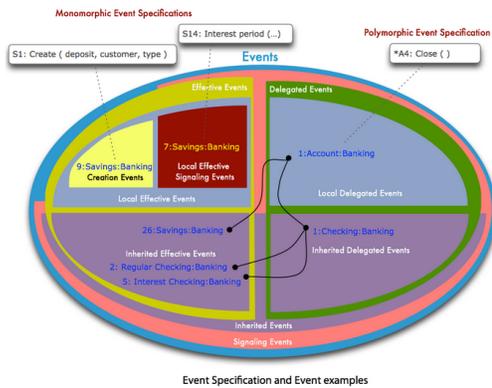
Figure 22: Original diagram 5 [19].
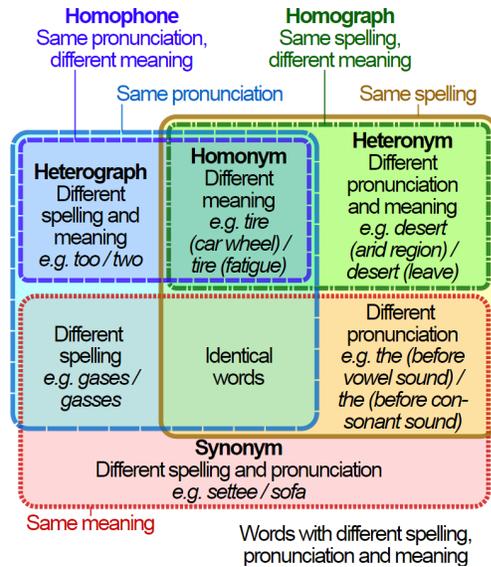
Figure 23: Original diagram 6 [37].
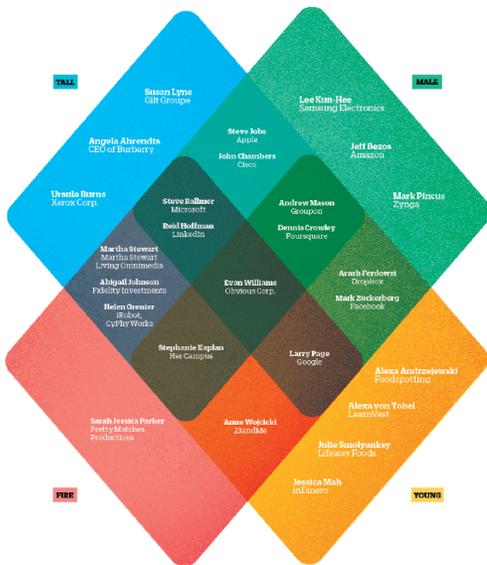
Figure 24: Original diagram 7 [9].
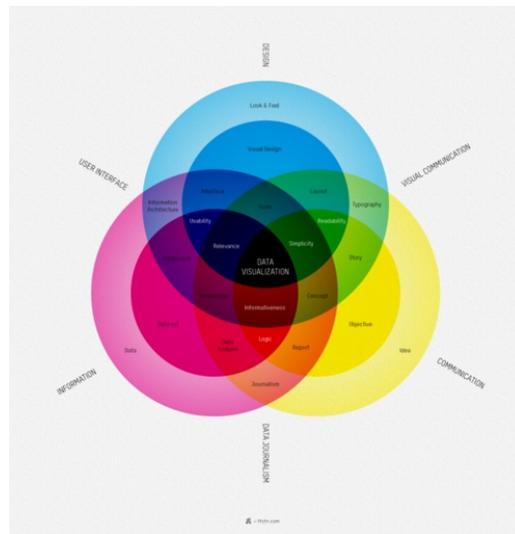
Figure 25: Original diagram 8 [8].

29
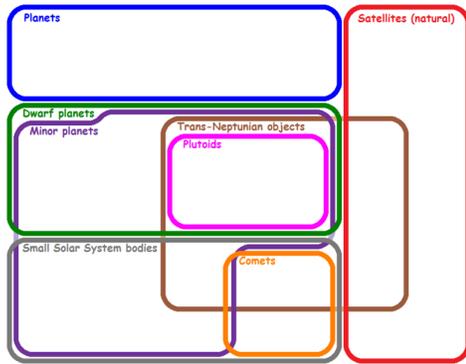
Figure 26: Original diagram 9 [38].
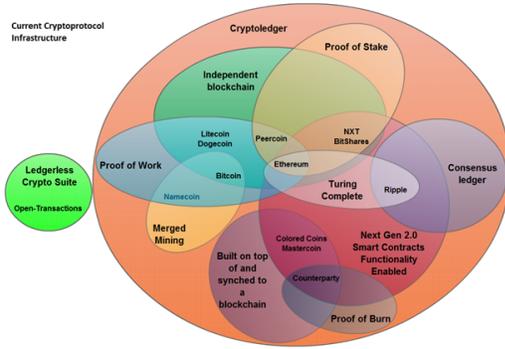
Figure 27: Original diagram 10 [21].
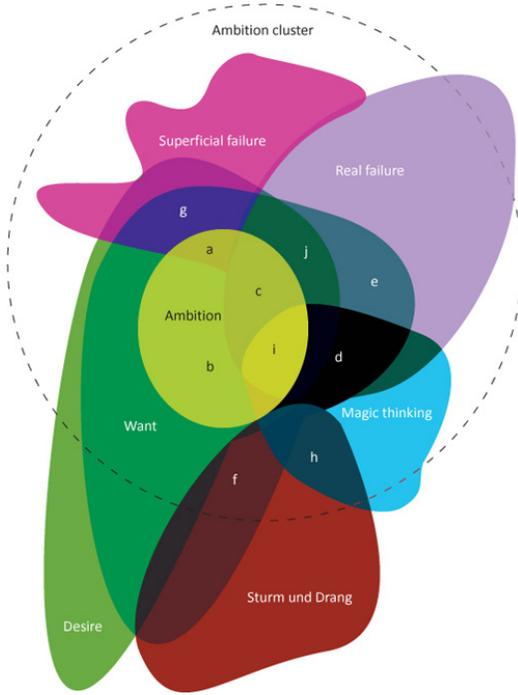
Figure 28: Original diagram 11 [17].
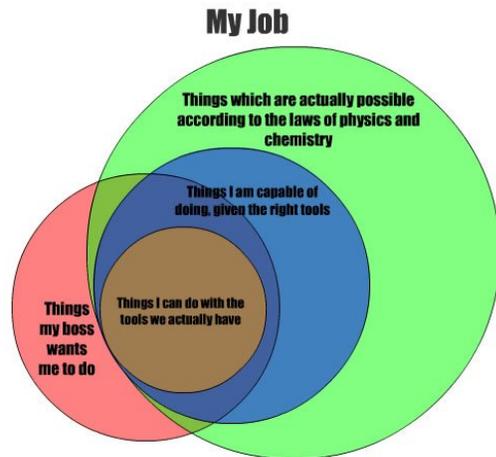
Figure 29: Original diagram 12 [36].

8. Appendix 2 - Real-world Diagrams
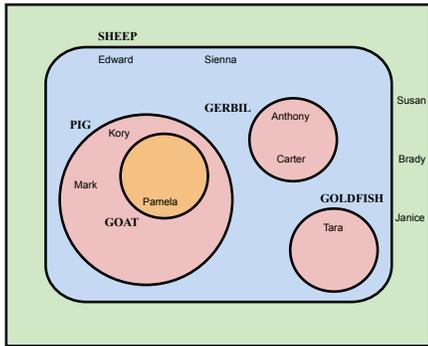
Figure 30: Real-world diagram 1.
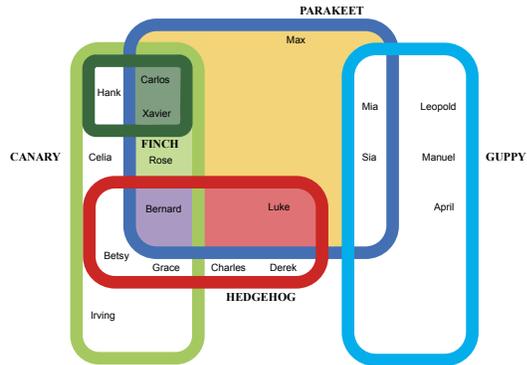
Figure 31: Real-world diagram 2.
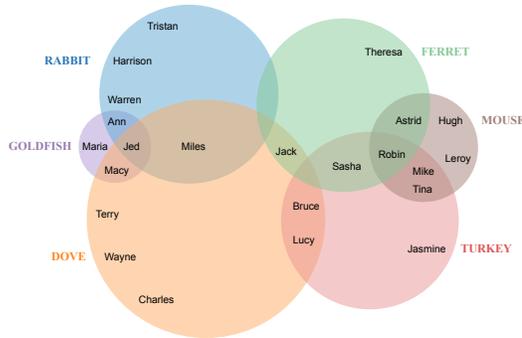
Figure 32: Real-world diagram 3.
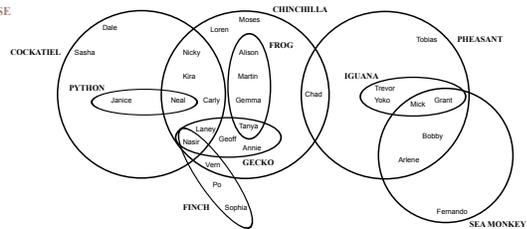
Figure 33: Real-world diagram 4.
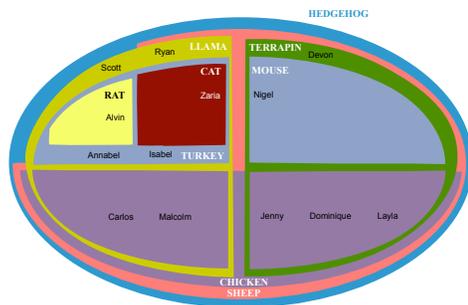
Figure 34: Real-world diagram 5.
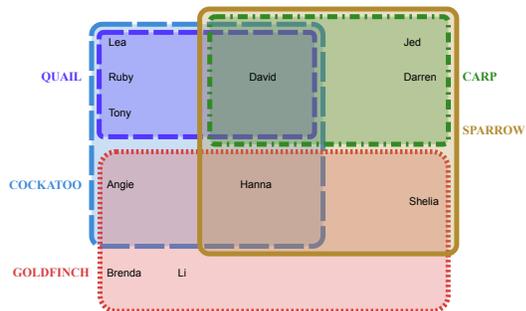
Figure 35: Real-world diagram 6.
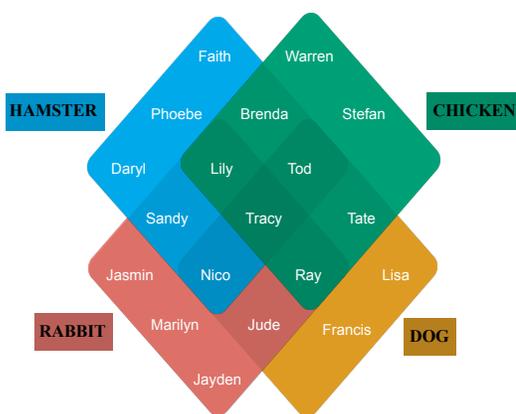
Figure 36: Real-world diagram 7.
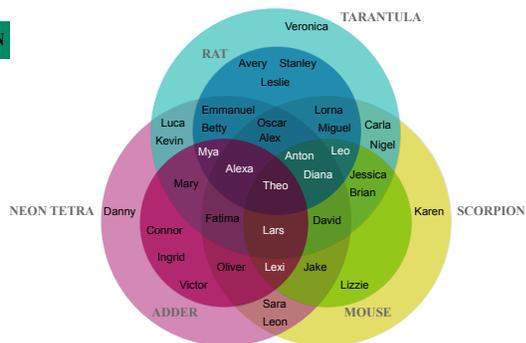
Figure 37: Real-world diagram 8.
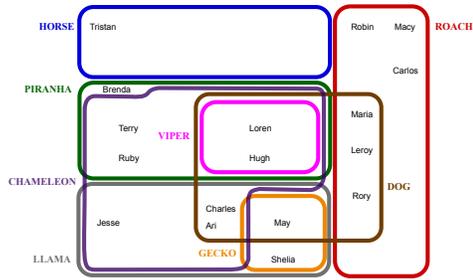
Figure 38: Real-world diagram 9.
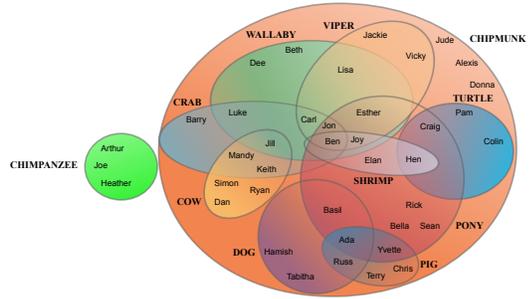
Figure 39: Real-world diagram 10.
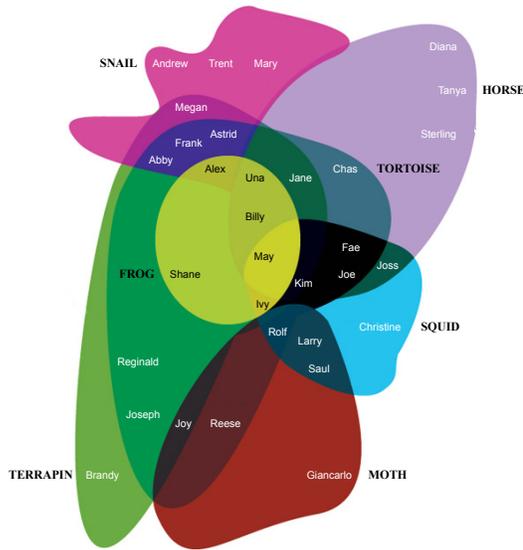
Figure 40: Real-world diagram 11.
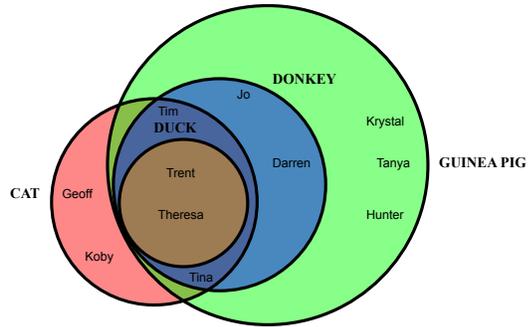
Figure 41: Real-world diagram 12.

9. Appendix 3 - Guided Diagrams

Figure 42: Guided diagram 1.

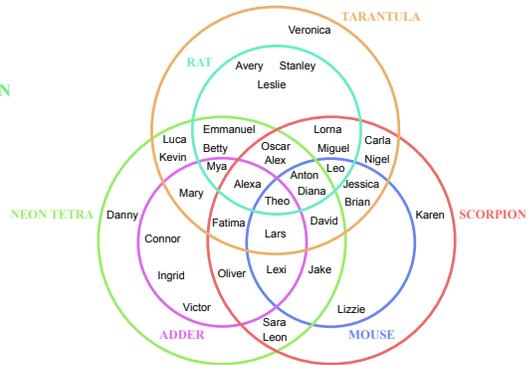Figure 43: Guided diagram 2.

Figure 44: Guided diagram 3.

Figure 45: Guided diagram 4.
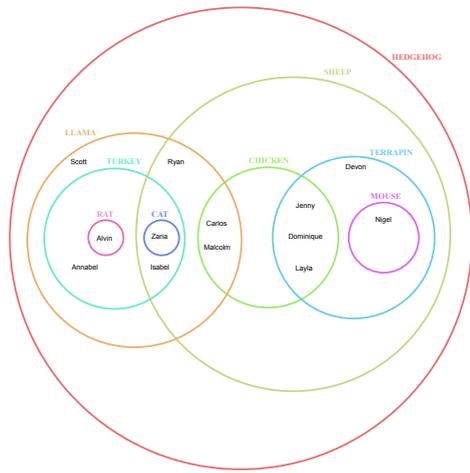
Figure 46: Guided diagram 5.
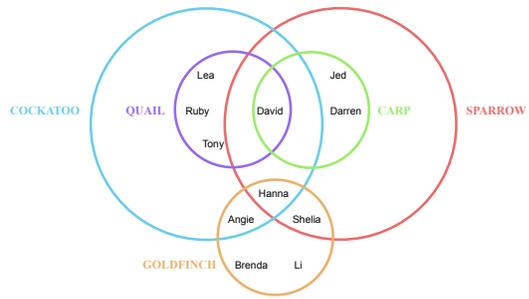
Figure 47: Guided diagram 6.
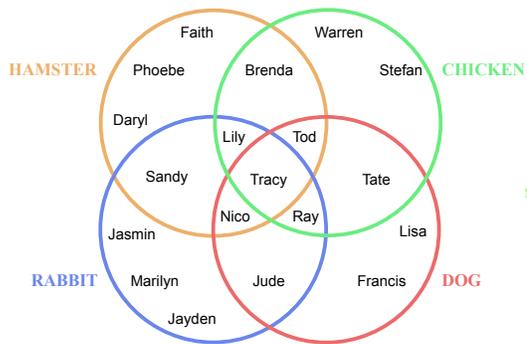
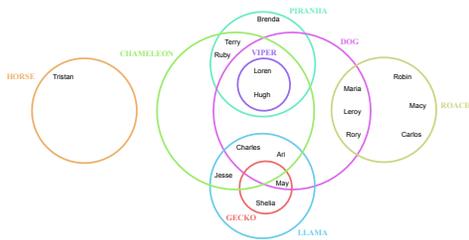Figure 48: Guided diagram 7.

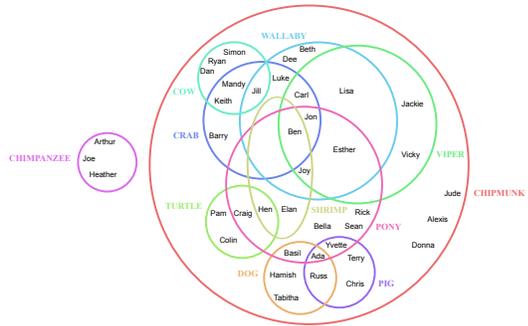Figure 49: Guided diagram 8.

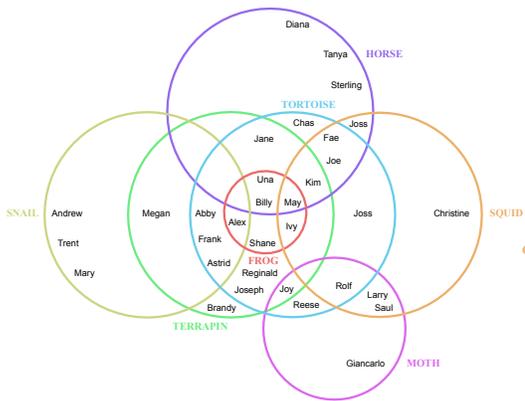Figure 50: Guided diagram 9.
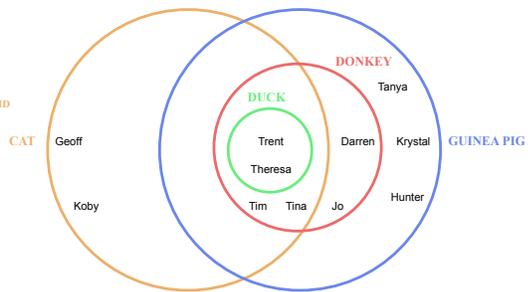

Figure 51: Guided diagram 10.

Figure 52: Guided diagram 11.

Figure 53: Guided diagram 12.