Conceptual Modelling with Euler^{+} Diagrams *

Uta Priss

Fakultät Informatik, Ostfalia University, Wolfenbüttel, Germany
www.upriss.org.uk

Abstract. This short paper introduces Euler^{+}diagrams as an enhanced version of traditional Euler diagrams and discusses how these can be utilised for conceptual modelling. Instead of the traditional interpretation of Euler diagrams as Boolean logic, Euler^{+}diagrams are considered 3-valued logic diagrams that are interpreted as First Order Logic (FOL) expressions. It is argued that such diagrams have a good usability because they are sufficiently simple yet reasonably expressive. Conditions for a translation between Euler^{+}diagrams and logical expressions and some consistency rules are provided. Questions still remain with respect to a detailed explanation of visual reasoning algorithms.

1 Introduction

Venn and Euler diagrams are frequently used as a tool for visualising logical and set theoretical expressions. A common interpretation of such diagrams evaluates existing zones into True and missing or shaded zones into False resulting in Boolean algebra. But set theory is more complex because the operations $\cup$ and $\cap$ result in sets, whereas $\subseteq$, $\subset$ and $=$ result in truth values. For example, $A \subseteq B$ is equivalent to (NOT $A) \cup B$ as a set-valued (Boolean) expression but to $\forall(x)\, x \in A \Rightarrow x \in B$ as a truth-valued, FOL expression. These two possible interpretations are not equivalent to each other because their negations are different as shown in Section 3. For a truth-valued interpretation, Euler diagrams should be assumed to be filled with 3 states: 'none', 'at least one' or 'any number of' elements. This enhanced version of Euler diagrams is introduced in this paper as 'Euler^{+}diagrams' which can additionally express functions and relations. Visually, the enhancement is simple: a quantifier and arrows are added to the diagrams. Furthermore in order to reduce complexity, diagrams can be split and concatenated using 'AND' and 'OR'. Last but not least, truth-valued set statements can also be added to Euler^{+}diagrams in a textual format.

One motivation for this paper were Chapman et al.'s (2011) 'Concept Diagrams' which present a different form of enhanced Euler diagrams used for modelling ontologies. Amongst several differences between Concept Diagrams and Euler^{+}diagrams, Concept Diagrams use a notation where dots represent variables which, in our opinion, is more difficult to visually parse. Nevertheless the visual reasoning algorithms described by Chapman et al. are relevant for Euler^{+}diagrams as well. A second motivation for this paper were discussions with students about Euler diagrams while teaching an introductory mathematics class. It highlighted the need for using simple notations

* Published in Ojeda-Aciego et al. Proc of ICCS'23, LNCS 14133, Springer-Verlag

that the students are already familiar with or learning anyway during the class. Students tend to be very critical users that point out any difficulties encountered when learning a notation. Previous experience showed that conventional diagrams for concept lattices (Ganter & Wille 1999) are not intuitive for students and require more teaching time (Priss 2017). Euler diagrams can express the same content as concept lattices (Priss 2023) but appear to be easier to read. Because students sometimes perceive diagrams for functions as in Fig. 1c also (incorrectly) as Euler diagrams, the idea arose to include arrows for functions in Euler$^+$diagrams as well. Apart from teaching purposes, we envision illustrations of scientific results for a general audience as a possible application of Euler$^+$diagrams.

Euler$^+$Diagrams support conceptual modelling because sets can be considered concepts as they have both an extensional listing of elements as well as an intensional, logical definition. For example, $\{x \mid x \in \mathbb{N} \text{ AND } x < 4\}$ has an extension $\{1, 2, 3\}$ and an intension 'natural numbers smaller than 4'. Set operations can be interpreted as conceptual operations. For example, if sets for 'dog' and 'pet' are defined, then so are 'dog AND pet' and 'dog OR pet'. Thus Euler$^+$diagrams visualise methods of concept formation and are also suitable for representing concept lattices (Priss 2023).

This short paper introduces Euler$^+$diagrams without presenting a detailed mathematical or logical description (which will be left for a future paper). Rodgers (2014) provides an overview of existing Euler diagram research. Later results can be found mostly in the DIAGRAMS conference series[1]. Stapleton, Shimojima & Jamnik (2018) discuss some aspects of existential quantifiers for Euler diagrams, but we believe that a clear distinction between (Boolean) Euler diagrams and more expressive Euler$^+$diagrams as presented in this paper is more convincing and more usable.

Section 2 presents a short definition of Euler$^+$diagrams and their semantics. Section 3 explains details and challenges of using Euler$^+$diagrams. Section 4 discusses how to add functions and relations to the diagrams. Section 5 presents some short examples of visual reasoning with Euler$^+$diagrams. The paper finishes with a conclusion.

2 Definition of Euler$^+$Diagrams

This section introduces Euler$^+$diagrams as an enhanced version of Euler diagrams (Fig. 1a and b). Euler diagrams consist of closed curves with labels representing sets. In this paper, Euler diagrams fulfil the 'well-formedness' condition that each visible area of the diagram is in one-to-one correspondence to a distinct intersection of sets. The areas are called 'zones' in this paper. For example, in Fig. 1a, exactly and only the zone just inside the outer zone corresponds to 'partial function $\cap$ NOT function'. Rodgers (2014) defines further terminology and well-formedness conditions which are not relevant for this paper. Venn diagrams are Euler diagrams that contain 2^n zones for n sets corresponding to all possible intersections (such as, $\emptyset$, A, B and $A \cap B$ for $n = 2$). In Euler diagrams, empty sets, such as a zone 'function AND NOT partial function' in Fig. 1a are left off or shaded as in Fig. 1b. It is not always possible to draw an Euler diagram without shading. For fewer than 4 sets, the curves of Euler diagrams can be drawn

[1] http://diagrams-conference.org/

as circles but for more than 3 sets, other curve shapes may be required. Priss (2023) explains that rounded rectangular curves as used in this paper have some advantages over the other shapes with respect to the number of diagrams that can be drawn.

Euler$^+$diagrams are Euler diagrams with the following enhancements:

D1 Diagrams can be combined using AND and OR and also with textual statements (truth-valued set expressions or definitions of sets using ':=' and set operations or relations).
D2 Elements of the sets can be written into the curves (as labels).
D3 Zones have 3 possible states: shaded, 'don't care' (nothing is written into the zone) or 'existential' (contains at least one element or an $\exists$).
D4 Arrows can be added between two zones or between two elements.

Some conditions are required:

C1 If a zone occurs more than once in a combined diagram, it must have the same state.
C2 The state of the outer zone must always be 'don't care'.
C3 Each curve and arrow must have exactly one label. Labels can occur more than once but only for the same item. The sets of labels for curves, sets and elements must be mutually disjoint. If it is clear what is meant, labels can sometimes be omitted.

The semantics of Euler$^+$diagrams is defined as follows:

S1 Labels of elements, curves and arrows are names of elements, sets and relations, respectively.
S2 For a combined diagram, each component is translated separately into a statement by interpreting each missing or shaded zone as a statement about not existing elements, each existential zone as a statement about existing elements and ignoring all 'don't care' zones. The resulting FOL statements are then combined with AND.
S3 Textual statements are interpreted as FOL statements.
S4 Arrows between zones are interpreted as binary relations. Arrow heads are in the middle of the lines for relations and at the end of the lines for functions. Arrows between elements are relation instances. An arrow head indicates a direction of a relation, for example, $a \leftarrow b$ corresponds to a pair (b, a). Relations and functions are interpreted as not empty and as total, i.e. all elements in the sets at both ends of the arrows must occur at least once in the relation.
S5 The negation of shaded or missing zones is existential zones and vice versa and the negation of 'don't care' is 'don't care'.

With further conditions:

C4 The set of drawn or deducible arrows is complete, i.e. whenever some elements in a zone relate to another zone, an arrow must exist between the two zones or be deducible from textual statements.
C5 Because relations are not empty, zones connected by arrows must be existential, i.e. contain an $\exists$-quantifier.

The conditions are not sufficient for avoiding contradictory diagrams but FOL also does not have conditions that stop a user from writing '$A = B$ AND $A \neq B$'. Thus, a diagram is non-contradictory if all its statements combine to an FOL statement that is free of contradictions. Further details about what is meant by some of the points of this definition are explained in the remainder of the paper. The focus of this paper is on the

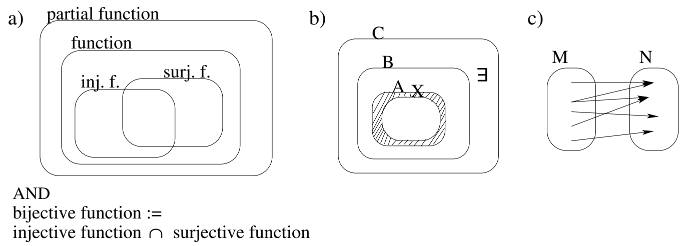

Fig. 1. Euler$^+$diagrams (left and middle) and a diagram of a function (right)

graphical aspects not on a more detailed description of formal semantics which is left for a future paper.

A challenge for Euler diagrams is that they can easily become too complex to be usable. Euler$^+$diagrams overcome this challenge by allowing to split a diagram into many parts which are then combined with AND and OR. Furthermore, textual statements (that are equivalent to Euler$^+$diagrams) are allowed because sometimes a diagram is simpler, sometimes a textual expression is simpler. Obviously this poses a new question as to how to split a diagram in a manner that still supports visual reasoning about facts that are distributed across different parts. The condition C1 avoids some problems. For example, it would not be useful to split $A = B$ into $A \subseteq B$ in one component and $B \subseteq A$ in another. Most likely OR should be used extremely sparingly for combining diagrams. NOT is only allowed as a set operation but not for combining statements. Combining and splitting diagrams is discussed, for example, by Priss (2021 and 2023).

3 Expressing Logical Statements with Euler$^+$Diagrams

Fig. 1a displays an Euler$^+$diagram visualising conceptual information, such as the fact that functions are partial functions. It is a typical diagram that might be used in the context of teaching showing students that functions are partial functions contrary to natural language where a noun modified by an adjective tends to denote a subconcept of the unmodified noun. The diagram also expresses a definition of 'bijective function' but as a textual statement because otherwise an intersection would need to be labelled which is difficult to visually parse. Fig. 1b demonstrates transitivity of the set containment relation ($A \subseteq B$ AND $B \subseteq C \Longrightarrow A \subseteq C$). The existence quantifier indicates that $B \subset C$. Whether $A \subset B$ is not known (i.e. 'don't care'). The diagram further shows an example of shading. If curves have exactly one label (C3), then showing the equality of sets ($A = X$) either requires shading or a textual statement 'AND $X := A$'. A textual statement would be clearer in this case.

Elements of sets can be written into the zones as in Fig. 2c ($a \in F$). Sets can only be shown as subsets but not as elements of other sets unless they are written as strings ($\{\} \in F$). Fig. 2 highlights difficulties expressing empty sets in Euler diagrams which also affect Euler$^+$diagrams. An empty set can either be shaded ($A \cap B \cap C = \emptyset$) or missing ($D \cap E = \emptyset$). The empty set J is a subset of $G \cap H \cap I$. But there are no graphical clues showing that J must not be drawn in any other location than the intersection of all

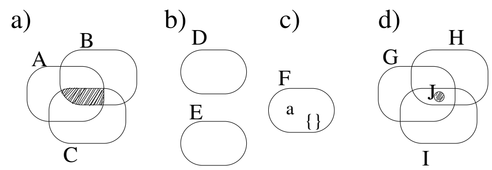

Fig. 2. Euler diagrams and the empty set

other sets, that it is a subset of $G \cap H \cap I$ even if not shown in the diagram and that all empty sets are equal to each other. These challenges may not be caused by the diagrams but by the fact that 'empty' tends to be a difficult concept.

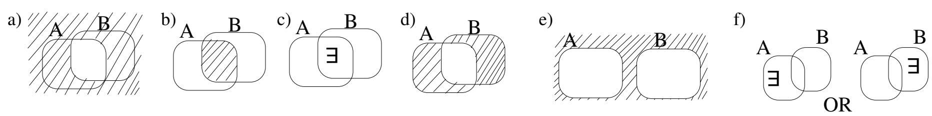

Fig. 3. Euler diagram negation: binary or truth-valued

Fig. 3 shows the difference between set- and truth-valued diagrams with respect to negation. According to C2, Fig. 3a and 3e are Euler diagrams, but not Euler$^+$diagrams because their outer zones are shaded. The set-valued negation of $A \cap B$ (Fig. 3a) is in 3b whereas 3b and 3c are truth-valued negations of each other. The set-valued negation of $A = B$ (in Fig. 3d) is in 3e and its truth-valued negation in 3f, in this case resulting in two diagrams connected with 'OR'. Fig. 4 shows all 4 possible quantifiers that can result from translating an Euler$^+$diagram into an FOL statement. A symbol for an all-quantifier is not included in the definition of Euler$^+$diagrams because it is implied by missing zones. According to S2, 'don't care' zones are ignored. If the quantifier in Fig. 4b was missing, then it would contain four 'don't care' zones and be interpreted as an empty FOL statement. Translations between Euler$^+$diagrams and FOL should be equivalent, but are not unique. For example, Fig. 4c can also be expressed as $\forall(x \in B)\, x \notin A$. Furthermore, Euler diagrams can always be drawn in different manners. A proof of logical equivalence of interpretations could follow strategies employed by Chapman et al (2011) and similar publications but is not included in this short paper.

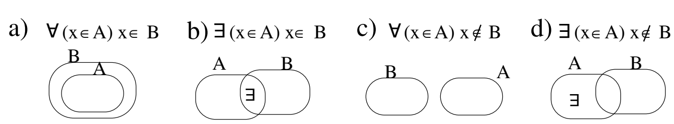

Fig. 4. Expressing quantifiers: a) ALL, b) SOME, c) NONE, d) NOT ALL

4 Euler$^+$Diagrams, Functions and Relations

As mentioned in the introduction, operations on sets or concepts support the formation of further concepts. For example, the concepts ‘pet’ and ‘cat’ support a discussion about ‘pet cats’. If functions or relations are added to the mixture, further sets or concepts can be defined. For example, a relation ‘childOf’ generates a set of parents and a set of children and a verb ‘to see’ distinguishes objects that can see or can be seen. For a universal set $\mathbb{U}$ of elements, a relation r and subsets $A, B \subseteq \mathbb{U}$, one can define the sets $r^{\lhd}(B) := \{x \mid x \in \mathbb{U}, \exists (b \in B)(x, b) \in r\}$ and $r^{\rhd}(A) := \{x \mid x \in \mathbb{U}, \exists (a \in A)(a, x) \in r\}$. It follows that $r \subseteq r^{\lhd}(\mathbb{U}) \times r^{\rhd}(\mathbb{U})$ is a relation that is total on both sides, which means that every element in $r^{\lhd}(\mathbb{U})$ relates to at least one element in $r^{\rhd}(\mathbb{U})$ and vice versa. For functions one can write the usual $f(A)$ instead of $f^{\rhd}(A)$. It follows that $f^{-1}(B) = f^{\lhd}(B)$ for bijective functions, $r^{\lhd}(\mathbb{U}) = r^{\lhd}(r^{\rhd}(\mathbb{U}))$, $r^{\rhd}(\mathbb{U}) = r^{\rhd}(r^{\lhd}(\mathbb{U}))$ and $f(f^{\lhd}(f(A))) = f(A)$ for functions. But in general $r^{\rhd}(r^{\lhd}(r^{\rhd}(A))) \neq r^{\rhd}(A)$ is possible. For example for a translation relation between English and Irish, one can start with an Irish word, look up its English translation, then their Irish translations and so on - a process that might only stop after many iterations or when $r^{\lhd}(\mathbb{U})$ and $r^{\rhd}(\mathbb{U})$ have been reached.

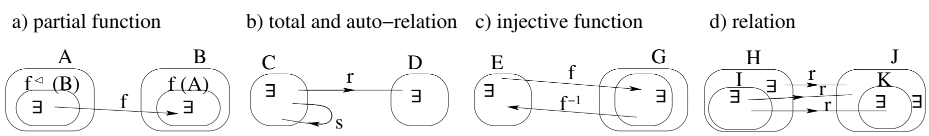

Fig. 5. Euler$^+$diagrams for functions and relations

Fig. 5a shows a partial function $f : A \nrightarrow B$ with $f : f^{\lhd}(B) \rightarrow B$ and $f(A) = f(f^{\lhd}(B)) \subseteq B$. Fig. 5b contains a total relation $r \subseteq C \times D$ and a total auto-relation $s \subseteq C \times C$. Because of C5, arrows connect zones that are existential. Furthermore C4 implies that if a zone I has more than one arrow for a single relation r then $J = r^{\rhd}(I)$ is true for the union J of the outermost zones (Fig. 5d). For a zone K with one arrow, the zone at the other end of the arrow can be defined using K and r as $I := r^{\lhd}(K)$. For the two outermost zones H and J, C4 implies that $H = r^{\lhd}(r^{\rhd}(H)) = r^{\lhd}(J)$ and $J = r^{\rhd}(r^{\lhd}(J)) = r^{\rhd}(H)$. If these two equations hold, sets H and J are called a ‘closed pair’ in this paper. Closed pairs can be modelled as concepts using Formal Concept Analysis (Ganter & Wille 1999) but that is left for another paper. Because of C4, a relation between sets with many subsets will lead to many arrows. Most likely separate diagrams should therefore be used for each function or relation. Instead of labelling each arrow, different colours could be used. Furthermore, a reduced drawing of arrows can be employed drawing only the arrows for closed pairs and adding textual statements that imply the remaining arrows. For example, in Fig. 5d only the arrow between H and J might be drawn and a statement ‘$I := r^{\lhd}(K)$’ added. Full and reduced drawing of arrows must not be mixed.

5 Reasoning with Diagrams

This section translates two examples from Chapman et al (2011) into Euler$^+$ diagrams. A translation of the first example, shown in Fig. 6, yields a function isPetOf: isPetOf$^{\lhd}(\mathbb{U}) \rightarrow$ person and a statement 'isPetOf$^{\lhd}(\mathbb{U}) \subseteq$ animal AND isPetOf(Rex) = Mick' which implies 'Rex $\in$ animal AND Mick $\in$ person' involving reasoning about the fact that if a function is applicable to instances then the instances must be elements of the domain and codomain of the function. Such reasoning is more easy to see in the diagrams than using the FOL statements. In this case the top right and left diagrams should be mentally combined into one diagram. The bottom right diagram summarises all of the information.

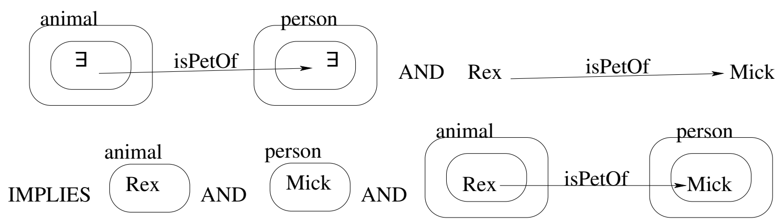

Fig. 6. Example of reasoning with diagrams

A second, slightly more complex example, is shown in Fig. 7. Translated into FOL, it defines a relation drives $\subseteq$ drives$^{\lhd}(\mathbb{U})\times$ drives$^{\rhd}(\mathbb{U})$ with driver := drives$^{\lhd}$(vehicle). The definition of driver is implied by the diagram because there is only one arrow into drives$^{\rhd}(\mathbb{U}) \cap$ vehicle. Thus drivers are people who drive at least one vehicle and possibly other non-vehicles. The statement 'drives$^{\lhd}(\mathbb{U}) \subseteq$ person AND driver $\subseteq$ adult AND ABC1 $\in$ vehicle AND drives(Mick) = ABC1' then implies 'Mick $\subseteq$ adult'. Visual reasoning consists of mentally inserting the diagram about the relation instance (Mick, ABC1) into the top right diagram using the fact that ABC1 is a vehicle and then realising that the top left diagram applies. The same information is contained in the FOL statements but these are more difficult to visually parse and combine without writing down each step.

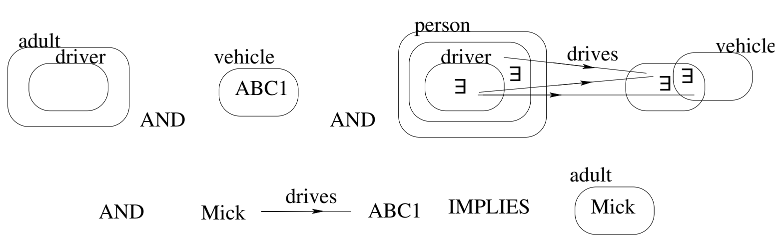

Fig. 7. Further example of reasoning with diagrams

6 Conclusion

This paper introduces Euler$^+$ diagrams as a means for visually representing statements about sets, functions and relations. A goal of this research is to produce simple diagrams for representing set (or conceptual) statements to a general audience of people who are not or not yet trained in mathematics. By splitting information into separate diagrams, complexity issues of larger Euler diagrams can be avoided by Euler$^+$ diagrams. Advantages of Euler diagrams for reasoning are known from the literature (eg. Stapleton et al. 2018) and apply to Euler$^+$ diagrams as well. Euler$^+$ diagrams do not solve consistency checking of statements. But because Euler$^+$ diagrams can be translated into FOL statements, algorithms for consistency checking of FOL statements can also be utilised for Euler$^+$ diagrams - although a more precise translation algorithm still needs to be provided in a future paper. Some rudimentary software for generating Euler diagrams from expressions is currently in development[2]. It is planned to extend this software into a tool for Euler$^+$ diagram generation and modification that is compatible with other software for conceptual structures, such as Formal Concept Analysis, Conceptual Graphs and ontologies.

References

1. Chapman, P.; Stapleton, G.; Howse, J.; & Oliver, I. (2011). Deriving sound inference rules for concept diagrams. In 2011 IEEE Symposium on Visual Languages and Human-Centric Computing. p. 87-94
2. Ganter, B.; Wille, R. (1999). Formal Concept Analysis. Mathematical Foundations. Berlin-Heidelberg-New York: Springer.
3. Priss, U. (2017). Learning Thresholds in Formal Concept Analysis. In: Bertet; et al. (eds.), Formal Concept Analysis. Proceedings of ICFCA'2017, Springer Verlag, LNCS 10308, p. 198-210.
4. Priss, U. (2021). Modelling Conceptual Schemata with Formal Concept Analysis. In: Proc. of FCA4AI'21.
5. Priss, U. (2023). Representing Concept Lattices with Euler Diagrams. In: Proc. of ICFCA'23. (to appear).
6. Rodgers, P. (2014). A survey of Euler diagrams. Journal of Visual Languages & Computing, 25, 3, p. 134-155.
7. Stapleton, G., Shimojima, A., & Jamnik, M. (2018). The observational advantages of Euler diagrams with existential import. In Diagrammatic Representation and Inference: 10th International Conference, Diagrams 2018, Springer, p. 313-329.

[2] https://upriss.github.io/educaJS/