Corresponding Regions in Euler Diagrams

John Howse, Gemma Stapleton, Jean Flower, and John Taylor

School of Computing & Mathematical Sciences
University of Brighton, Brighton, UK
{John.Howse,G.E.Stapleton,J.A.Flower,John.Taylor}@bton.ac.uk
http://www.it.bton.ac.uk/research/vmg/VisualModellingGroup.html

Abstract. Euler diagrams use topological properties to represent set-theoretical concepts and thus are 'intuitive' to some people. When reasoning with Euler diagrams, it is essential to have a notion of correspondence among the regions in different diagrams. At the semantic level, two regions correspond when they represent the same set. However, we wish to construct a purely syntactic definition of corresponding regions, so that reasoning can take place entirely at the diagrammatic level. This task is interesting in Euler diagrams because some regions of one diagram may be missing from another. We construct the correspondence relation from 'zones' or minimal regions, introducing the concept of 'zonal regions' for the case in which labels may differ between diagrams. We show that the relation is an equivalence relation and that it is a generalization of the counterpart relations introduced by Shin and Hammer.

1 Introduction

Euler diagrams [1] illustrate relations between sets. This notation uses topological properties of enclosure, exclusion and intersection to represent the set-theoretic notions of subset, disjoint sets, and intersection, respectively. The diagram d_2 in figure 1 is an Euler diagram with interpretation A is disjoint from B. Venn [13] adapted Euler's notation to produce a system of diagrams representing logical propositions. In a Venn diagram all intersections between contours must occur. The diagram d_1 in figure 1 is a Venn diagram. Some extensions of Euler diagrams allow shading, as in Venn diagrams, but since we are interested in a syntactic correspondence between regions shading is irrelevant. Thus we treat Venn diagrams as a special case of Euler diagrams, and ignore shading. Peirce [10] extended Venn's notation to include existential quantification and disjunctive information.

Shin [11] developed sound and complete reasoning rules for a system of Venn-Peirce diagrams. This work was seminal in that the rules were stated at the diagrammatic level and all reasoning took place at that level. This was the first complete formal diagrammatic reasoning system; until then diagrammatic reasoning was a mixture of informal reasoning at the diagrammatic level and formal (and informal) reasoning at the semantic level. Hammer [3] developed a sound and complete set of reasoning rules for a simple Euler system; it only considered inferences from a single diagram and contained only three reasoning rules.

In order to compare regions in different diagrams, Shin and Hammer developed counterpart relations [4, 11]. This paper considers an alternative, but related, approach to these counterpart relations and generalizes it to comparing regions in Euler diagrams. The solution of this problem is very important in extending diagrammatic reasoning to systems which have practical applications. Euler diagrams form the basis of more expressive diagrammatic notations such as Higraphs [5] and constraint diagrams [2], which have been developed to express logical properties of systems. These notations are used in the software development process, particularly in the modelling of systems and frequently as part of, or in conjunction with, UML [9]. Indeed, some of the notations of UML are based on Euler diagrams. The development of software tools to aid the software development process is very important and it is essential that such tools work at the diagrammatic level and not at the underlying semantic level so that feedback is given to developers in the notations that they are using and not in some mathematical notation that the developers may find difficult to understand. Thus it is necessary to construct a purely syntactic definition of corresponding regions across diagrams.

The task of defining such a correspondence relation is interesting, and very much non-trivial, in Euler diagrams because some regions of one diagram may be missing from another. For example, in figure 1 the region within the contours A and B in d_1 is missing from d_2. Diagram d_1 asserts that $A \cap B$ may or may not be empty, whereas d_2 asserts that $A \cap B = \emptyset$. What are the corresponding regions in this case?

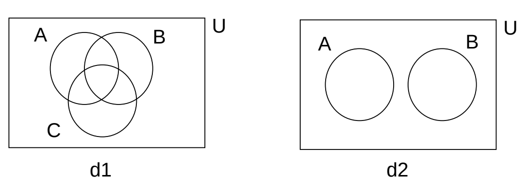

Fig. 1. A Venn diagram and an Euler diagram.

In §2 we give a concise informal description of Euler diagrams and a formal definition of its syntax. In §3 we define the correspondence relation between regions in the more straightforward case of Venn diagrams. In §4 we discuss the problems of defining corresponding regions in Euler diagrams and in the particularly difficult case of a system involving the disjunction of diagrams, before giving a general definition of the correspondence relation and showing that it is an equivalence relation. We then show, in §5, that it is a generalization of the counterpart relations developed by Shin and Hammer.

2 Syntax of Euler Diagrams

We now give a concise informal description of Euler diagrams. A *contour* is a simple closed plane curve. A *boundary rectangle* properly contains all other contours. Each contour has a unique label. A *district* (or *basic region*) is the bounded area of the plane enclosed by a contour or by the boundary rectangle. A *region* is defined, recursively, as follows: any district is a region; if r_1 and r_2 are regions, then the union, intersection and difference of r_1 and r_2 are regions provided these are non-empty. A *zone* (or *minimal region*) is a region having no other region contained within it. Contours and regions denote (possibly empty) sets. Every region is a union of zones. In figure 2 the zone within A, but outside B is missing from the diagram; the set denoted by such a "missing" zone is empty. An Euler diagram containing all possible zones is called a Venn diagram.

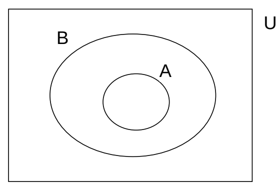

Fig. 2. An Euler diagram.

Given two diagrams we can connect them with a straight line to produce a *compound diagram* [6]. This connection operation is interpreted as the disjunction of the connected diagrams. A multi-diagram is a collection of compound diagrams and is interpreted as the conjunction of the compound diagrams. In this system a multi-diagram is in conjunctive normal form (cf. Shin's Venn II system [11]). In figure 3 diagrams d_1 and d_2 are to be taken in disjunction, thus $\{d_1, d_2\}$ is a compound diagram, as is $\{d_3\}$ (any unitary diagram is a compound diagram); the diagram $\{\{d_1, d_2\}, \{d_3\}\}$ is a multi-diagram.

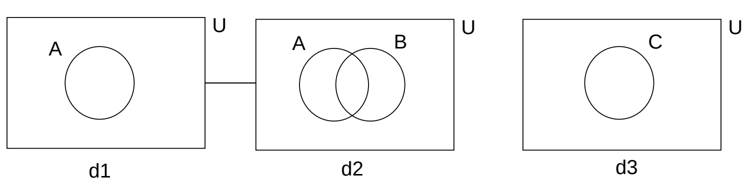

Fig. 3. Two compound diagrams.

A *unitary Euler diagram* is a tuple $d = \langle L, U, Z \rangle = \langle L(d), U(d), Z(d) \rangle$ whose components are defined as follows:

1. L is a finite set whose members are called *contours*. The element U, which is not a member of L, is called the *boundary rectangle*.
2. The set $Z \subseteq \mathbb{P}L$ is the set of *zones*. A zone $z \in Z$ is *incident* on a contour $c \in L$ if $c \in z$. Let $R = \mathbb{P}Z - \emptyset$ be the set of *regions*.

If $Z = \mathbb{P}L$, d is a *Venn diagram*. At this level of abstraction we identify a contour and its label. A zone is defined by the contours that contain it and is thus represented as a set of contours. The set of labels of a zone, z, is thus $L(z) = z$. A region is just a non-empty set of zones. The Euler diagram d in figure 2 has $L(d) = \{A, B\}$ and $Z(d) = \{\emptyset, \{B\}, \{A, B\}\}$.

A *compound diagram*, D, is a finite set of unitary diagrams taken in disjunction. A *multi-diagram*, Δ, is a finite set of compound diagrams taken in conjunction [6]. The set of labels of a compound diagram, D, is $L(D) = \bigcup_{d \in D} L(d)$. The set of labels of a multi-diagram, Δ, is $L(\Delta) = \bigcup_{D \in \Delta} L(D)$. In figure 3

$$L(\{\{d_1, d_2\}, \{d_3\}\}) = \{A, B, C\}.$$

3 Venn Diagrams

We will identify corresponding regions across Venn diagrams that do not necessarily have the same label sets. As an example, region $\{z_1, z_2, z_3, z_4\}$ in d_1 and region $\{z_5, z_6\}$ in d_2 in figure 4 are corresponding. We introduce the concept of a zonal region in order to identify this formally. Intuitively a zonal region is a region that becomes a zone when contours are removed. This is illustrated in figure 4. The contour with label C is removed and region $\{z_1, z_2\}$ becomes a zone, $\{z_5\}$, in the second diagram.

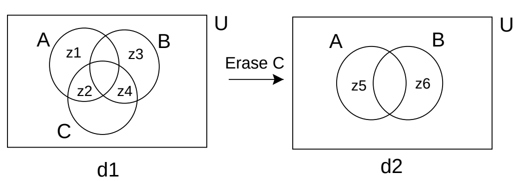

Fig. 4. Two Venn diagrams with different label sets.

3.1 Zonal Regions and Splits

In figure 4, consider how we might describe or identify the region $\{z_1, z_2\}$. Informally, it has description 'everything inside A but outside B'. Thus we associate $\{z_1, z_2\}$ with an ordered pair of sets, $\{A\}$ and $\{B\}$, which we shall write as

$\langle\{A\},\{B\}\rangle$. Similarly the region $\{z_1\}$ is associated with $\langle\{A\},\{B,C\}\rangle$, intuitively meaning 'everything inside A, but outside B and C'. In order to define zonal regions formally, and to allow us to compare regions across diagrams, we introduce the notion of a 'split'.

Definition 1. *A* ***split*** *is a pair of sets, $\langle P,Q\rangle$, such that $P\cap Q=\emptyset$; if $P\cup Q\subseteq X$ then $\langle P,Q\rangle$ is said to be a* ***split on*** *X.*

Addition is defined on splits with the following axioms:

1. $\langle P_1,Q_1\rangle=\langle P_2,Q_2\rangle \Leftrightarrow P_1=P_2\wedge Q_1=Q_2$
2. $\forall\, A\notin P\cup Q,\quad \langle P,Q\rangle=\langle P\cup\{A\},Q\rangle+\langle P,Q\cup\{A\}\rangle$
3. $\sum_{i=1}^{n}\langle P_i,Q_i\rangle=\sum_{j=1}^{m}\langle R_j,S_j\rangle$ if $\forall\, i\,\exists\, j \bullet \langle P_i,Q_i\rangle=\langle R_j,S_j\rangle$ and $\forall\, j\,\exists\, i \bullet \langle R_j,S_j\rangle=\langle P_i,Q_i\rangle$

Lemma 1. *Addition is commutative and associative. Each element is idempotent. If $\langle P,Q\rangle$ is a split and S is a finite set such that $(P\cup Q)\cap S=\emptyset$ then*

$$\langle P,Q\rangle=\sum_{W\subseteq S}\langle P\cup W,Q\cup(S-W)\rangle$$

This lemma follows from axioms 2 and 3. The last part of the lemma generalizes axiom 2 and is illustrated below.

$$\begin{aligned}\langle\{A\},\{B\}\rangle &= \sum_{W\subseteq\{C,D\}}\langle\{A\}\cup W,\{B\}\cup(\{C,D\}-W)\rangle\\ &= \langle\{A\},\{B,C,D\}\rangle+\langle\{A,D\},\{B,C\}\rangle+\\ &\quad\langle\{A,C\},\{B,D\}\rangle+\langle\{A,C,D\},\{B\}\rangle\end{aligned}$$

Definition 2. *For unitary Venn diagram d, let $\langle P,Q\rangle$ be a split on $L(d)$. Then the* ***zonal region*** *associated with $\langle P,Q\rangle$ is*

$$\{z\in Z(d):P\subseteq L(z)\wedge Q\subseteq\overline{L(z)}\}$$

where $\overline{L(z)}=L(d)-L(z)$, [8].

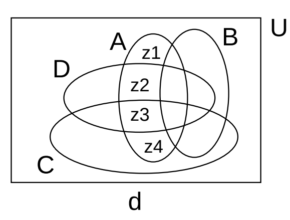

Fig. 5. Venn-4.

In figure 5, zonal regions $\{z_1\}$, $\{z_2\}$, $\{z_3\}$ and $\{z_4\}$ are associated with $\langle\{A\}, \{B,C,D\}\rangle$, $\langle\{A,D\},\{B,C\}\rangle$, $\langle\{A,C,D\},\{B\}\rangle$ and $\langle\{A,C,\},\{B,D\}\rangle$. The zonal region $\{z_1,z_2,z_3,z_4\}$ is associated with $\langle\{A\},\{B\}\rangle$. We have $\{A\} = L(z_1) \cap L(z_2) \cap L(z_3) \cap L(z_4)$ and $\{B\} = \overline{L(z_1)} \cap \overline{L(z_2)} \cap \overline{L(z_3)} \cap \overline{L(z_4)}$.

Lemma 2. *For any unitary Venn diagram d, if a zonal region zr is associated with $\langle P,Q\rangle$ then $P = \bigcap_{z\in zr} L(z)$ and $Q = \bigcap_{z\in zr} \overline{L(z)}$.*

Hence each zonal region is associated with a unique split. There is a parallel between axiom 2, $\langle P,Q\rangle = \langle P\cup\{A\},Q\rangle + \langle P,Q\cup\{A\}\rangle$, and lemma 3 below.

Lemma 3. ***The split rule.*** *Let zr be a zonal region of unitary Venn diagram d. If zr is associated with $\langle P,Q\rangle$ and $A \in L(d)-(P\cup Q)$ then $zr = zr_1 \cup zr_2$ where zr_1 and zr_2 are zonal regions associated with $\langle P\cup\{A\},Q\rangle$ and $\langle P,Q\cup\{A\}\rangle$ respectively.*

For example, consider the diagram in figure 6. Zonal regions $zr_1 = \{z_2,z_4,z_5,z_6\}$, $zr_2 = \{z_2,z_4\}$ and $zr_3 = \{z_5,z_6\}$ are associated with $\langle\{B\},\emptyset\rangle$, $\langle\{A,B\},\emptyset\rangle$ and $\langle\{B\},\{A\}\rangle$ respectively. From the split rule $zr_1 = zr_2 \cup zr_3$ and from axiom 2,

$$\langle\{B\},\emptyset\rangle = \langle\{A,B\},\emptyset\rangle + \langle\{B\},\{A\}\rangle.$$

Informally, this is splitting zr_1 into the part contained in A and the part excluded from A. A more general version of the split rule now follows.

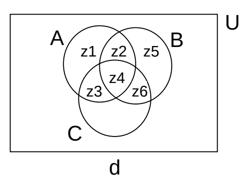

Fig. 6. Venn-3.

Corollary 1. ***The derived split rule.*** *If zr is a zonal region of unitary Venn diagram d associated with $\langle P,Q\rangle$ and $S \subseteq L(d)-(P\cup Q)$ then*

$$zr = \bigcup_{W\subseteq S} zr_W$$

where zr_W is the zonal region associated with $\langle P\cup W, Q\cup(S-W)\rangle$.

Taking $zr = \{z_1,z_2,z_3,z_4\}$ to be the zonal region associated with $\langle\{A\},\emptyset\rangle$ in figure 6, using the derived spit rule with $S = \{B,C\}$ gives

$$zr = \{z_1\} \cup \{z_2\} \cup \{z_3\} \cup \{z_4\}$$

since $\{z_1\}$, $\{z_2\}$, $\{z_3\}$ and $\{z_4\}$ are associated with $\langle\{A\}\cup\emptyset,\emptyset\cup\{B,C\}\rangle$, $\langle\{A\}\cup\{B\},\emptyset\cup\{C\}\rangle$, $\langle\{A\}\cup\{C\},\emptyset\cup\{B\}\rangle$ and $\langle\{A\}\cup\{B,C\},\emptyset\cup\emptyset\rangle$ respectively. In general, if we set $S = L(d) - (P \cup Q)$ in the lemma above and take $zr = \{z_1, z_2, ..., z_n\}$ we get $zr = \bigcup_{i=1}^{n}\{z_i\}$.

Note that the split associated with a zone involves all the labels in the diagram: $\{z\}$ is associated with $\langle P, Q\rangle$ where $P = L(z)$ and $Q = \overline{L(z)} = L(d) - L(z)$. Since any region is a set of zones, we can use this to define a function, ρ, from regions to splits.

Definition 3. *Let d be a unitary Venn diagram.*

(i) *If $z \in Z(d)$ then $\rho(\{z\}) = \left\langle L(z), \overline{L(z)}\right\rangle$*

(ii) *If $r = \{z_1, z_2, ..., z_n\} \in R(d)$ then $\rho(r) = \sum_{i=1}^{n} \rho(\{z_i\})$*

For example, in figure 6, $\rho(\{z_1\}) = \langle\{A\},\{B,C\}\rangle$. Under ρ the region $\{z_3, z_5, z_6\}$ maps to

$$\langle\{A,C\},\{B\}\rangle + \langle\{B\},\{A,C\}\rangle + \langle\{B,C\},\{A\}\rangle = \langle\{A,C\},\{B\}\rangle + \langle\{B\},\{A\}\rangle$$

Lemma 4. *Let zr be a zonal region of unitary Venn diagram d associated with $\langle P, Q\rangle$. Then $\rho(zr) = \langle P, Q\rangle$.*

The zonal region $\{z_1, z_2, z_3, z_4\}$ in figure 6 is associated with $\langle\{A\},\emptyset\rangle$ and

$$\begin{aligned}\rho(\{z_1, z_2, z_3, z_4\}) &= \langle\{A\},\{B,C\}\rangle + \langle\{A,B\},\{C\}\rangle + \\ &\quad \langle\{A,C\},\{B\}\rangle + \langle\{A,B,C\},\emptyset\rangle \\ &= \langle\{A\},\{C\}\rangle + \langle\{A,C\},\emptyset\rangle \\ &= \langle\{A\},\emptyset\rangle\end{aligned}$$

Lemma 4 does not follow over to Euler diagrams, as we shall see in section 4. If we know certain relationships between zonal regions, we can make deductions about their images under ρ, and vice versa.

Lemma 5. *Let zr_1 and zr_2 be zonal regions of unitary Venn diagram d. If $\rho(zr_1) = \langle P_1, Q_1\rangle$ and $\rho(zr_2) = \langle P_2, Q_2\rangle$ then*

$$zr_1 \subseteq zr_2 \Leftrightarrow P_2 \subseteq P_1 \wedge Q_2 \subseteq Q_1$$

If $(P_1 \cup P_2) \cap (Q_1 \cup Q_2) = \emptyset$ then $zr_1 \cap zr_2 = zr_3$ where

$$\rho(zr_3) = \langle P_1 \cup P_2, Q_1 \cup Q_2\rangle$$

From lemma 5 we can deduce that the zonal region associated with $\langle\{A\},\emptyset\rangle$ in diagram d, figure 6, is not a subset of the zonal region associated with $\langle\{B\},\emptyset\rangle$. The zonal region associated with $\langle\{A\},\emptyset\rangle$ is $\{z_1, z_2, z_3, z_4\}$. The zonal region associated with $\langle\{B\},\emptyset\rangle$ is $\{z_2, z_4, z_5, z_6\}$ and $\{z_1, z_2, z_3, z_4\} \not\subseteq \{z_2, z_4, z_5, z_6\}$.

Lemma 5 also tells us the zonal regions associated with $\langle\{A\},\emptyset\rangle$ and $\langle\{B\},\emptyset\rangle$ intersect to give a zonal region associated with $\langle\{A,B\},\emptyset\rangle$, that is $\{z_1,z_2,z_3,z_4\}\cap\{z_2,z_4,z_5,z_6\}=\{z_2,z_4\}$. We now define correspondence between zonal regions. Corresponding zonal regions have the same semantic interpretation.

Definition 4. *Let zr_1 and zr_2 be zonal regions of Venn diagrams d_1 and d_2 respectively. Regions zr_1 and zr_2 are **corresponding** zonal regions* [8], *denoted $zr_1 \equiv_c zr_2$, if and only if $\rho(zr_1)=\rho(zr_2)$.*

In figure 4 zonal region $\{z_1,z_2\}$ in d_1 corresponds to zonal region $\{z_5\}$ in d_2 since $\rho(\{z_1,z_2\})=\rho(\{z_5\})=\langle\{A\},\{B\}\rangle$.

Theorem 1. *The relation $\equiv_c$ is an equivalence relation on zonal regions.*

3.2 Corresponding Regions in Venn Diagrams

The definition of correspondence is now extended to regions.

Definition 5. *Let r_1 and r_2 be a regions of Venn diagrams d_1 and d_2 respectively. Regions r_1 and r_2 are **corresponding** regions* [8], *denoted $r_1 \equiv_c r_2$, if and only if $\rho(r_1)=\rho(r_2)$.*

At the semantic level, corresponding regions represent the same set [6]. In figure

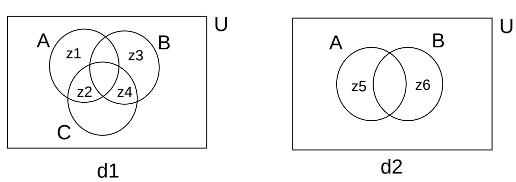

Fig. 7. Two Venn diagrams.

7, region $r_1=\{z_1,z_2,z_3,z_4\}$ in diagram d_1 has

$$\rho(r_1)=\langle\{A\},\{B,C\}\rangle+\langle\{A,C\},\{B\}\rangle+\langle\{B\},\{A,C\}\rangle+\langle\{B,C\},\{A\}\rangle$$

Region $r_2=\{z_5,z_6\}$ in diagram d_2 has

$$\rho(r_2)=\langle\{A\},\{B\}\rangle+\langle\{B\},\{A\}\rangle$$

Using axiom 2, $\langle P,Q\rangle=\langle P\cup\{A\},Q\rangle+\langle P,Q\cup\{A\}\rangle$, we obtain

$$\langle\{A\},\{B\}\rangle+\langle\{B\},\{A\}\rangle = \langle\{A\},\{B,C\}\rangle+\langle\{A,C\},\{B\}\rangle+\langle\{B\},\{A,C\}\rangle+\langle\{B,C\},\{A\}\rangle$$

Thus $r_1 \equiv_c r_2$.

Theorem 2. *The relation $\equiv_c$ is an equivalence relation on regions.*

Proofs for some of the results in this section can be found in [12]. Ideally, we want to be able to reason with diagrams that are Euler diagrams. The focus of this paper now turns to diagrams of this nature. Of the definitions related to Venn diagrams, 2 and 3 carry over to Euler diagrams. Also lemma 3 and corollary 1 apply to Euler diagrams.

4 Euler Diagrams

In this section we investigate problems related to zonal regions and their associated splits in Euler diagrams. It is no longer necessarily true that, for a zonal region zr associated with $\langle P, Q\rangle$, $\rho(zr) = \langle P, Q\rangle$ because the associated $\langle P, Q\rangle$ is no longer unique. In figure 8, the zonal region $\{z_1\}$ is associated with both $\langle\{A\}, \emptyset\rangle$ and $\langle\{A\}, \{B\}\rangle$ but $\rho(\{z_1\}) = \langle\{A\}, \{B\}\rangle \neq \langle\{A\}, \emptyset\rangle$. Thus lemma 4 fails. However, we can think of $\langle\{A\}, \emptyset\rangle$ and $\langle\{A\}, \{B\}\rangle$ as being 'equivalent in the context of d' because $\langle\{A\}, \emptyset\rangle = \langle\{A, B\}, \emptyset\rangle + \langle\{A\}, \{B\}\rangle$ and the zone corresponding to $\langle\{A, B\}, \emptyset\rangle$ is 'missing' from the diagram.

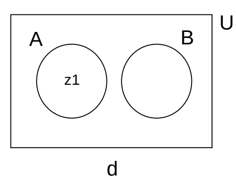

Fig. 8. An Euler diagram with a missing zonal region.

In some diagrams there may be a split on $L(d)$ with no zonal region associated with it. There is no zonal region associated with $\langle\{A, B\}, \emptyset\rangle$, in diagram d, in figure 8. Informally, in our 'algebra of splits' we can think of $\langle\{A, B\}, \emptyset\rangle$ as representing zero. If we allow this, we see that

$$\begin{aligned}\rho(\{z_1\}) &= \langle\{A\}, \{B\}\rangle \\ &= \langle\{A\}, \{B\}\rangle + \langle\{A, B\}, \emptyset\rangle \\ &= \langle\{A\}, \emptyset\rangle\end{aligned}$$

We have here the idea of equality *in the context of* a diagram.

Definition 6. *The **context** of unitary diagram d, denoted $\chi(d)$, is*

$$\chi(d) = \{\langle P, Q\rangle : P \in \mathbb{P}L(d) - Z(d) \wedge Q = L(d) - P\}$$

*If $\langle P, Q\rangle \in \chi(d)$ then $\langle P, Q\rangle$ is **zero in the context of** d, denoted $\langle P, Q\rangle =_d 0$.*

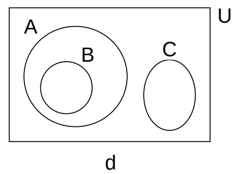

Fig. 9. A unitary Euler diagram.

The diagram in figure 9 has $Z(d) = \{\emptyset, \{A\}, \{A, B\}, \{C\}\}$, so

$$\chi(d) = \{\langle\{B\}, \{A, C\}\rangle, \langle\{A, C\}, \{B\}\rangle, \langle\{B, C\}, \{A\}\rangle, \langle\{A, B, C\}, \emptyset\rangle\}$$

corresponding to the four zones that are present in the Venn diagram with labels $\{A, B, C\}$ but are missing in d.

Lemma 6. *If d is a unitary Venn diagram $\chi(d) = \emptyset$.*

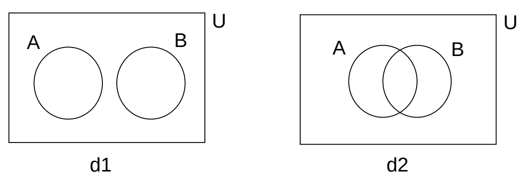

Fig. 10. Two Euler diagrams.

When considering more than one diagram, we need to take care when deciding what is the context. Considering diagrams d_1 and d_2, figure 10, in *conjunction* we may deduce that $\langle\{A, B\}, \emptyset\rangle$ is zero in context, since

$$\{z \in Z(d_1) : \{A, B\} \subseteq L(z) \wedge \emptyset \subseteq \overline{L(z)}\} = \emptyset$$

At the semantic level, the sets represented by the contours labelled A and B are disjoint. Thus we would want the zonal region associated with $\langle\{A\}, \emptyset\rangle$ in d_2 to correspond to that associated with $\langle\{A\}, \{B\}\rangle$, also in d_2. However if we were to take the diagrams in *disjunction*, we cannot deduce that the sets represented by the contours labelled A and B are disjoint. Thus we would not want $\langle\{A, B\}, \emptyset\rangle$ to be zero. In the disjunctive case it is incorrect for $\rho(\{z_1\}) = \langle\{A\}, \emptyset\rangle$.

In order to define the context of compound and multi-diagrams we first define a function ζ_δ, called *zonify*, from splits on $L(\delta)$ to sets of splits on $L(\delta)$, where δ is a unitary, compound or multi-diagram,

$$\zeta_\delta(\langle P, Q\rangle) = \{\langle P_i, Q_i\rangle : P \subseteq P_i \wedge Q_i = L(\delta) - P_i\}$$

The zonify function delivers the set of splits corresponding to the zones that are elements of the zonal region associated with $\langle P, Q\rangle$ in the Venn diagram with labels $L(\delta)$. Taking $\Delta = \{\{d_1, d_2\}, \{d_3\}\}$ in figure 11,

$$\zeta_\Delta(\langle\{A\}, \{B\}\rangle) = \{\langle\{A\}, \{B, C\}\rangle, \langle\{A, C\}, \{B\}\rangle\}$$

Consider the compound diagram $D = \{d_1, d_2\}$ in figure 12. The shaded zones

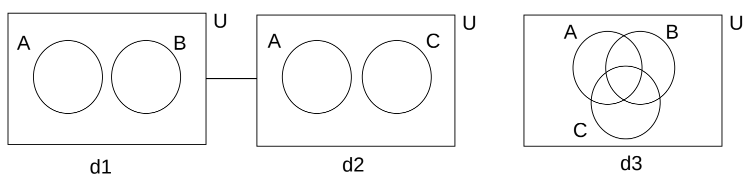

Fig. 11. A multi-diagram.

in the Venn diagram, d, with $L(d) = L(D)$, represent those sets we can deduce empty at the semantic level. Each of these shaded zones is associated with a split that partitions $L(D)$.

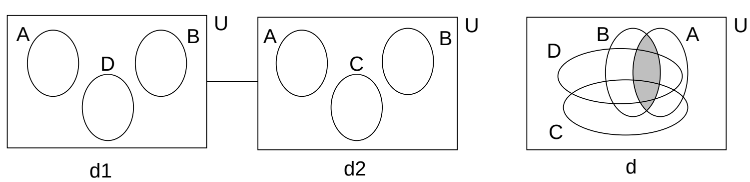

Fig. 12. A compound diagram and a Venn diagram.

Definition 7. *Let D be a compound diagram. The* ***context*** *of D is defined to be*

$$\chi(D) = \bigcap_{d \in D} \left(\bigcup_{\langle P,Q\rangle \in \chi(d)} \zeta_D(\langle P, Q\rangle) \right)$$

If $\langle P, Q\rangle \in \chi(D)$ then $\langle P, Q\rangle$ is ***zero in the context of*** *D, denoted $\langle P, Q\rangle =_D 0$.*

The context of $\{d_1, d_2\}$ in figure 13 is $\chi(\{d_1, d_2\}) = \emptyset$, since $\chi(d_2) = \emptyset$. The contexts of diagrams d_3 and d_4 are

$$\chi(d_3) = \{\langle\{A\}, \{B\}\rangle\}$$
$$\chi(d_4) = \{\langle\{A, C\}, \{B\}\rangle, \{\langle\{A\}, \{B, C\}\rangle\}$$

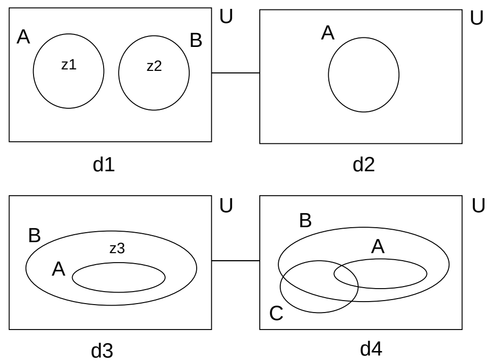

Fig. 13. A multi-diagram containing two compound diagrams.

We use the zonify function to find $\chi(\{d_3, d_4\})$.

$$\begin{aligned}\chi(\{d_3, d_4\}) &= \left(\bigcup_{\langle P,Q\rangle \in \chi(d_3)} \zeta_D(\langle P, Q\rangle)\right) \quad \cap \quad \left(\bigcup_{\langle P,Q\rangle \in \chi(d_4)} \zeta_D(\langle P, Q\rangle\right) \\ &= \{\langle\{A, C\}, \{B\}\rangle, \langle\{A\}, \{B, C\}\rangle\} \cap \\ &\qquad (\{\langle\{A, C\}, \{B\}\rangle\} \cup \{\langle\{A\}, \{B, C\}\rangle\}) \\ &= \{\langle\{A, C\}, \{B\}\rangle, \langle\{A\}, \{B, C\}\rangle\}\end{aligned}$$

Definition 8. *Let Δ be a multi-diagram. The* ***context*** *of Δ is defined to be*

$$\chi(\Delta) = \bigcup_{D \in \Delta} \left(\bigcup_{\langle P,Q\rangle \in \chi(D)} \zeta_\Delta(\langle P, Q\rangle)\right)$$

If $\langle P, Q\rangle \in \chi(\Delta)$ then $\langle P, Q\rangle$ is ***zero in the context of*** *Δ, denoted $\langle P, Q\rangle =_\Delta 0$.*

The context of $\Delta = \{\{d_1, d_2\}, \{d_3, d_4\}\}$ in figure 13 is

$$\chi(\Delta) = \{\langle\{A, C\}, \{B\}\rangle, \langle\{A\}, \{B, C\}\rangle\}$$

Therefore $\langle\{A, C\}, \{B\}\rangle =_\Delta 0$ and $\langle\{A\}, \{B, C\}\rangle =_\Delta 0$.

Definition 9. *Let Δ be a multi-diagram, $\sum_{i=1}^{n}\langle P_i, Q_i\rangle$ and $\sum_{j=1}^{m}\langle P'_j, Q'_j\rangle$ be sums of splits. $\sum_{i=1}^{n}\langle P_i, Q_i\rangle$ and $\sum_{j=1}^{m}\langle P'_j, Q'_j\rangle$ are said to be* ***equal in the context of*** *Δ, denoted $\sum_{i=1}^{n}\langle P_i, Q_i\rangle =_\Delta \sum_{j=1}^{m}\langle P'_j, Q'_j\rangle$, if and only if there exists $\sum_{i=1}^{k}\langle R_i, S_i\rangle$ and $\sum_{j=1}^{l}\langle R'_j, S'_j\rangle$ such that*

$\sum_{i=1}^{n}\langle P_i, Q_i\rangle = \sum_{i=1}^{k}\langle R_i, S_i\rangle$, $\sum_{j=1}^{m}\langle P'_j, Q'_j\rangle = \sum_{j=1}^{l}\langle R'_j, S'_j\rangle$ *and*
$\forall\, i\, (\exists\, j \bullet \langle R_i, S_i\rangle = \langle R'_j, S'_j\rangle) \vee \langle R_i, S_i\rangle =_\Delta 0$ *and*
$\forall\, j\, (\exists\, i \bullet \langle R'_j, S'_j\rangle = \langle R_i, S_i\rangle) \vee \langle R'_j, S'_j\rangle =_\Delta 0.$

In figure 13, taking $\Delta = \{\{d_1, d_2\}, \{d_3, d_4\}\}$ we have,

$$\begin{aligned}\langle\{A\},\{B\}\rangle + \langle\{B\},\{A\}\rangle &= \langle\{A,C\},\{B\}\rangle + \langle\{A\},\{B,C\}\rangle + \langle\{B\},\{A\}\rangle \\ &=_\Delta \langle\{B\},\{A\}\rangle\end{aligned}$$

Definition 10. *Let Δ be a multi-diagram and d_1, d_2 be unitary diagrams such that $d_1 \in D_1, d_2 \in D_2$ where $\{D_1, D_2\} \subseteq \Delta$. Let r_1 and r_2 be regions of d_1 and d_2 respectively. Region r_1 is said to* ***correspond in the context of*** *Δ to region r_2, denoted $r_1 \equiv_\Delta r_2$, if and only if $\rho(r_1) =_\Delta \rho(r_2)$.*

Corresponding regions have the same semantic interpretation. Consider regions $r_1 = \{z_1\}$, $r_2 = \{z_1, z_2\}$, $r_3 = \{z_2\}$ and $r_4 = \{z_3\}$ in figure 13.

$$\begin{aligned}\rho(r_1) &= \langle\{A\},\{B\}\rangle =_\Delta 0 \\ \rho(r_2) &= \langle\{A\},\{B\}\rangle + \langle\{B\},\{A\}\rangle \\ &=_\Delta \langle\{B\},\{A\}\rangle \\ \rho(r_3) &= \langle\{B\},\{A\}\rangle \\ \rho(r_4) &= \langle\{B\},\{A\}\rangle\end{aligned}$$

Thus $r_1 \not\equiv_\Delta r_2$ and $r_2 \equiv_\Delta r_4$. Interestingly, we also have $r_2 \equiv_\Delta r_3$ (r_2 and r_3 are different regions in the same diagram).

Theorem 3. *The relation $\equiv_\Delta$ is an equivalence relation on regions of unitary diagrams contained in Δ.*

5 The Counterpart Relations of Shin and Hammer

The basic idea of the counterpart relation on Venn diagrams is to identify corresponding basic regions (i.e., the region enclosed by a closed curve) and then to recursively define the relation on unions, intersections and complements of regions. Shin only defines the counterpart on basic regions and leaves the rest implicit. Hammer defines the relation as follows for Venn diagrams:

The counterpart relation is an equivalence relation defined as follows. Two basic regions are counterparts if and only if they are both regions enclosed by rectangles or else both regions enclosed by curves having the same label. If r and r' are regions of diagram D, s and s' are regions of diagram D', r is the counterpart of s, and r' is the counterpart of s', then $r \cup r'$ is the counterpart of $s \cup s'$ and $\bar{r}$ is the counterpart of $\bar{s}$.

This definition works very well for Venn diagrams where all minimal regions must occur. In figure 14, the two regions enclosed by the rectangles are counterparts, and so are the two crescent-shaped regions within the circles labelled A but outside the circles labelled B; the region within all three curves in the

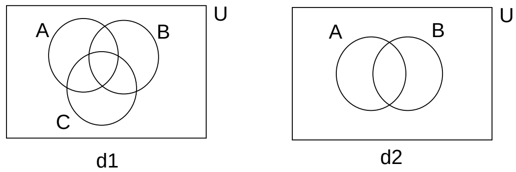

Fig. 14. Two Venn diagrams.

left-hand diagram has no counterpart in the other one. The counterpart relation is obviously equivalent to the correspondence relation defined in §3.

Hammer defines a counterpart relation on Euler diagrams, but only for diagrams with the same label set:

Suppose m and m' are minimal regions of two Euler diagrams D and D', respectively. Then m and m' are counterparts if and only if there are curves $B_1, \ldots, B_m$ and $B_{m+1}, \ldots, B_n$ of D and curves $B'_1, \ldots, B'_m$ and $B'_{m+1}, \ldots, B'_n$ of D' such that (1) for each i, $1 \leq i \leq n$, B_i and B'_i are tagged by the same label; (2) m is the minimal region within $B_1, \ldots, B_m$ but outside $B_{m+1}, \ldots, B_n$; and (3) m' is the minimal region within $B'_1, \ldots, B'_m$ but outside $B'_{m+1}, \ldots, B'_n$.

This definition is sufficient for Hammer's purposes, but it only covers a special case of Euler diagrams. Consider the two Euler diagrams in figure 15. Minimal region 1 is the counterpart of minimal region a, 2 is the counterpart of b and 3 is the counterpart of d. Minimal region c has no counterpart in the left-hand diagram.

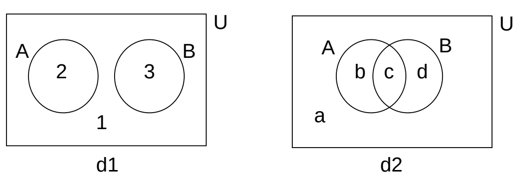

Fig. 15. Two Euler diagrams.

The correspondence relation defined in §4 agrees with this interpretation when the context is the disjunction of the two diagrams. It also agrees in the case in which the context is the conjunction of the two diagrams but adds in further correspondences such as region 2 corresponds with region $b \cup c$.

6 Conclusions and Further Work

We have constructed a purely syntactic definition of corresponding regions in Euler diagrams and shown it to be an equivalence relation and a generalization of the counterpart relations introduced by Shin and Hammer. At the semantic level, two corresponding regions represent the same set. The system of Euler diagrams we considered in this paper is in conjunctive normal form. However, we wish to reason in the more general case where we consider any combination of disjuncts and conjuncts of diagrams such as in constraint trees [7]. The correspondence relation defined in this paper can be adapted for such systems.

The general aim of this work is to provide the necessary mathematical underpinning for the development of software tools to aid reasoning with diagrams. In particular, we aim to develop the tools that will enable diagrammatic reasoning to become part of the software development process.

Acknowledgements Author Stapleton would like to thank the UK EPSRC for support under grant number 01800274. Authors Howse, Flower and Taylor were partially supported by UK EPSRC grant GR/R63516.

References

1. L. Euler. Lettres a une princesse d'allemagne, 1761.
2. J. Gil, J. Howse, and S. Kent. Towards a formalization of constraint diagrams. In *Proc Symp on Human-Centric Computing*. IEEE Press, Sept 2001.
3. E. Hammer. *Logic and Visual Information*. CSLI Publications, 1995.
4. E. Hammer and S-J Shin. Euler's visual logic. In *History and Philosophy of Logic*, pages 1–29, 1998.
5. D. Harel. On visual formalisms. In J. Glasgow, N. H. Narayan, and B. Chandrasekaran, editors, *Diagrammatic Reasoning*, pages 235–271. MIT Press, 1998.
6. J. Howse, F. Molina, and J. Taylor. On the completness and expressiveness of spider diagram systems. In *Proceedings of Diagrams 2000*, pages 26–41. Springer-Verlag, 2000.
7. S. Kent and J. Howse. Constraint trees. In A. Clark and J. Warner, editors, *Advances in object modelling with ocl*. Spinger Verlag, to appear, 2002.
8. F. Molina. *Reasoning with extended Venn-Peirce diagrammatic systems*. PhD thesis, University of Brighton, 2001.
9. OMG. UML specification, version 1.3. Available from www.omg.org.
10. C. Peirce. *Collected Papers*, volume Vol. 4. Harvard Univ. Press, 1933.
11. S.-J. Shin. *The Logical Status of Diagrams*. Cambridge University Press, 1994.
12. G. Stapleton. Comparing regions in spider diagrams. Available at www.it.brighton.ac.uk/research/vmg/papers.html.
13. J. Venn. On the diagrammatic and mechanical representation of propositions and reasonings. *Phil.Mag*, 1880.