SoSym manuscript No.
(will be inserted by the editor)

Nesting in Euler Diagrams: syntax, semantics and construction

Jean Flower, John Howse, John Taylor

Visual Modelling Group,

School of Computing, Mathematical and Information Sciences,

University of Brighton, Brighton, U.K.

www.cmis.bton.ac.uk/research/vmg

Received: date / Revised version: date

Abstract This paper considers the notion of *nesting* in Euler diagrams, and how nesting affects the interpretation and construction of such diagrams. After setting up the necessary definitions for concrete Euler diagrams (drawn in the plane) and abstract diagrams (having just formal structure), the notion of nestedness is defined at both concrete and abstract levels. The concept of a dual graph is used to give an alternative condition for a drawable abstract Euler diagram to be nested. The notions of nesting at the two levels are shown to be equivalent for drawable abstract diagrams under the morphism from concrete to abstract diagrams. The natural progression to the diagram semantics is explored and we present a "nested form" for

diagram semantics. We describe how this work supports tool-building for diagrams, and how effective we might expect this support to be in terms of the proportion of nested diagrams.

1 Introduction

Euler diagrams [3] 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. Euler diagrams form the basis of more expressive diagrammatic notations such as Higraphs [7], constraint diagrams [6] and some of the notations of UML [10]. A *concrete* Euler diagram is one which is drawn in the plane. We can abstract away from any irrelevant geometric and topological information to produce an *abstract* diagram which has only formal structure. The distinction between concrete diagrams and abstract diagrams was highlighted in [8]. The problem of converting an abstract Euler diagram into a concrete representative, necessary for the development of software tools using notations based on Euler diagrams, was addressed in [4]. This paper extends work on Euler diagrams by incorporating the notion of a *nested* diagram. Informally, an Euler diagram is nested if the (concrete) diagram contains disjoint components (other choices of name for this concept could have been *disconnected* or *separated*). Section 2 begins with the necessary background notation and definitions for Euler diagrams. Depending on the well-formedness conditions of concrete Euler diagrams, some abstract diagrams are not drawable.

The concept of nesting is most obvious, visually, for concrete diagrams. However, the notion extends to abstract diagrams. In section 3 we define the notion of nesting in concrete and abstract Euler diagrams. We use the powerful notion of a dual graph to give an alternative condition for a drawable abstract Euler diagram to be nested. The two notions of nesting are shown to be equivalent for drawable abstract diagrams under the morphism from concrete to abstract diagrams (theorems 3, 4). Nesting in diagrams gives rise to different ways of presenting the semantics of diagrams, and in section 4 we establish a "nested form" for diagram semantics.

One application of this work is in diagram generation algorithms which are used to drive software tools, particularly those involved with software system modelling, for example [9]. This application of the nesting concept is discussed in section 5. Finally, in section 6, some data are presented to show how much leverage can be gained from making use of nesting in abstract diagrams.

2 The context: Euler diagrams

Work in this section is largely based upon work from [4]. An *abstract Euler diagram* comprises a set whose elements are called *contours* and a set of *zones* which are subsets of the contour set. An abstract Euler diagram encapsulates some of the information conveyed in a diagram, encompassing enough information to create a semantic interpretation of the diagram but discarding some topological details.

Definition 1 *An* ***abstract Euler diagram*** *is a pair:* $d = \langle \mathcal{C}(d), \mathcal{Z}(d) \rangle$ *where*

(i) $\mathcal{C}(d)$ *is a finite set whose members are called* ***contours***

(ii) $\mathcal{Z}(d) \subseteq \mathcal{P}\mathcal{C}(d)$ *is the set of* ***zones*** *of* d*, so* $z \in \mathcal{Z}(d)$ *is* $z \subseteq \mathcal{C}(d)$

(iii) $\varnothing \in \mathcal{Z}(d)$

(iv) $\bigcup_{z \in \mathcal{Z}(d)} z = \mathcal{C}(d)$

The set of abstract diagrams is denoted $\mathcal{D}$*.*

Example 1 The abstract diagram $\langle \{a, b, c\}, \{\{\}, \{a\}, \{a, b\}, \{b\}, \{c\}\} \rangle \in \mathcal{D}$ has three contours and five zones.

Abstract diagrams can easily be given semantic interpretations (see section 4) and each abstract diagram may have many concrete representations, with differing topological details.

A *concrete Euler diagram* is a set of labelled *contours* (simple closed curves) in the plane, each with a unique label. A *zone* is a connected component of the complement of the contour set.

Definition 2 *A* ***concrete Euler diagram*** *is a triple* $\hat{d} = \langle \widehat{\mathcal{L}}(\hat{d}), \widehat{\mathcal{C}}(\hat{d}), \widehat{\mathcal{Z}}(\hat{d}) \rangle$ *such that:*

(i) $\widehat{\mathcal{C}}(\hat{d})$ *is a finite set of simple closed curves,* ***contours****, in the plane* $\mathbb{R}^2$*. Each contour has a unique label from the set* $\widehat{\mathcal{L}}(\hat{d})$*, so that the labelling mapping* $\widehat{\mathcal{C}}(\hat{d}) \to \widehat{\mathcal{L}}(\hat{d})$ *is a bijection.*

(ii) contours meet transversely.

(iii) each component $\hat{z} \in \mathbb{R}^2 - \bigcup_{\hat{c} \in \widehat{\mathcal{C}}(\hat{d})} \hat{c}$ *is a* ***zone****. These make up the set* $\widehat{\mathcal{Z}}(\hat{d})$.

(iv) each zone $\hat{z}$ *is uniquely identified by a set of contours* $\widehat{\mathcal{C}}(\hat{z}) \subset \widehat{\mathcal{C}}(\hat{d})$ *with* $\hat{z} = \bigcap_{\hat{c} \in \widehat{\mathcal{C}}(\hat{z})} \mathit{interior}\,(\hat{c}) \cap \bigcap_{\hat{c} \in \widehat{\mathcal{C}}(\hat{d}) - \widehat{\mathcal{C}}(\hat{z})} \mathit{exterior}\,(\hat{c})$*. The mapping from contours to labels gives unique label sets for each zone,* $\widehat{\mathcal{C}}(\hat{z})$ *mapping to* $\widehat{\mathcal{L}}(\hat{z})$*. This defines* $\widehat{\mathcal{L}}(\hat{z})$.

The set of concrete diagrams is denoted $\widehat{\mathcal{D}}$.

Part (ii) of this definition prohibits examples with concurrent contours or tangential contours. This restriction is justifiable from the point of view of the usability of the diagrams as a representation of information, or as a reasoning tool. The decision could be seen as arbitrary. However the condition that contours cross transversely (which prohibits tangential contours and concurrent contours) is necessary to ensure the consistent definition of nestedness at the concrete and abstract levels. The use of tangential contours would be counter to our intuition of nestedness at the concrete level.

Figure 1 shows examples of diagrams which include concurrent or tangential contours, and are thus not examples of concrete Euler diagrams.

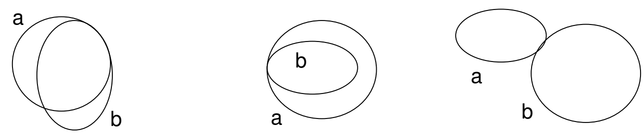

Fig. 1 Diagrams which include concurrent or tangential contours

Parts (iii) and (iv) of definition 2 combine to prohibit diagrams with *disconnected zones*, such as those illustrated in figure 2.

Fig. 2 Diagrams with disconnected zones

Diagrams that satisfy definition 2 are well-formed Euler diagrams. They exclude diagrams with concurrent or tangential contours and those with disconnected zones. Figure 3 shows examples of (well-formed) concrete Euler diagrams.

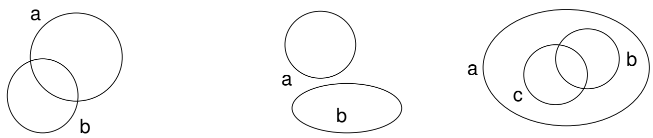

Fig. 3 Well-formed concrete Euler diagrams

Example 2 Let $\hat{d}$ be the first concrete diagram given in figure 3. $\widehat{\mathcal{C}}(\hat{d})$ has two elements (the two contours shown), $\widehat{\mathcal{L}}(\hat{d}) = \{a, b\}$ and $\widehat{\mathcal{Z}}(\hat{d})$ has four elements, uniquely determined by the label sets $\{\}$, $\{a\}$, $\{b\}$ and $\{a, b\}$.

For each concrete Euler diagram, there is an abstract Euler diagram which has the same structural information as the concrete diagram but loses some geometric and topological details. The mapping from a concrete diagram to its abstract diagram is defined next.

Definition 3 *The mapping* $ab : \widehat{\mathcal{D}} \to \mathcal{D}$ *("ab" for "abstraction") forgets the positioning of the contours. It is defined by*

$$ab : \langle \widehat{\mathcal{L}}(\hat{d}), \widehat{\mathcal{C}}(\hat{d}), \widehat{\mathcal{Z}}(\hat{d}) \rangle \mapsto \langle \widehat{\mathcal{L}}(\hat{d}), \{\widehat{\mathcal{L}}(\hat{z}) : \hat{z} \in \widehat{\mathcal{Z}}(\hat{d})\} \rangle$$

When mapping from concrete diagrams to abstract diagrams, identify contours and their labels. This identification reduces a three-tuple for a concrete diagram to a pair defining the abstract diagram.

Example 3 Let $\hat{d}$ be the concrete diagram given in figure 4. Its abstract diagram has:

$$\mathcal{C}(ab(\hat{d})) = \{a, b, c, d\}$$

$$\mathcal{Z}(ab(\hat{d})) = \{\{\}, \{a\}, \{b\}, \{a, b\}, \{a, c\}, \{a, d\}, \{a, c, d\}\}$$

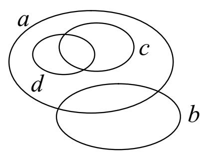

Fig. 4 A concrete diagram

Definition 4 *A concrete diagram $\hat{d}$* ***represents*** *or* ***complies with*** *an abstract diagram d if and only if $d = ab(\hat{d})$. An abstract diagram which has a compliant well-formed concrete representation is* ***drawable****.*

An abstract Euler diagram is either undrawable, or it has many concrete representations. One example of an undrawable abstract diagram has two contours and two zones: $\mathcal{C}(d) = \{a, b\}$ and $\mathcal{Z}(d) = \{\{\}, \{a, b\}\}$. A concrete representation of this must have the two contours running concurrently as the shared boundary between the two zones, giving a concrete diagram which is not well-formed. Adding one more abstract zone to d gives an abstract diagram which has many concrete representations: for example, the diagram $d = \langle\{a, b\}, \{\{\}, \{a\}, \{a, b\}\}\rangle$ has concrete representations illustrated in figure 5.

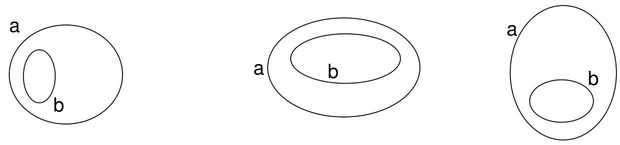

Fig. 5 Different concrete representations of the same abstract Euler diagram

A powerful idea for the analysis of concrete and abstract Euler diagrams is that of the *dual graph*. The dual graph of an abstract Euler diagram is a particular kind of abstract labelled graph, defined next.

Definition 5 *An* ***abstract labelled graph*** *is a triple* $\langle \mathcal{L}(G), \mathcal{V}(G), \mathcal{E}(G)\rangle$ *where the components are defined as follows.*

(i) $\mathcal{L}(G)$ *is a set of labels.*

(ii) $\mathcal{V}(G)$*is a set of vertices. Each vertex* v *is labelled with* $\mathcal{L}(v) \subseteq \mathcal{L}(G)$.

(iii) $\mathcal{E}(G)$ *is a set of edges. Each edge is a pair of vertices in* $\mathcal{V}(G)$ *where the vertex labels must have a singleton symmetric difference; that is, one set exceeds the other by a single additional element. This element can be used to label the edge.*

The set of abstract labelled graphs is denoted $\mathcal{LG}$.

Definition 6 *The map dual* : $\mathcal{D} \to \mathcal{LG}$ *is defined by*

$$\langle \mathcal{C}(d), \mathcal{Z}(d)\rangle \mapsto \langle \mathcal{C}(d), \mathcal{Z}(d), \mathcal{E}(G)\rangle$$

where the edges include all possible $e = (v_1, v_2)$ *where* v_1 *and* v_2 *have singleton symmetric difference.*

The dual graph of an abstract diagram is illustrated in figure 6.

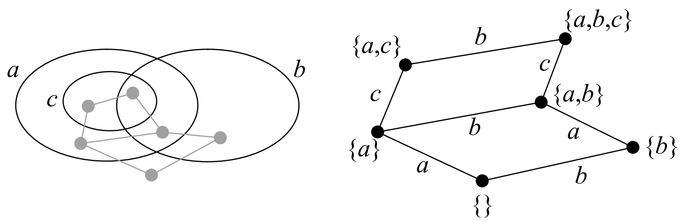

Fig. 6 A diagram and its dual graph

This map assigns an abstract labelled graph to each abstract Euler diagram. The diagram with $\mathcal{C}(d) = \{a, b\}$ and $\mathcal{Z}(d) = \{\{\}, \{a\}, \{a, b\}\}$ maps to a graph with two labels (one for each abstract contour), three vertices (one for each abstract zone) and two edges:

$$\mathcal{L}(dual(d)) = \{a, b\},$$

$$\mathcal{V}(dual(d)) = \{\{\}, \{a\}, \{a, b\}\},$$

$$\mathcal{E}(dual(d)) = \{\{\{\}, \{a\}\}, \{\{a\}, \{a, b\}\}\}.$$

This definition of the dual graph of an abstract diagram extends to a definition of the dual graph of a concrete diagram.

Definition 7 *The map dual* : $\widehat{\mathcal{D}} \to \mathcal{LG}$ *is defined by*

$$\langle \widehat{\mathcal{L}}(\hat{d}), \widehat{\mathcal{C}}(\hat{d}), \widehat{\mathcal{Z}}(\hat{d})\rangle \mapsto dual(ab(\langle \widehat{\mathcal{L}}(\hat{d}), \widehat{\mathcal{C}}(\hat{d}), \widehat{\mathcal{Z}}(\hat{d})\rangle)).$$

Note that this dual graph is not a topological construction. It is possible for two zones which are not topologically adjacent in $\hat{d}$ to correspond to adjacent vertices in the dual (the simplest example is any representation of a Venn diagram with four contours: see figure 7). However, if two zones are adjacent in $\hat{d}$ then the vertices are necessarily adjacent in the dual.

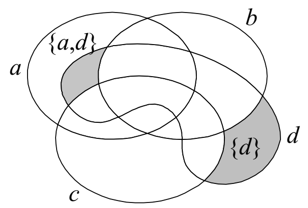

Fig. 7 The topological dual graph of the Venn diagram on four contours has fewer edges than the abstract dual graph

Definition 8 (The connectivity conditions) *An abstract labelled graph* $\langle \mathcal{L}(G), \mathcal{V}(G), \mathcal{E}(G) \rangle$ *satisfies the* ***connectivity conditions*** *if it is connected and, for all labels* $l \in \mathcal{L}(G)$*, the subgraphs* $G^{+}(l)$ *generated by vertices whose labels include* l*, and* $G^{-}(l)$ *generated by vertices whose labels exclude* l *are connected.*

Theorem 1 (The connectivity theorem) *Let* $\hat{d}$ *be a concrete diagram. Then* $dual(\hat{d})$ *satisfies the connectivity conditions. Hence, if an abstract diagram* d *is drawable then* $dual(d)$ *satisfies the connectivity conditions.*

This theorem is proved in [4] and can be used to prove that the abstract diagram $d = \langle \{a, b\}, \{\{\}, \{a, b\}\} \rangle$ is indeed undrawable. Its dual graph $dual(d)$ has two labels $\{a, b\}$, two vertices labelled $\{\}$ and $\{a, b\}$ and no edges. The dual is disconnected, so the diagram is not drawable.

This completes the framework we need to discuss the main points of this paper. The next section focusses on the concepts of nested and atomic diagrams.

3 Defining atomic and nested diagrams

In this section we identify nesting within a diagram. Initially we will define nesting as a property of concrete diagrams and later extend the concept to abstract diagrams.

3.1 Nesting in concrete diagrams

Definition 9 *A concrete diagram $\hat{d}$ is* ***nested*** *if there exist $n \geq 2$ subdiagrams $\hat{d}_1, ..., \hat{d}_n$ such that*

1. *for each i, the contour set of $\hat{d}_i$ is non-empty*
2. *for each $i \neq j$ no contour in $\widehat{\mathcal{C}}(\hat{d}_i)$ meets any contour in $\widehat{\mathcal{C}}(\hat{d}_j)$*

A diagram which is not nested is called ***atomic****.*

Proposition 1 *Let $\hat{d}$ be a concrete Euler diagram. The following four conditions are equivalent:*

1. *$\hat{d}$ is nested;*
2. *the union of the contours of $\hat{d}$ is a disconnected subset of the plane;*
3. *there exists a simple closed curve γ which does not meet any of the contours of $\hat{d}$, and splits the plane into two parts, each of which includes at least one contour of $\hat{d}$;*
4. *there exist subdiagrams $\hat{d}_1$ and $\hat{d}_2$ of $\hat{d}$, each of which has a non-empty contour set, and there is a zone $\hat{z} \in \widehat{\mathcal{Z}}(\hat{d}_1)$ such that all contours in $\widehat{\mathcal{C}}(\hat{d}_2)$ are contained within $\hat{z}$.*

Figure 8 illustrates these equivalent approaches to nesting in concrete Euler diagrams. Two examples are shown from three different points of view.

The first point of view partitions the contours (shown here as solid contours and dashed contours). This partition supports the definition of nesting (partitioning into sub-diagrams) and the notion that the set comprising the contour points is a disconnected set (one component is solid, the other is dashed).

The second point of view adds a path which is not a contour of the diagram, the path γ in proposition 1(3).

The final point of view shows two diagrams, $\hat{d}_1$ shown as solid lines and $\hat{d}_2$ shown as dashed lines. One zone of $\hat{d}_1$ is shaded, and that zone contains all contours in $\hat{d}_2$. This illustrates proposition 1 (4).

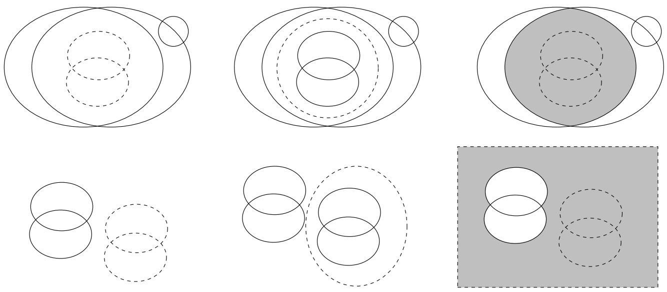

Fig. 8 Criteria for nesting in a Euler diagrams

3.2 Nesting in abstract diagrams

The notions of crossing contours, topological connectedness or topological containment are unavailable to us when we define the notion of nesting in the abstract case. However, we can identify the abstract diagram equivalent of condition 4 of proposition 1.

Definition 10 *An abstract diagram d is* ***nested*** *if there exist abstract diagrams $d_1 = \langle \mathcal{C}_1, \mathcal{Z}_1 \rangle$ and $d_2 = \langle \mathcal{C}_2, \mathcal{Z}_2 \rangle$ and a zone $z^\star \in \mathcal{Z}(d_1)$ such that*

(i) $\{\mathcal{C}_1, \mathcal{C}_2\}$ *is a partition of* $\mathcal{C}(d)$

(ii) $\mathcal{Z}(d) = \mathcal{Z}_1 \cup \mathcal{Z}_2^\star$ *and* $\mathcal{Z}_1 \cap \mathcal{Z}_2^\star = \{z^\star\}$ *where* $\mathcal{Z}_2^\star = \{z^\star \cup z_2 : z_2 \in \mathcal{Z}_2\}$.

We say that d_2 is ***embedded*** *in the zone $z^\star$ of d_1, and write $d = d_2 \xrightarrow{z^\star} d_1$.*

Example 4 The diagram d whose concrete representation is given (on the left) in figure 9 has $\mathcal{C}(d) = \{a, b, c, e\}$ and $\mathcal{Z}(d) = \{\{\}, \{a\}, \{a, b\}, \{b\}, \{a, c\}, \{a, c, e\}, \{a, e\}\}$.

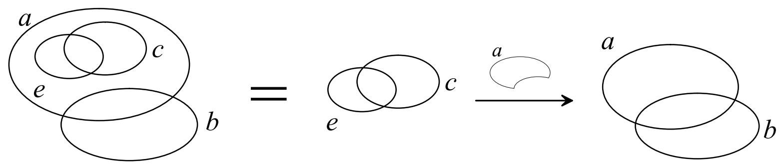

Fig. 9 Embedding d_2 in d_1

Here $d_2 = \langle \{c, e\}, \{\{\}, \{c\}, \{c, e\}, \{e\}\} \rangle$ is embedded in the zone $\{a\}$ of $d_1 = \langle \{a, b\}, \{\{\}, \{a\}, \{a, b\}, \{b\}\} \rangle$. The contour set $\mathcal{C}(d)$ partitions as $\mathcal{C}_1 \cup \mathcal{C}_2 = \{a, b\} \cup \{c, e\}$ and the zone set $\mathcal{Z}(d) = \mathcal{Z}_1 \cup \mathcal{Z}_2^\star$ where $\mathcal{Z}_1 = \{\{\}, \{a\}, \{a, b\}, \{b\}\}$, $\mathcal{Z}_2^\star = \{\{a\}, \{a, c\}, \{a, c, e\}, \{a, e\}\}$ and $\mathcal{Z}_1 \cap \mathcal{Z}_2^\star = \{a\}$.

Example 5 Although the definition of nesting in abstract Euler diagrams is motivated by the concrete nesting criterion given in proposition 1 (4), it is still meaningful for undrawable diagrams.

Let $d = \langle\{a, b, c, d\}, \{\{\}, \{a\}, \{b\}, \{a, b\}, \{a, b, c, e\}\}\rangle$. Then $\mathcal{C}(d) = \{a, b\} \cup \{c, e\}$ and $\mathcal{Z}(d) = \{\{\}, \{a\}, \{b\}, \{a, b\}\} \cup \{\{a, b\}, \{a, b, c, e\}\}$ where $\{\{\}, \{a\}, \{b\}, \{a, b\}\} \cap \{\{a, b\}, \{a, b, c, e\}\} = \{a, b\}$. Hence $d_2 = \langle\{c, e\}, \{\{\}, \{c, e\}\}\rangle$ is embedded in zone $\{a, b\}$ of $d_1 = \langle\{a, b\}, \{\{\}, \{a\}, \{b\}, \{a, b\}\}\rangle$. In this example, d_1 is drawable but d and d_2 are undrawable. Figure 10 attempts to represent d as a (non-well-formed) concrete diagram which has coincident contours labelled c and d (violating condition 2 of definition 2).

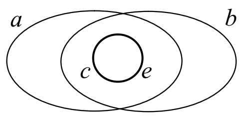

Fig. 10 An undrawable nested diagram

Example 6 The diagram d whose concrete representation is given in figure 11 is $\langle\mathcal{C}(d), \mathcal{Z}(d)\rangle = \langle\{a, b, c, e\}, \{\{\}, \{a\}, \{a, b\}, \{c\}, \{c, e\}, \{e\}\}\rangle$. Let $d_1 = \langle\{a, b\}, \{\{\}, \{a\}, \{a, b\}\}\rangle$ and $d_2 = \langle\{c, e\}, \{\{\}, \{c\}, \{c, e\}, \{e\}\}\rangle$. Then d is obtained by embedding d_2 in the zone $z^\star = \{\}$ of d_1. Since the embedding zone $z^\star$ is the empty set of contours, we could also regard d as being obtained by embedding d_1 in d_2.

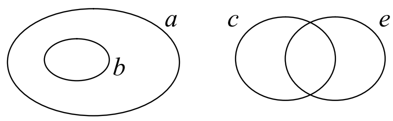

Fig. 11 Independently embedded diagrams

When the embedding zone is $z^\star = \varnothing$, we say that diagrams d_1 and d_2 are **independently embedded**. (The motivation for this terminology comes from concrete diagrams: if $\hat{d}$ is a concrete representation of d, then there exist disjoint, simply-connected regions of the plane $\mathbb{R}^2$ containing $\hat{d}_1$ and $\hat{d}_2$ respectively.)

In this example, since d_1 is itself a nested diagram, we may also describe d in several additional ways, as follows:

(i) $\left(d_{1,1} \xrightarrow{\{a\}} d_{1,2}\right) \xrightarrow{\{\}} d_2$ where $d_{1,1} = \langle\{b\},\{\{\},\{b\}\}\rangle$ and $d_{1,2} = \langle\{a\},\{\{\},\{a\}\}\rangle$,

(ii) $d_{1,1} \xrightarrow{\{a\}} \left(d_{1,2} \xrightarrow{\{\}} d_2\right)$ (with $d_{1,1}$ and $d_{1,2}$ as in (i)),

(iii) $d_{1,1} \xrightarrow{\{a\}} d_3$ where $d_3 = \langle\{a,c,e\},\{\{\},\{a\},\{c\},\{c,e\},\{e\}\}\rangle$.

Motivated by example 6, the following proposition identifies the elementary properties of the 'embedding relation' $\xrightarrow{\square}$.

Proposition 2 *The embedding relation satisfies the following.*

(i) $d_1 \xrightarrow{\varnothing} d_2 \quad = \quad d_2 \xrightarrow{\varnothing} d_1$

(ii) $\left(d_1 \xrightarrow{z_a} d_2\right) \xrightarrow{z_b} d_3 \quad = \quad d_1 \xrightarrow{z_a \cup z_b} \left(d_2 \xrightarrow{z_b} d_3\right)$

□

An abstract *drawable* diagram is nested if and only if its dual graph has a cut vertex. In order to prove this result we need to develop some terminology and prove a lemma. Let d be an abstract drawable diagram. The dual of d has $n > 1$ subgraphs $S_1, ..., S_n$, called **cut components**, obtained by removing the cut vertex and replacing it back into each component in turn: see figure 12. For each $1 \leq i \leq n$, let $C_i \subseteq \mathcal{C}(d)$ be the set of contour labels appearing as edge labels of cut component S_i.

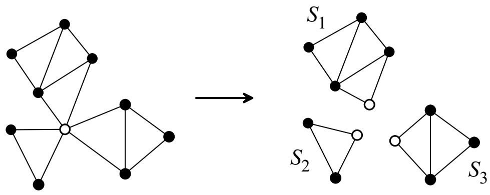

Fig. 12 The cut components of a graph with cut vertex

Lemma 1 *Let d be an abstract drawable diagram whose dual graph dual(d) contains a cut vertex. Then the sets $C_1, ..., C_n$ of edge labels of the cut components of dual(d) partition the contour set $\mathcal{C}(d)$*

Proof Every contour in $\mathcal{C}(d)$ appears in one of the sets $C_1, ..., C_n$. We will show that the sets $C_1, ..., C_n$ are disjoint, by contradiction.

Let $c \in C_i \cap C_j$ where $i \neq j$. There are edges of e_i in S_i, e_j in S_j which are both labelled c. Let e_i have ends v_i, w_i and e_j have ends v_j, w_j where $c \in v_i, v_j$ and $c \notin w_i, w_j$.

Let $v^\star$ be a cut vertex of *dual*(d). Assume first that $c \notin v^\star$ so that $v_i \in S_i - v^\star$, $v_j \in S_j - v^\star$. Any path from v_i to v_j must pass through the cut vertex $v^\star$, but the drawability of d tells us that the dual satisfies the connectivity conditions, including the fact that the subgraph restricted to those vertices which contain c is connected. This is a contradiction.

If, on the other hand, $c \in v^\star$ then follow a similar line of argument using w_i and w_j, and the contradiction comes from the connectivity condition that the subgraph of S built from vertices which exclude c is connected.

Thus $C_1, ..., C_n$ partition the contour set. $\square$

Theorem 2 *Let d be an abstract drawable diagram. Then d is nested if and only if the dual graph $dual(d)$ contains a cut vertex.*

Proof Let d be an abstract nested drawable diagram. Then there exist abstract diagrams $d_1 = \langle \mathcal{C}_1, \mathcal{Z}_1 \rangle$ and $d_2 = \langle \mathcal{C}_2, \mathcal{Z}_2 \rangle$ and a zone $z^\star \in \mathcal{Z}(d_1)$ such that $\{\mathcal{C}_1, \mathcal{C}_2\}$ is a partition of $\mathcal{C}(d)$ and $\mathcal{Z}(d) = \mathcal{Z}_1 \cup \mathcal{Z}_2^\star$ and $\mathcal{Z}(d) = \mathcal{Z}_1 \cap \mathcal{Z}_2^\star = \{z^\star\}$, where $\mathcal{Z}_2^\star = \{z_2 \cup z^\star : z_2 \in \mathcal{Z}_2\}$. Then $\mathcal{Z}(d)$ is partitioned by $\mathcal{Z}(d) = (\mathcal{Z}_1 - \{z^\star\}) \cup \{z^\star\} \cup (\mathcal{Z}_2^\star - \{z^\star\})$.

Let $\triangle$ denote set symmetric difference. Then

$$\begin{aligned}
z \in \mathcal{Z}_1 - \{z^\star\} \quad &\Rightarrow \quad z \triangle z^\star \subseteq \mathcal{C}(d_1) \wedge z \triangle z^\star \neq \varnothing \\
z \in \mathcal{Z}_2^\star - \{z^\star\} \quad &\Rightarrow \quad z \triangle z^\star \subseteq \mathcal{C}(d_2) \wedge z \triangle z^\star \neq \varnothing \\
z_1 \in \mathcal{Z}_1 - \{z^\star\} \wedge z_2 \in \mathcal{Z}_2^\star - \{z^\star\} \quad &\Rightarrow \quad z_1 \triangle z_2 = (z_1 \triangle z^\star) \cup (z_2 \triangle z^\star)\,.
\end{aligned}$$

The symmetric difference of abstract zones in sets $\mathcal{Z}_1 - \{z^\star\}$ and $\mathcal{Z}_2^\star - \{z^\star\}$ contains at least two elements, so no two are adjacent. The zone $z^\star$ acts as a pathway in the dual graph from $\mathcal{Z}_1 - \{z^\star\}$ to $\mathcal{Z}_2^\star - \{z^\star\}$, and is a cut vertex of the dual graph.

Conversely, let $z^\star$ be a cut vertex in $dual(d)$. The dual has $n > 1$ cut components $S_1, ..., S_n$. By construction of the S_i we have, for distinct i and j, $\mathcal{V}(S_i) \cap \mathcal{V}(S_j) = \{z^\star\}$.

For each $1 \leq i \leq n$, let $C_i \subseteq \mathcal{C}(d)$ be the set of contour labels appearing as edge labels of S_i. By lemma 1, the C_i partition $\mathcal{C}(d)$. Without loss of generality, S_1 contains a vertex labelled by the empty set (possibly other cut components do too, if the cutvertex is the null vertex). Then choose

$d_1 = \langle C_1, \mathcal{V}(S_1)\rangle$ and $d_2 = \langle \bigcup_{i\neq 1} C_i, \{v - z^\star : v \in \bigcup_{i\neq 1} \mathcal{V}(S_i)\}\rangle$. Then $\{\mathcal{C}(d_1), \mathcal{C}(d_2)\}$ is a partition of $\mathcal{C}(d)$ and $\mathcal{Z}(d) = \mathcal{Z}_1 \cup \mathcal{Z}_2^\star$ and $\mathcal{Z}_1 \cap \mathcal{Z}_2^\star = \{z^\star\}$, where $\mathcal{Z}_2^\star = \bigcup_{i\neq 1} \mathcal{V}(S_i) = \{z_2 \cup z^\star : z_2 \in \mathcal{Z}_2\}$. Hence d is nested. $\square$

Example 7 Let $d_1 = \langle \mathcal{C}(d) = \{a, b\}, \mathcal{Z}(d) = \{\{\}, \{a\}, \{a, b\}\}\rangle$. Then d_1 is nested by embedding $\langle\{b\}, \{\{\}, \{b\}\}\rangle$ in the zone $\{a\}$ of $\langle\{a\}, \{\{\}, \{a\}\}\rangle$. The dual graph of d_1, which has cut vertex labelled $\{a\}$ is given in figure 13. A more interesting example can be seen by building the dual graph of the second example in figure 8. Let $d_2 = \langle\{a, b, c, d\}, \{\{\}, \{a\}, \{b\}, \{a, b\}, \{c\}, \{d\}, \{c, d\}\}\rangle$. In this case the diagrams $\langle\{a, b\}, \{\{\}, \{a\}, \{b\}, \{a, b\}\}\rangle$ and $\langle\{a, b\}, \{\{\}, \{a\}, \{b\}, \{a, b\}\}\rangle$ are independently embedded in d_2. In the dual graph, the cutvertex is labelled $\{\}$ – see figure 13.

The addition of a single abstract zone $\{a, c\}$ turns the nested diagram d_2 into an atomic diagram d_3 – the third dual graph in figure 13 has no cut vertex.

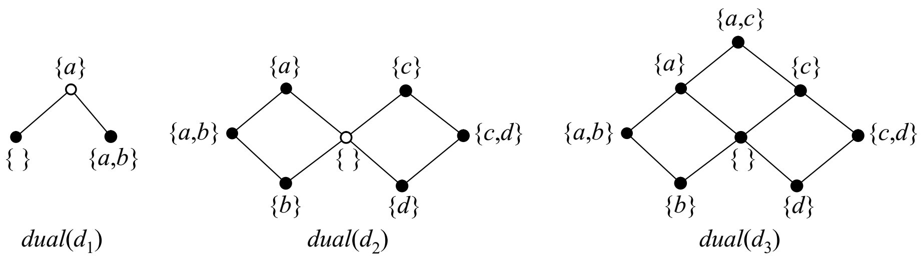

Fig. 13 Cut vertices in dual graphs

Example 8 The condition in theorem 2 that the diagram d is *drawable* is necessary. Consider the diagrams $d_1 = \langle\{a, b, c\}, \{\{\}, \{a, b\}, \{a, c\}\}\rangle$ and $d_2 = \langle\{a, b, c, d\}, \{\{\}, \{a, b\}, \{a, b, c, d\}\}\rangle$. The dual graphs of d_1 and d_2

are isomorphic. Each dual graph comprises three isolated vertices and no edges, hence neither diagram is drawable. However the diagram d_1 is atomic whereas d_2 is nested. Figure 14 gives the dual graphs and non-well-formed concrete representations of d_1 and d_2.

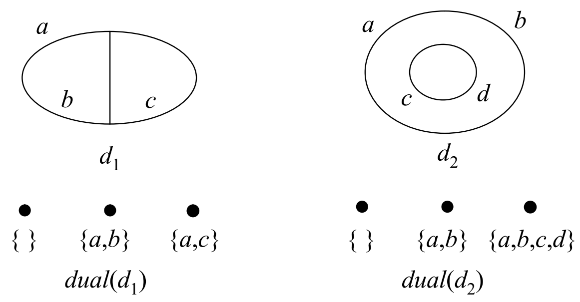

Fig. 14 Atomic and nested undrawable diagrams with isomorphic dual graphs

In the examples shown, the abstract diagram of a nested (atomic) concrete diagram is itself nested (atomic). In the next section, we generalise this observation to show that the definitions of nestedness at the concrete and abstract levels are consistent with the abstraction mapping given in definition 3.

3.3 Consistency between abstract and concrete nesting

Figure 13 illustrates the relationship between nested drawable diagrams and the presence of a cut vertex in the dual graph. We will now show two results: if a concrete diagram is nested, then its abstract diagram is also nested, and if an abstract diagram is nested, then all concrete representations of it will be nested.

Theorem 3 *Given an abstract diagram d which is nested, let $\hat{d}$ be any well-formed concrete representation. Then $\hat{d}$ must be nested.*

Proof Let $v^\star$ be a cut vertex of $dual(d)$. By lemma 1, the cut components of $dual(d)$ induce a partition $C_1, C_2..., C_n$ of the contour set $\mathcal{C}(d)$. There is a bijection from $\mathcal{C}(d) \rightarrow \widehat{\mathcal{C}}(\hat{d})$ which induces a partition $\hat{C}_1, ..., \hat{C}_n$ of the concrete contour set.

It will be sufficient to show that a contour $\hat{c}_i \in \hat{C}_i$ can never cross a contour $\hat{c}_j \in \hat{C}_j$. If $\hat{c}_i$ meets $\hat{c}_j$ and the diagram has transverse crossings, then there must be zones z, $z \cup \{c_i\}$,$z \cup \{c_j\}$ and $z \cup \{c_i, c_j\}$ in the abstract diagram d. The dual edges between z and $z \cup \{c_i\}$ and between $z \cup \{c_j\}$ and $z \cup \{c_i, c_j\}$ lie in subgraph S_i, and the dual edges between z and $z \cup \{c_j\}$ and between $z \cup \{c_i\}$ and $z \cup \{c_i, c_j\}$ lie in subgraph S_j. But the subgraphs S_i and S_j share only one vertex, the cut vertex $v^\star$. This is a contradiction, so the partition of the contour set shows that the concrete diagram $\hat{d}$ is nested. $\square$

The following result only holds in the presence of the well-formedness rules. (For example, $A \subseteq B$ can be represented by a non-nested concrete diagram if we allow tangential contours: see figure 1.)

Theorem 4 *Given a concrete diagram $\hat{d}$ which is nested, then its abstract diagram $ab(\hat{d})$ is nested.*

Recall that topological adjacency implies dual adjacency but the converse does not hold.

Proof Let $\hat{d}$ be nested, and let $\hat{\mathcal{C}}_2$ be the contours in an innermost connected component of the union of contours of $\hat{d}$. Let $\hat{\mathcal{C}}_1$ be $\widehat{\mathcal{C}}(\hat{d}) - \hat{\mathcal{C}}_2$. Think of contours in $\hat{\mathcal{C}}_2$ as being inside some simple closed curve, γ, and contours in $\hat{\mathcal{C}}_1$ being outside γ (see proposition 1 which equates alternative concrete approaches to nestedness to understand the path γ).

This enables us to partition $\widehat{\mathcal{Z}}(\hat{d}) = \hat{\mathcal{Z}}_{in} \sqcup \{\hat{z}_\gamma\} \sqcup \hat{\mathcal{Z}}_{out}$ where the zones in $\hat{\mathcal{Z}}_{in}$ have boundaries made up of contours from $\hat{C}_2$, the zones in $\hat{\mathcal{Z}}_{out}$ have boundaries made up of contours from $\hat{C}_1$ and the zone which has a boundary meeting both contours from $\hat{C}_1$ and $\hat{C}_2$ is called $\hat{z}_\gamma$:

$$\begin{aligned}
\hat{z} \in \hat{\mathcal{Z}}_{in} \wedge \hat{c} \in \hat{\mathcal{C}}_1 \quad &\Rightarrow \quad \partial\hat{z} \cap \hat{c} = \varnothing \\
\hat{z} \in \hat{\mathcal{Z}}_{out} \wedge \hat{c} \in \hat{\mathcal{C}}_2 \quad &\Rightarrow \quad \partial\hat{z} \cap \hat{c} = \varnothing \\
\exists \hat{c_1} \in \hat{\mathcal{C}}_1 \wedge \hat{c_2} \in \hat{\mathcal{C}}_2 \ \text{ such that } &\ \partial\hat{z}_\gamma \cap \hat{c_1} \neq \varnothing \wedge \partial\hat{z}_\gamma \cap \hat{c_2} \neq \varnothing
\end{aligned}$$

Given any zone $\hat{z} \in \hat{\mathcal{Z}}_{in}$, there is a path α inside γ from a point in $\hat{z}$ to a point in $\hat{z}_\gamma$. The symmetric difference between the abstract zones $z \in \mathcal{Z}_{in}$ and z_γ consists of contours in C_2. The partition of concrete zones induces a partition of abstract zones $\mathcal{Z}(d) = \mathcal{Z}_{in} \sqcup \{z_\gamma\} \sqcup \mathcal{Z}_{out}$ with the following symmetric difference properties).

$$\begin{aligned}
z \in \mathcal{Z}_{in} \quad &\Rightarrow \quad z \vartriangle z_\gamma \subseteq \mathcal{C}_1 \wedge z \vartriangle z_\gamma \neq \varnothing \\
z \in \mathcal{Z}_{out} \quad &\Rightarrow \quad z \vartriangle z_\gamma \subseteq \mathcal{C}_2 \wedge z \vartriangle z_\gamma \neq \varnothing \\
z_1 \in \mathcal{Z}_{in} \wedge z_2 \in \mathcal{Z}_{out} \quad &\Rightarrow \quad z_1 \vartriangle z_2 = (z_1 \vartriangle z_\gamma) \sqcup (z_2 \vartriangle z_\gamma)\,.
\end{aligned}$$

The symmetric difference of abstract zones in sets $\mathcal{Z}_{in}$ and $\mathcal{Z}_{out}$ contains at least two elements, so no two are adjacent. The zone z_γ acts as a pathway in the dual graph from $\mathcal{Z}_{in}$ to $\mathcal{Z}_{out}$, and is a cut vertex of the dual graph.

$\square$

4 The semantics of nested diagrams

The contours and zones of an Euler diagram represent sets and topological containment (for concrete diagrams) represents the subset relation. For example, the diagram in figure 15 denotes $(C \subseteq A \cap \overline{B}) \wedge (D \subseteq A \cap \overline{B})$ where an uppercase letter denotes the set represented by the contour with the corresponding lowercase label. We begin by formalising this intuitive 'semantic reading' of diagrams.

Definition 11 *A* ***set assignment to contours*** *for diagram d is a pair (U, ψ) where U is some universal set and $\psi : \mathcal{C}(d) \longrightarrow \mathcal{P}U$. The mapping ψ extends to $\psi : \mathcal{P}\mathcal{C}(d) \rightarrow \mathcal{P}U$ defined by*

$$z \mapsto \bigcap_{c \in z} \psi(c) \cap \bigcap_{c \in \mathcal{C}(d) - z} \overline{\psi(c)}$$

for any $z \subseteq \mathcal{C}(d)$.

The overline used here means set complement in the context of the universal set U, $\overline{\psi(c)} = U - \psi(c)$. The extension of ψ to $\mathcal{P}\mathcal{C}(d)$ ensures that two different zones correspond to disjoint sets.

Definition 12 *A set assignment to contours (U, ψ) is a* ***model*** *for diagram d if the* ***plane tiling condition*** *is satisfied:*

$$\bigcup_{z \in \mathcal{Z}(d)} \psi(z) = U.$$

The plane tiling condition simply asserts that the union of the sets represented by all the zones is the universal set.

Example 9 Consider the diagram d given in figure 15. Let the universal set be $U = \{1, 2, 3, 4, 5\}$. Define $\psi : \mathcal{C}(d) \to U$ by $a \mapsto \{1, 2, 3, 4\}; b \mapsto \{4, 5\}; c \mapsto \{1, 2\}; e \mapsto \{2, 3\}$.

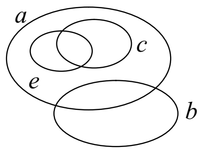

Fig. 15 Diagram for examples 9 and 10

This extends to the mapping on zones $\psi : \mathcal{Z}(d) \to \mathcal{P}U$ where:

$$\begin{array}{rr} \{a\} \mapsto \{\}, & \{b\} \mapsto \{5\}, \\ \{a,b\} \mapsto \{4\}, & \{a,c\} \mapsto \{1\}, \\ \{a,e\} \mapsto \{3\}, & \{a,c,e\} \mapsto \{2\}, \\ \{\} \mapsto \{\}. & \end{array}$$

This defines a model of d since the plane tiling condition is satisfied:

$$\bigcup_{z \in \mathcal{Z}(d)} \psi(z) = \{\} \cup \{5\} \cup \{4\} \cup \{1\} \cup \{3\} \cup \{2\} \cup \{\} = \{1, 2, 3, 4, 5\} = U.$$

Example 10 With the same diagram d given in figure 15 and again taking $U = \{1, 2, 3, 4, 5\}$, define $\psi : \{a, b, c, d\} \to \mathcal{P}U$ by $a \mapsto \{1, 2, 3, 4\}; b \mapsto \{3, 4, 5\}; c \mapsto \{1, 2\}; e \mapsto \{2, 3\}$. This extends to the mapping of zones ψ :

$\mathcal{Z}(d) \to \mathcal{P}U$ where:

$$\begin{aligned} \{a\} &\mapsto \{\}, & \{b\} &\mapsto \{\}, \\ \{a,b\} &\mapsto \{4\}, & \{a,c\} &\mapsto \{1\}, \\ \{a,e\} &\mapsto \{\}, & \{a,c,e\} &\mapsto \{2\}. \\ \{\} &\mapsto \{\}. \end{aligned}$$

The union of the sets representing zones is $\{1, 2, 4\} \neq U$ so the plane tiling condition is not satisfied. Hence the set assignment does not define a model of d.

An alternative semantic condition asserts that each of the sets representing those 'zones' *not* present in the diagram is empty. For any diagram d, we say that an element of $\mathcal{PC}(d) - \mathcal{Z}(d)$ is a **missing zone** of d. The following theorem formalises the alternative semantics.

Theorem 5 *Let (U, ψ) be a set assignment to contours for a diagram d. The plane tiling condition for d is equivalent to the following* ***missing zones condition****.*

$$\bigcup_{z \in \mathcal{PC}(d) - \mathcal{Z}(d)} \psi(z) = \varnothing.$$

Example 11 Consider the diagram whose concrete representation is given in figure 16. Let $\psi : \mathcal{C}(d) \longrightarrow U$ given by $\psi(a) = A, \psi(b) = B, \psi(c) = C$.

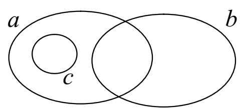

Fig. 16 Diagram for example 11

The set of zones of d is $\mathcal{Z}(d) = \{\{\}, \{a\}, \{a, c\}, \{a, b\}, \{b\}\}$. Therefore the plane tiling condition is:

$$(\overline{A} \cap \overline{B} \cap \overline{C}) \cup (A \cap \overline{B} \cap \overline{C}) \cup (A \cap \overline{B} \cap C) \cup (A \cap B \cap \overline{C}) \cup (\overline{A} \cap \ B \cap \overline{C}) = U.$$

It is easier to use the missing zones condition if we wish to obtain a simpler and more natural semantic expression. The zones missing from d are $\{c\}, \{b, c\}$ and $\{a, b, c\}$. Hence the missing zones condition is:

$$(\overline{A} \cap \overline{B} \cap C) \cup (\overline{A} \cap B \cap C) \cup (A \cap B \cap C) = \varnothing.$$

From this we may deduce $\overline{A} \cap C = \varnothing$ and $B \cap C = \varnothing$. Hence $C \subseteq A$ and $C \subseteq \overline{B}$ or, more simply, $C \subseteq A \cap \overline{B}$ which is the 'natural' reading of figure 16.

In the case of nested diagrams, we can exploit the nesting to simplify the semantics predicate. In the previous example, the diagram d represented by figure 16 is nested: we can regard it as being formed by embedding $d_2 = \langle \{c\}, \{\{\}, \{c\}\} \rangle$ in the zone $z^\star = \{a\}$ of $d_1 = \langle \{a, b\}, \{\{\}, \{a\}, \{b\}, \{a, b\}\} \rangle$. In this case, the diagrams d_1 and d_2 give no semantic information – their semantics predicates are both *true*. The natural reading of the diagram, $C \subseteq A \cap \overline{B}$ comes from the embedding.

Lemma 2 *Let d be a diagram such that $\mathcal{C}(d)$ has a partition $\{\mathcal{C}_1, \mathcal{C}_2\}$. Let (U, ψ) be a set assignment to contours for d. Let $z \subseteq \mathcal{C}(d)$. Then $z = z_1 \cup z_2$ where $z_1 \subseteq \mathcal{C}(d_1), z_2 \subseteq \mathcal{C}(d_2)$. Further,*

$$\psi(z_1 \cup z_2) = \psi_1(z_1) \cap \psi_2(z_2)$$

where, for $i = 1, 2$, ψ_i is the extension to $\mathcal{PC}(d_i)$ of the restriction of ψ to $\mathcal{C}(d_i)$:

$$\psi_i(z_i) = \bigcap_{c \in z_i} \psi(c) \cap \bigcap_{c \in \mathcal{C}(d_i) - z_i} \overline{\psi(c)}.$$

□

Theorem 6 *Let d be a nested diagram formed by embedding d_2 in zone $z^\star$ of d_1: $d = d_2 \xrightarrow{z^\star} d_1$. Let (U, ψ) be a set assignment to contours for d and let $SP(d)$ denote the semantics predicate of diagram d (which, by theorem 5, could be either the plane tiling condition or the missing zones condition). Then*

$$SP(d) = SP(d_1) \wedge SP(d_2) \wedge \bigwedge_{c \in \mathcal{C}_2} \psi(c) \subseteq \psi_1(z^\star).$$

*The condition $\bigwedge_{c \in \mathcal{C}_2} \psi(c) \subseteq \psi_1(z^\star)$ is called the **embedding condition**.*

The situation in theorem 6 is illustrated in figure 17.

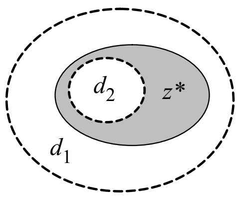

Fig. 17 Embedding d_2 in zone $z^\star$ of d_1

Proof We use the missing zones condition. The strategy of the proof is to show that the zones missing from d are missing because either they are missing from d_1 or they are missing from d_2 or they are missing as a result of the embedding. Recall that the set of zones of d is the union $\mathcal{Z}(d_1) \cup \mathcal{Z}_2^\star$

where $\mathcal{Z}_2^\star = \{z^\star \cup z_2 : z_2 \in \mathcal{Z}(d_2)\}$. We may rewrite this as

$$\mathcal{Z}(d) = \{z_1 \cup z_2 : (z_1 = z^\star \wedge z_2 \in \mathcal{Z}(d_2)) \vee (z_1 \in \mathcal{Z}(d_1) \wedge z_2 = \varnothing)\}.$$

Now

$$\begin{aligned} \mathcal{PC}(d) = \{&z_1 \cup z_2 : z_1 \in \mathcal{PC}(d_1) \wedge z_2 \in \mathcal{PC}(d_2) - \mathcal{Z}(d_2)\} \\ &\cup \{z_1 \cup z_2 : z_1 \in \mathcal{PC}(d_1) - \mathcal{Z}(d_1) \wedge z_2 \in \mathcal{PC}(d_2)\} \\ &\cup \{z_1 \cup z_2 : z_1 \in \mathcal{Z}(d_1) \wedge z_2 \in \mathcal{Z}(d_2)\}. \end{aligned}$$

Therefore the missing zones in d are

$$\begin{aligned} &\{z_1 \cup z_2 : z_1 \in \mathcal{PC}(d_1) \wedge z_2 \in \mathcal{PC}(d_2) - \mathcal{Z}(d_2)\} \\ \cup\ &\{z_1 \cup z_2 : z_1 \in \mathcal{PC}(d_1) - \mathcal{Z}(d_1) \wedge z_2 \in \mathcal{PC}(d_2)\} \\ \cup\ &\{z_1 \cup z_2 : z_1 \in \mathcal{PC}(d_1) \wedge z_1 \neq z^\star \wedge z_2 \in \mathcal{PC}(d_2) \wedge z_2 \neq \varnothing\}. \end{aligned}$$

The missing zones condition for d can therefore be expressed as

$$\begin{aligned} SP(d) = &\bigwedge_{\substack{z_1 \in \mathcal{PC}(d_1) - \mathcal{Z}(d_1) \\ z_2 \in \mathcal{PC}(d_2)}} \psi(z_1 \cup z_2) = \varnothing \\ \wedge &\bigwedge_{\substack{z_1 \in \mathcal{PC}(d_1) \\ z_2 \in \mathcal{PC}(d_2) - \mathcal{Z}(d_2)}} \psi(z_1 \cup z_2) = \varnothing \\ \wedge &\bigwedge_{\substack{z_1 \in \mathcal{PC}(d_1) \wedge z_1 \neq z^\star \\ z_2 \in \mathcal{PC}(d_2) \wedge z_2 \neq \varnothing}} \psi(z_1 \cup z_2) = \varnothing. \end{aligned}$$

Using lemma 2, we may rewrite this as

$$SP(d) = \bigwedge_{\substack{z_1 \in \mathcal{PC}(d_1) - \mathcal{Z}(d_1) \\ z_2 \in \mathcal{PC}(d_2)}} \psi_1(z_1) \cap \psi_2(z_2) = \varnothing \tag{1}$$

$$\wedge \bigwedge_{\substack{z_1 \in \mathcal{PC}(d_1) \\ z_2 \in \mathcal{PC}(d_2) - \mathcal{Z}(d_2)}} \psi_1(z_1) \cap \psi_2(z_2) = \varnothing \tag{2}$$

$$\wedge \bigwedge_{\substack{z_1 \in \mathcal{PC}(d_1) \wedge z_1 \neq z^\star \\ z_2 \in \mathcal{PC}(d_2) \wedge z_2 \neq \varnothing}} \psi_1(z_1) \cap \psi_2(z_2) = \varnothing. \tag{3}$$

Consider (1). For any $z_1 \in \mathcal{PC}(d_1) - \mathcal{Z}(d_1)$, we have

$$\begin{aligned} \bigcup_{z_2 \in \mathcal{Z}(d_2)} \psi_1(z_1) \cap \psi_2(z_2) = \varnothing \quad &\Rightarrow \quad \psi_1(z_1) \cap \bigcup_{z_2 \in \mathcal{Z}(d_2)} \psi_2(z_2) = \varnothing \\ &\Rightarrow \quad \psi_1(z_1) = \varnothing, \end{aligned}$$

since $\bigcup_{z_2 \in \mathcal{Z}(d_2)} \psi_2(z_2) = U$, by the plane tiling condition for d_2. Therefore (1) reduces to the missing zones condition for d_1. By interchanging the roles of d_1 and d_2 it is clear that (2) reduces to the missing zones condition for d_2. Therefore

$$\begin{aligned} SP(d) &= SP(d_1) \wedge SP(d_2) \wedge \bigwedge_{\substack{z_1 \in \mathcal{PC}(d_1) \wedge z_1 \neq z^\star \\ z_2 \in \mathcal{PC}(d_2) \wedge z_2 \neq \varnothing}} \psi_1(z_1) \cap \psi_2(z_2) = \varnothing \\ &= SP(d_1) \wedge \quad SP(d_2) \wedge \bigwedge_{\substack{z_2 \in \mathcal{PC}(d_2) \\ z_2 \neq \varnothing}} \left(\bigcup_{\substack{z_1 \in \mathcal{PC}(d_1) \\ z_1 \neq z^\star}} \psi_1(z_1) \right) \cap \psi_2(z_2) = \varnothing. \end{aligned}$$

By the plane tiling condition for d_1 and the fact that distinct zones represent disjoint sets, we have

$$\bigcup_{\substack{z_1 \in \mathcal{PC}(d_1) \\ z_1 \neq z^\star}} \psi_1(z_1) = U - \psi_1(z^\star) = \overline{\psi_1(z^\star)}.$$

Hence, finally,

$$\begin{aligned} SP(d) &= SP(d_1) \wedge SP(d_2) \wedge \bigwedge_{\substack{z_2 \in \mathcal{PC}(d_2) \\ z_2 \neq \varnothing}} \psi_2(z_2) \cap \overline{\psi_1(z^\star)} = \varnothing \\ &= SP(d_1) \wedge SP(d_2) \wedge \bigwedge_{c \in \mathcal{C}(d_2)} \psi(c) \cap \overline{\psi_1(z^\star)} = \varnothing \\ &= SP(d_1) \wedge SP(d_2) \wedge \bigwedge_{c \in \mathcal{C}(d_2)} \psi(c) \subseteq \psi_1(z^\star), \end{aligned}$$

which completes the proof. $\square$

Corollary 1 *Let $d = d_1 \xrightarrow{\varnothing} d_2$ and let (ψ, U) be a set assignment to contours for d. Then*

$$SP(d) = SP(d_1) \wedge SP(d_2) \wedge \bigwedge_{\substack{c_1 \in \mathcal{C}_1 \\ c_2 \in \mathcal{C}_2}} \psi(c_1) \cap \psi(c_2) = \varnothing$$

where $SP(d)$ again denotes the semantic predicate for d (either the plane tiling condition or the missing zones condition). $\square$

Example 12 Let d be the diagram represented by figure 18 and let $d_1 = \langle\{a, b, c\}, \{\{\}, \{a\}, \{b\}, \{a, b\}, \{a, c\}, \{a, b, c\}\}\rangle$ and $d_2 = \langle\{e, f\}, \{\{\}, \{e\}, \{f\}\}\rangle$. Then d is formed by embedding d_2 in zone $\{b\}$ of d_1.

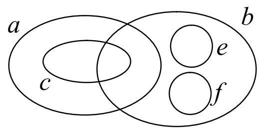

Fig. 18 Using the 'nested semantics' of theorem 6

Let $\psi : \mathcal{C}(d) \to U$ be a set assignment to contours given by $\psi(a) = A, \psi(b) = B, \psi(c) = C, \psi(e) = E, \psi(f) = F$. The missing zones of d_1 are $\{c\}$ and $\{b, c\}$ so the missing zones condition for d_1 is $(\overline{A} \cap \overline{B} \cap C) \cup (\overline{A} \cap B \cap C) = \varnothing$. This gives $\overline{A} \cap C = \varnothing$ or $C \subseteq A$.

The diagram d_2 has missing zone $\{e, f\}$ so the missing zone condition for d_2 is $E \cap F = \varnothing$.

The embedding condition is $(E \subseteq \overline{A} \cap B \cap \overline{C}) \wedge (F \subseteq \overline{A} \cap B \cap \overline{C})$.

Therefore the semantics predicate of d is:

$$(C \subseteq A) \wedge (E \cap F = \varnothing) \wedge (E \subseteq \overline{A} \cap B \cap \overline{C}) \wedge (F \subseteq \overline{A} \cap B \cap \overline{C}).$$

Since $\overline{A} \subseteq \overline{C}$, this simplifies to

$$(C \subseteq A) \wedge (E \cap F = \varnothing) \wedge (E \subseteq \overline{A} \cap B) \wedge (F \subseteq \overline{A} \cap B)$$

which, again, is the 'natural' reading of d.

5 Constructing atomic and nested diagrams

An algorithm has been devised and implemented to create drawings of drawable Euler diagrams [4]. To enhance the efficiency of the algorithm and the readability of its output, we describe here an approach to make use of nesting in the abstract Euler diagrams.

Given an abstract Euler diagram d whose dual has a cut vertex, let $S_1, ..., S_n$ be the cut components of the $dual(d)$ (obtained by removing the cut vertex and replacing it, in turn, to each component). Without loss of generality, S_1 contains a vertex labelled by the empty set. (Possibly other subgraphs do too, if the cutvertex is the null vertex). Draw a concrete representation for the diagram whose dual is S_1, and add to it places to insert $n-1$ other diagrams inside the zone corresponding to the cut-vertex. We can think of the diagram as a *template*, as in [5].

The other $n-1$ subgraphs of the dual have vertex labels which are all supersets of the cut vertex. Replace each abstract zone z with $z - v^{\star}$. Then each of subgraph can be represented by a concrete diagram. These concrete diagrams are inserted into the template, to make up $\hat{d}$.

Figure 19 shows an example where two cut vertices are identified in the dual graph, $\{\}$ and $\{a\}$. These are used to split the dual graph into

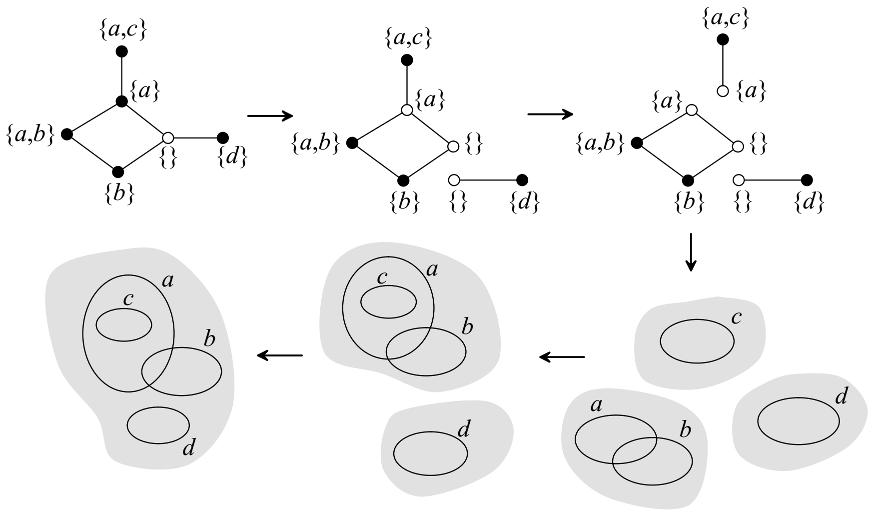

Fig. 19 Cut vertices and diagram construction

three components, and each component is used to draw a concrete Euler diagram. Information about the cut vertices is then used to combine the three concrete Euler diagrams into a single concrete representation.

Figure 20 shows the same example where one diagram component is interpreted as a *containing diagram*, and is used as template. The other components are inserted into two different zones of the containing diagram. The implementation challenges here are to be able to identify rectangles in the concrete zones, rectangles within which to insert other diagrams.

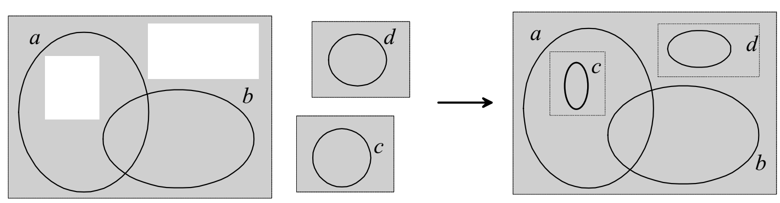

Fig. 20 Constructing concrete nested diagrams using templates

One implementation of this template application uses information about containment of sub-diagrams to enlarge zones. Output from this implementation is shown in figure 21. Some smoothing work can be done to improve the appearance of the output, but the nesting has been used to create containing diagram and contained diagram independently before insertion.

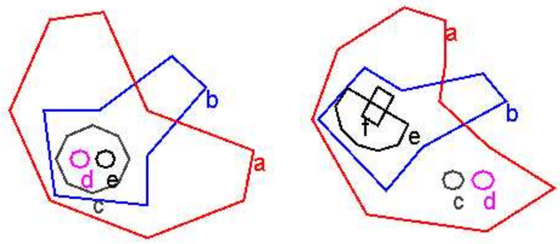

Fig. 21 Output from a java implementation using nesting

6 Counting atomic and nested diagrams

To see the leverage gained by using the nesting concept in semantics or drawing problems, consider how the numbers of abstract diagrams grow with the number of contours. The following table shows how many well-formed diagrams there are with a given number of contours (by row) and a given number of zones (by column). The number of diagrams is seen to grow quickly, but the number of atomic diagrams, shown in brackets, grows much less quickly. Drawing nested diagrams using templates as described in the previous section can handle the vast majority of diagrams, leaving just a few atomic examples to be drawn without using a template.

number of zones

	nested:atomic	3z	4z	5z	6z	7z	8z
	2c	2:0	0:1	0:0	0:0	0:0	0:0
number of contours	3c	0:0	4:0	4:0	0:3	0:3	0:1
	4c	0:0	0:0	9:0	15:0	20:0	16:14
	5c	0:0	0:0	0:0	20:0	50:0	101:0

7 Summary

In this paper we extended work on Euler diagrams by considering the notion of a *nested* diagram. The concept of nesting is most obvious, visually, for concrete diagrams; however we showed that the concept extends to abstract diagrams. We defined the notion of nesting for both concrete and abstract Euler diagrams. We showed further that a *drawable* abstract diagram diagram is nested if and only if its dual graph contains a cut vertex. The notions of nesting at the concrete and abstract levels are shown to be equivalent for drawable abstract diagrams under the morphism from concrete to abstract diagrams. Nesting in diagrams gives rise to different ways of writing down diagram semantics, and we developed a “nested form” for diagram semantics.

Most Euler diagrams are nested. The table in section 6 is interesting in that in almost all entries at least one of the numbers is zero. This can be explained by the fact that any Euler diagram with n contours containing fewer than $2n$ zones is nested and any Euler diagram with n contours containing greater than $2^{n-1}+1$ zones is atomic. Research into methods for counting the diagrams is ongoing; we can use, for example, tree-counting [2] or *Polya's Counting Theorem* [1] to count some classes of Euler diagrams.

One of the aims of the work presented in this paper 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

The authors are grateful to Gem Stapleton and Andrew Fish for their valuable comments on an earlier draft of this paper. This research was partially supported by UK EPSRC grant GR/R63516

References

1. Biggs N. Discrete Mathematics 2nd edition, OUP, 2002.
2. Cayley, A. "On the analytical forms called trees", American Journal of Mathematics 4, 1881, p.266-268.
3. Euler L. *Lettres a Une Princesse d'Allemagne*, vol 2. 1761. Letters No. 102-108.
4. Flower, J. and Howse, J., "Generating Euler Diagrams", Diagrammatic Representation and Inference, Proc. Diagrams 2002, LNAI 2317, Springer Verlag, 2002, p.61-75.

5. Flower J., Howse, J. and Taylor, J.,“A framework for modelling with heterogeneous notations”, Proc IEEE Symp on Human-Centric computing, 2002, p.71-73.

6. Gil, J., Howse, J., Kent, S.,“Towards a formalisation of constraint diagrams”, Proc IEEE Symp on Human-Centric computing, 2001, p.72-79.

7. Harel D. On visual formalisms. In J. Glasgow, N. H. Narayan, and B. Chandrasekaran, editors, *Diagrammatic Reasoning*, pages 235–271. MIT Press, 1998.

8. Howse, J., Molina, F, Shin, S.J. and Taylor, J., “On Diagram tokens and types”, Diagrammatic Representation and Inference, Proc. Diagrams 2002, LNAI 2317, Springer Verlag, 2002, p.146-160.

9. The Kent Modelling Framework, KMF 1.0. www.cs.ukc.ac.uk/kmf.

10. Object Management Group. “UML Specification, Version 1.3:” www.omg.org.