Reasoning with Spider Diagrams

John Howse, Fernando Molina, John Taylor
School of Computing and Mathematical Sciences
University of Brighton, UK
{John.Howse, F.Molina, John.Taylor}@brighton.ac.uk

Stuart Kent
Computing Laboratory
University of Kent, Canterbury, UK
S.J.H.Kent@ukc.ac.uk

Abstract

Spider diagrams combine and extend Venn diagrams and Euler circles to express constraints on sets and their relationships with other sets. These diagrams can usefully be used in conjunction with object-oriented modelling notations such as the Unified Modelling Language. This paper summarises the main syntax and semantics of spider diagrams and introduces four inference rules for reasoning with spider diagrams and a rule governing the equivalence of Venn and Euler forms of spider diagrams. This paper also details rules for combining two spider diagrams to produce a single diagram which retains as much of their combined semantic information as possible and discusses disjunctive diagrams as one possible way of enriching the system in order to combine spider diagrams so that no semantic information is lost.

Keywords Diagrammatic reasoning, visual formalisms.

1. Introduction

In object-oriented software development, diagrammatic modelling notations are used to specify systems. Recently, the Unified Modelling Language (UML) (Rumbaugh, Jacobson, Booch, 1999) has become the Object Management Group's (OMG) standard for such notations. In UML, constraints, such as invariants, preconditions and postconditions, are expressed using the Object Constraint Language (OCL) (Warmer and Kleppe, 1998), essentially a stylised, but textual, form of first-order predicate logic, which is part of the UML standard. *Constraint diagrams* (Kent, 1997) provide a diagrammatic notation for expressing constraints and can be used in conjunction with UML and OCL. *Spider diagrams* (Gil, Howse, Kent, 1999) emerged from work on constraint diagrams. They combine and extend Venn diagrams and Euler circles to express constraints on sets and their relationships with other sets.

Euler circles (Euler, 1761) were introduced to illustrate relations between classes. This notation uses the topological properties of enclosure, exclusion and intersection to represent the set-theoretic notions of subset, disjoint sets, and set intersection, respectively. Venn (1880) modified this notation to represent logical propositions. In Venn diagrams, all possible intersections of the closed curves must be shown and shading is used to show that a particular region represents the empty set. Peirce (1933) modified Venn diagrams by including *X-sequences* to introduce elements and disjunctive information into the system. Recently, full semantics and inference rules have been developed for Venn-Peirce diagrams (Shin, 1994) and Euler diagrams (Hammer, 1995). This paper extends these diagrammatic inference rules to spider diagrams.

A more detailed discussion of spider diagrams is conducted in section 2, where the main syntax and semantics of the notation is introduced. Section 3 introduces inference rules for reasoning with spider diagrams together with outline proofs of the validity of some of the rules. Section 3 also introduces a rule which governs the equivalence of Venn and Euler forms of spider diagrams. Section 4 is concerned with the rules for combining two spider diagrams to produce a single diagram which retains as much of their combined semantic information as possible. Section 5 indicates one possible way of enriching the system in order to combine spider diagrams so that no semantic information is lost.

2. Spider diagrams

This section introduces the main syntax and semantics of spider diagrams; see (Gil, Howse, Kent, 1999) for more details and examples. Spider diagrams are Euler circles augmented with *shaded regions* and *spiders*. Spider diagrams also include the concepts of *Schrödinger spiders* and *projections*; these are not necessary for this paper and are omitted from this discussion. In (Gil, Howse, Kent, 1999), the distinction is made between given and existential spiders; in this paper, all spiders are given.

2.1. Syntactic elements of spider diagrams

A *contour* is a simple closed plane curve. A *boundary contour* is not contained in and does not intersect with any other contour. A *district* (or *basic region*) is the bounded region of the plane enclosed by a contour. A *region* is defined as follows: any district is a region; if r_1 and r_2 are regions, then the union, intersection, or 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 sets.

A *spider* is a tree with nodes (called *feet*) placed in different zones; the connecting edges (called *legs*) are straight lines. A spider *touches* a zone if one of its feet appears in that region. A spider may only touch a zone once. A spider is said to *inhabit* the region which is the union of the zones it touches. For any spider *s*, the *habitat* of *s*, denoted $\eta(s)$, is the region inhabited by *s*. The set of spiders touching region *r* is denoted by $S(r)$. Spiders are used to denote elements. In this paper, all spiders represent given elements. Two distinct spiders denote distinct elements, unless they are joined by a *tie* or by a *strand*.

A *tie* is a double, straight line (an equals sign) connecting two feet, from different spiders, placed in the same zone. The *nest* of spiders *s* and *t*, written $\tau(s, t)$, is the union of those zones *z* having the property that there is a sequence of spiders

$$s = s_0, s_1, s_2, \ldots, s_n = t$$

such that, for $i = 0, \ldots, n-1$, s_i and s_{i+1} are connected by a tie in *z*. Two spiders which have a non-empty nest are referred to as *mates*. If both the elements denoted by spiders *s* and *t* are in the set denoted by the same zone in the nest of *s* and *t*, then *s* and *t* denote the same element.

A *strand* is a wavy line connecting two feet, from different spiders, placed in the same zone. The *web* of spiders *s* and *t*, written $\zeta(s, t)$, is the union of zones *z* having the property that there is a sequence of spiders

$$s = s_0, s_1, s_2, \ldots, s_n = t$$

such that, for $i = 0, \ldots, n-1$, s_i and s_{i+1} are connected by a tie or by a strand in *z*. So $\tau(s, t)$ is a subregion of $\zeta(s, t)$. Two spiders with a non-empty web are referred to as *friends*. Two spiders *s* and *t* may (but not necessarily must) denote the same element if that element is in the set denoted by the web of *s* and *t*. Clearly, if there is a tie between feet, then a strand between those feet is redundant. Similarly, multiple strands or ties between the same pairs of feet are redundant.

In later sections, we will need to compare webs and nests of spiders across diagrams. To facilitate this, we extend the notation and use $\zeta(s, t, D)$ and $\tau(s, t, D)$ to denote the web and nest respectively of spiders *s* and *t* in the diagram *D*.

Every region is a union of zones. A region is *shaded* if each of its component zones is shaded. A shaded region containing no spiders denotes the empty set. Shading a region *r* which includes spiders has the effect of placing an upper limit on the number of elements in the set denoted by the region. An upper bound is $|S(r)|$, but this might not be a least upper bound.

A *spider diagram* is a finite collection of contours (exactly one of which must be a boundary contour), spiders, strands, ties and shaded regions. For any spider diagram *D*, we use $C = C(D)$, $R = R(D)$, $Z = Z(D)$, $Z^* = Z^*(D)$ and $S = S(D)$ to denote the sets of contours, regions, zones, shaded zones and spiders of *D*, respectively.

The *Venn form* of a spider diagram contains every possible intersection of contours; otherwise, the diagram is in *Euler form*. A spider diagram with *n* (non-boundary) contours has 2^n zones if and only if it is in Venn form.

The spider diagram *D* in figure 2.1 is in Venn form. It has three non-boundary contours *A*, *B*, *C* and two spiders *s* and *t*. The label *s* refers to the whole spider and not just to any particular node. There is a tie between *s* and *t* in $A \cap C - B$ and a strand between *s* and *t* in $B - (A \cup C)$. Here are some properties of (the denotation of) *D*:

$A - (B \cup C) = \{\}$, $|(B \cap C) - A| \leq 1$,

$s \in (B - C) \cup (A \cap C - B)$,

$t \in (B - A \cap B \cap C) \cup (A \cap C - B)$,

$s, t \in A \cap C - B \Rightarrow s = t$, $s, t \in A \cap B - C \Rightarrow s \neq t$.

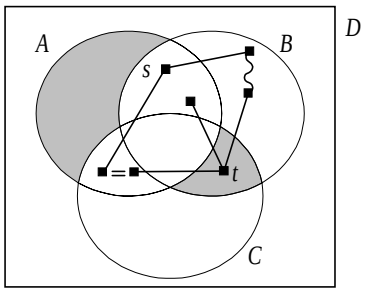


Figure 2.1

2.2. Semantics of spider diagrams

The semantics of a spider diagram *D* is given in terms of the semantic functions

$$\Psi : C \to Set\ \mathbf{U}, \quad \psi : S \to \mathbf{U}$$

where **U** is a given universal set of *D* and *Set* **U** denotes the power set of **U**. Contours are interpreted as subsets of **U**, and spiders as elements of **U**. The boundary contour is interpreted as **U**.

A zone is uniquely defined by the contours containing it and the contours not containing it; its interpretation is the intersection of the sets denoted by the contours containing it and the complements of the sets denoted by those contours not containing it. We extend the domain of Ψ to interpret regions as subsets of **U**. First define $\Psi : Z \to Set\ \mathbf{U}$ by

$$\Psi(z) = \bigcap_{c \in C^+(z)} \Psi(c) \cap \bigcap_{c \in C^-(z)} \overline{\Psi(c)}$$

where $C^+(z)$ is the set of contours containing the zone *z*, $C^-(z)$ is the set of contours not containing *z* and

$\overline{\Psi(c)} = U - \Psi(c)$, the *complement* of $\Psi(c)$. Since any region is a union of zones, we may define $\Psi: R \to Set\,\mathbf{U}$ by

$$\Psi(r) = \bigcup_{z \in Z(r)} \Psi(z)$$

where, for any region r, $Z(r)$ is the set of zones contained in r.

The semantics of a diagram D is the conjunction of the following conditions.

Plane Tiling Condition: All elements fall within sets denoted by zones:

$$\bigcup_{z \in Z} \Psi(z) = \mathbf{U}$$

Spider Condition: The element denoted by a spider is in the set denoted by the habitat of the spider:

$$\bigwedge_{s \in S} \psi(s) \in \Psi(\eta(s))$$

Strangers Condition: The elements denoted by two distinct spiders are distinct unless they fall within the set denoted by the spiders' web:

$$\bigwedge_{\substack{s,\, t \in S \\ s \neq t}} \psi(s) = \psi(t) \;\Rightarrow\; \psi(s), \psi(t) \in \Psi(\zeta(s, t))$$

Mating Condition: If the elements denoted by two distinct spiders fall within the set denoted by the same zone in the spiders' nest, then the elements are equal:

$$\bigwedge_{s,\, t \in S} \bigwedge_{z \in Z(\tau(s,\, t))} \psi(s), \psi(t) \in \Psi(z) \;\Rightarrow\; \psi(s) = \psi(t)$$

Shading Condition: The set denoted by a shaded zone contains no elements other than those denoted by the spiders:

$$\bigwedge_{z \in Z^*} \Psi(z) \subseteq \bigcup_{s \in S} \{\psi(s)\}$$

3. Reasoning with spider diagrams

We introduce rules for manipulating single diagrams. We also give outline proofs of the validity of some of the inference rules.

3.1. Rules of transformation

We introduce five rules for manipulating single diagrams. The first four are inference rules that allow us to obtain one diagram from a given diagram by removing, adding or modifying diagrammatic elements. The last rule governs the equivalence of the Euler and Venn forms of spiders diagrams.

Rule 1: Introduction of a strand. A strand may be drawn between the feet of any two spiders in the same zone. Similarly, any tie may be replaced with a strand.

Example 3.1 Introducing a strand between two non-connected feet in a zone weakens the information contained in the diagram. In figure 3.1, the spiders *s* and *u* in diagram *D* represent distinct elements but in *D′* they may represent the same element of *B* – *A*.

Similarly, replacing a tie between the feet of two spiders with a strand also weakens the semantic information given by the diagram. If the element denoted by *s* lies in *A* – *B*, then, in *D*, *s* and *t* are necessarily equal whereas in *D′* they need not be.

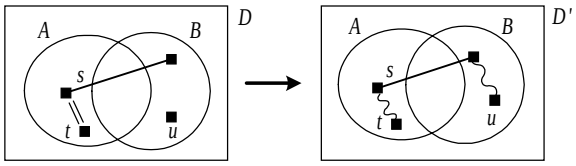


Figure 3.1

Rule 2: Erasure of a diagrammatic element. We may erase:

- the shading in an entire zone.
- a complete spider on any non-shaded region and any strand or tie connected to it. If removing a spider disconnects any component of the 'strand-tie graph' in a zone, then the components so formed should be reconnected using one or more strands to restore the original component.
- a contour. When a contour is erased:
  - any shading remaining in only a part of a zone should also be erased.
  - if a spider has feet in two regions which combine to form a single zone with the erasure of the contour, then these feet are replaced with a single foot connected to the rest of the spider and any ties connecting it in the new zone should be replaced by strands.

Example 3.2 In Figure 3.2, erasing the spider *u* and its two connecting strands disconnects spiders *s* and *t* in the zone *A* – *B*. However, the web of *s* and *t* is the region *A* – *B*, and this must not change with the deletion of *u*. Hence in *D′* the spiders are explicitly 'reconnected' by joining them with a strand.

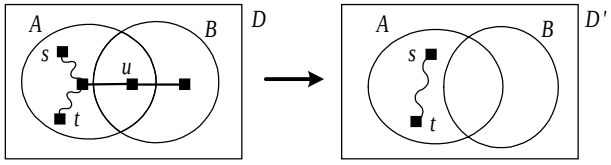


Figure 3.2

Example 3.3 The requirement that the region from which a spider is removed should be non-shaded is a necessary one. Figure 3.3 illustrates that the removal of a spider from a shaded zone may result in an invalid inference (see section 3.2). In diagram *D*, the set corresponding to region $A - B$ contains a single element, whereas in D', the corresponding set is empty.

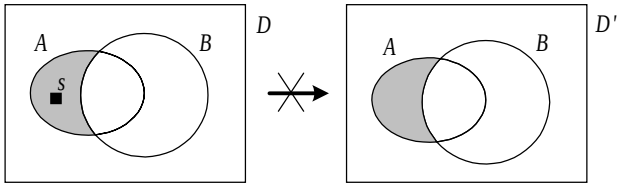


Figure 3.3

Example 3.4 Erasing a contour can cause both syntactic and semantic difficulties.

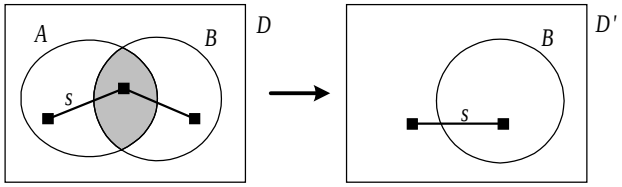


Figure 3.4

Figure 3.4 illustrates the syntactic difficulties. Simply erasing the contour *A* in the diagram *D*, the (new) zone *B* becomes partially shaded and the spider *s* has two feet in the new zone *B*. To ensure that the resulting diagram D' is well-formed, the partial shading must be erased and the feet of *s* in *B* should be replaced with a single foot.

The last part of rule 2 concerns semantic difficulties connected with erasing a contour and is a little more subtle.

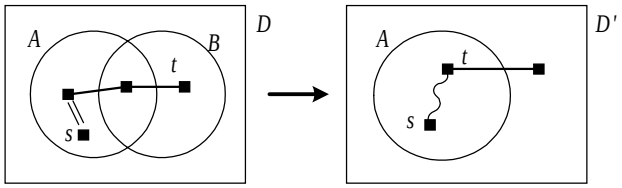


Figure 3.5

Consider the diagram *D* shown in figure 3.5. The diagram has a model in which the elements corresponding to spiders *s* and *t* both belong to the set *A* but are distinct; namely, the model where $s \in A - B$ and $t \in A \cap B$. When the contour *B* is removed, these two zones $A \cap B$ and $A - B$ 'combine' to form the single zone *A* in D'. Since it is possible for *s* and *t* to represent distinct elements of *A*, the tie connecting them must be replaced with a strand.

Rule 3: Spreading the feet of a spider. If a diagram has a spider *s*, then we may draw a node in any zone *z* which does not contain a foot of *s* and connect it to *s*. If *z* contains the foot of another spider *t*, then we may join the feet of *s* and *t* with a strand or a tie or leave the feet separated in *z*.

Example 3.5 Rule 3 is illustrated by the diagrams in figure 3.6. The inference from *D* to D' requires two applications of rule 3, but is clearly valid since it just represents a weakening of information. From *D* we know that the element corresponding to *s* belongs to $A - B$. Having spread its feet in D', we may only infer that this element belongs to $A \cup B$.

In the zone corresponding to $A \cap B$, we have chosen to keep the feet of *s* and *t* separated; in the zone corresponding to $B - A$, we have joined the feet of *s* and *t* with a strand.

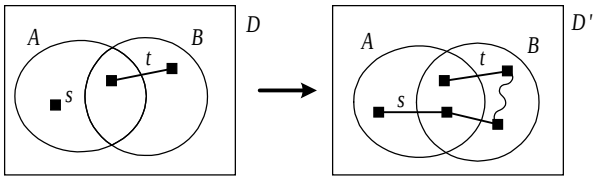


Figure 3.6

Rule 4: Introduction of a contour. A new contour may be drawn interior to the bounding rectangle observing the partial-overlapping rule: each zone splits into two zones with the introduction of the new contour. Each foot of a spider is replaced with a connected pair of feet, one in each new zone. Likewise, each strand or tie bifurcates and becomes a pair of strands or ties, one in each new zone.

Example 3.6 In figure 3.7, a new contour *B* is introduced satisfying the partial overlapping rule. Each zone in *D* becomes a pair of zones in D' and each foot of spiders *s*, *t* and *u* bifurcates to become two feet, one in each new zone. The strand and tie also bifurcate.

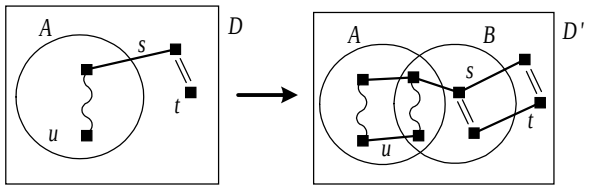


Figure 3.7

Rule 5: Equivalence of Venn and Euler forms. We may replace a diagram *D* in which some regions do not exist by a diagram *V*(*D*) in Venn form where those regions are shaded. All other diagrammatic elements—

other shaded regions, spiders, strands and ties—remain unchanged.

Conversely, we may replace a diagram *D* in Venn form which has a set of shaded zones containing no spider by a diagram *E* where (some of) those regions do not exist. Again, all other diagrammatic elements—other shaded regions, spiders, strands and ties—remain unchanged.

The transition from the Euler to the Venn form of a spider diagram is algorithmic. There are various known algorithms for constructing a Venn diagram with *n* contours—for example, see (Hammer, 1995). Given a spider diagram *D* in Euler form, first construct the underlying Venn diagram whose set of contours is *C*(*D*). Shade any zones which were not present in the original Euler form *D*. Finally add spiders, strands and ties in order to replicate the strand-tie graph in each zone of *D*. The resulting spider diagram is *V*(*D*), the Venn form of *D*.

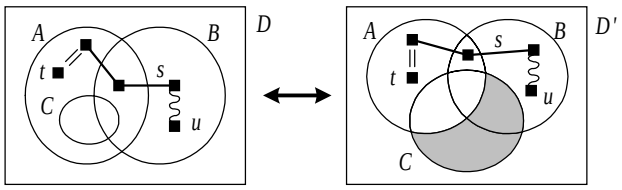


Figure 3.8

Example 3.7 Figure 3.8 illustrates the equivalence between the Euler and Venn forms of a spider diagram. The Euler form *D* does not contain zones corresponding to $\overline{A} \cap \overline{B} \cap C$ or $\overline{A} \cap B \cap C$. In the Venn form *D*′, the corresponding regions are shaded, but the strand-tie graph in every other zone is the same as the corresponding graph in *D*.

3.2. Validity of the inference rules.

In this section, we outline the proofs of validity of some of the inference rules introduced in the previous section. To say that a rule is *valid*, we mean that whenever a diagram *D*′ is obtained from another diagram *D* by a (single) application of the rule, we may infer the semantics of *D*′ from the semantics of *D*. (For further details, see (Howse, Molina, Taylor, 1999).)

Several of the rules amount to 'throwing away' some of the semantic information contained in a diagram, in the sense described in the following lemma. Note that we adopt the convention that the conjunction of an empty set of propositions equates to true.

Lemma 3.1 If diagrams *D* and *D*′ have semantics of the form $\bigwedge_{i \in I} P_i$ and $\bigwedge_{i \in J} P_i$ respectively, where $J \subseteq I$, then *D*′ is a valid inference from *D*.

Rule 1: Introducing a strand. Suppose two spiders *s* and *t* have feet which are separated (that is, not joined by a strand or a tie) in a zone *z* belonging to diagram *D*. Let *D*′ be the diagram obtained from *D* by adding a strand between the feet of *s* and *t* in *z*. Then

$$\zeta(s, t, D') = \zeta(s, t, D) \cup z.$$

The Strangers Condition is the only semantic condition which involves web of *s* and *t*; for these spiders the condition is

$$\psi(s) = \psi(t) \Rightarrow \psi(s), \psi(t) \in \Psi(\zeta(s, t, D)) .$$

Since $\zeta(s, t, D) \subseteq \zeta(s, t, D')$, we can infer the corresponding condition for *D*′. All the other semantic conditions are identical for *D* and *D*′, so the first part of rule 1 is valid.

To justify the validity of the second part of the rule, suppose *D* and *D*′ are as described above except that, in *D*, the spiders *s* and *t* are joined by a tie in *z*. In this case, the web of *s* and *t* is unchanged, but their nest changes between the diagrams:

$$\tau(s, t, D') = \tau(s, t, D) - z.$$

Thus it is only the Mating Condition which changes in *D*′. For *s* and *t*, the Mating Condition is a conjunction of terms of the form

$$\big(\psi(s) \in \Psi(z) \wedge \psi(t) \in \Psi(z)\big) \Rightarrow \psi(s) = \psi(t) ,$$

one term for each zone *z* in the nest of *s* and *t*. By lemma 3.1, we may infer the Mating Condition of *D*′ from that of *D*.

Rule 2: Erasure of a diagrammatic element. Erasing the shading in a zone only changes the Shading Condition by removing conjuncts, so the validity of the first part of rule 2 follows by lemma 3.1.

The validity of the rule for erasing a spider follows similarly. However, in passing from the semantics of *D* to that of *D*′, one or more conjuncts may be lost from the Spider, Strangers and Mating conditions.

Lack of space prevents us from including the proofs of validity of the remaining rules. These proofs may be found in (Howse, Molina, Taylor, 1999).

4. Combining Diagrams

Given two diagrams, D_1 and D_2, we wish to combine them to produce a single diagram *D* which retains as much of their combined semantic information as possible. Of course, this is only meaningful if the pair D_1, D_2 is consistent.

In this section we describe the construction of such a combined diagram *D*. Even in simple cases, some information contained in the pair D_1, D_2 will be lost in the combination. In the next section, we will indicate

one possible way of enriching the system of spider diagrams to overcome this problem.

4.1. Comparing regions

To carry through the process of combining diagrams, we will need to be able to identify corresponding regions in different diagrams. For simplicity, we consider the case where a diagram D' is obtained from a diagram D by adding contours, so that

$$C(D) \subseteq C(D').$$

There is a natural mapping

$$\alpha\colon Z(D) \to R(D')$$

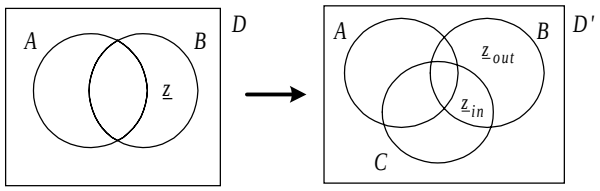
which may be defined inductively, with the inductive step as follows. Suppose that D' is obtained from D by adding a single contour. According to Rule 4, each zone z in D bifurcates into two zones z_{in} and z_{out} in D'; z_{in} is that part of z enclosed within the new contour and z_{out} is that part of z lying outside the new contour (see Figure 4.1). In this case, we define

$$\alpha(z) = z_{in} \cup z_{out}.$$

Given any zone z' in D', there is a unique zone z in D such that $z' \subseteq \alpha(z)$. The association $z' \mapsto z$ defines a mapping

$$\beta\colon Z(D') \to Z(D).$$

The mappings α and β are illustrated in figure 4.1.



$\alpha(z) = z_{in} \cup z_{out}$, $\beta(z_{in}) = z = \beta(z_{out})$

Figure 4.1

By taking unions of zones, these mappings extend to mappings

$$\alpha\colon R(D) \to R(D'), \quad \beta\colon R(D') \to R(D).$$

These mappings are related as follows. For all regions $r \in R(D)$, $\beta\alpha(r) = r$ and for all regions $r' \in R(D')$, $r' \subseteq \alpha\beta(r')$. (The first of these statements says that β is a left inverse for α and α is a right inverse for β. It follows that α is injective and β is surjective.)

4.2. The process of combining diagrams

Suppose two diagrams D_1 and D_2 are given which do not contain conflicting information. To simplify the process of combination, we first construct the equivalent Venn form of each diagram, $V(D_1)$ and $V(D_2)$ respectively. The combined diagram clearly must contain any contour which appears in either D_1 or D_2, so the first step in combining the diagrams is to construct a Venn diagram whose set of contours is

$$C(D_1) \cup C(D_2).$$

From this underlying Venn diagram, we add diagrammatic elements—shading, spiders, strands and ties—to produce the final combined diagram D. Since D is obtained from each of the diagrams $V(D_1)$ and $V(D_2)$ by adding contours, the 'corresponding region' mappings introduced in the previous section are defined between $V(D_1)$ and D and between $V(D_2)$ and D. These are denoted, respectively, α_1, β_1 and α_2, β_2.

Any shaded zone in the Venn forms $V(D_1)$ or $V(D_2)$ must correspond to a shaded region in D. Hence a zone z of D is shaded if and only if $\beta_1(z) \in Z^*(V(D_1))$ or $\beta_2(z) \in Z^*(V(D_2))$. As a consequence, we have:

$$\bigcup_{z \in Z^*(D)} z = \bigcup_{z \in Z^*(V(D_1))} \alpha_1(z) \cup \bigcup_{z \in Z^*(V(D_2))} \alpha_2(z).$$

This step is illustrated in figure 4.2 (where $D_1 = V(D_1)$ and $D_2 = V(D_2)$).

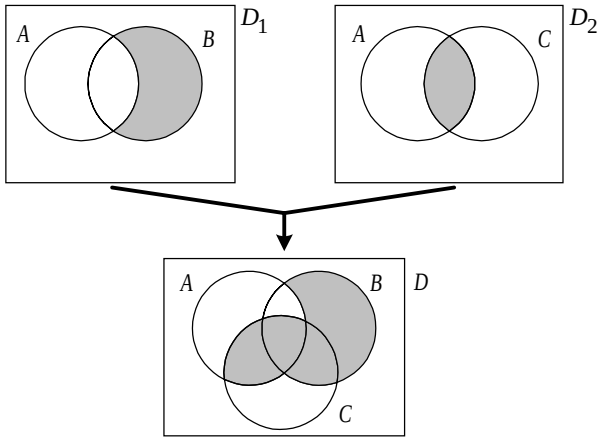


Figure 4.2

Next, we add spiders to D. Since $\eta(s)$ defines the region to which s belongs, intuition suggests that, for each spider, its habitat in D should be the intersection of the corresponding habitats in $V(D_1)$ and $V(D_2)$. This is not quite correct, however, since it does not take account of regions which are known to be empty.

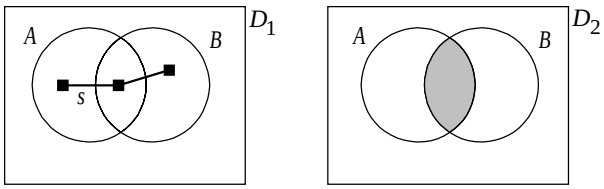


Figure 4.3

This is illustrated in figure 4.3. The habitat of the spider s in the combined diagram must exclude the region $A \cap B$ since, from D_2, this corresponds to an empty set.We define a region of a spider diagram D to be *empty* if it is shaded and contains no spiders. We denote by $E(V(D))$ the set of the empty zones of $V(D)$:

$$E(V(D)) = Z^*(V(D)) \cap \{z \in Z(V(D)) \mid S(z) = \{\}\}$$

For each spider $s \in S(V(D_1)) \cup S(V(D_2))$, we need to define its habitat in D. There are essentially two cases. If s belongs to both diagrams D_1 and D_2 then its habitat in D is the intersection of its habitats in each diagram:

$$s \in S(V(D_1)) \cap S(V(D_2))$$
$$\Rightarrow \quad \eta(s, D) = \alpha_1(\eta(s, (V(D_1))) \cap \alpha_2(\eta(s, V(D_2)))$$

If s belongs to exactly one of the diagrams D_1 and D_2 then its habitat in D is reduced by removing from it the empty zones in the other diagram:

$$s \in S(V(D_1)) - S(V(D_2))$$
$$\Rightarrow \quad \eta(s, D) = \alpha_1(\eta(s, (V(D_1))) - \left(\bigcup_{z \in E(V(D_2))} \alpha_2(z) \right).$$

With these definitions, the composition of the two diagrams in figure 4.3 is given in figure 4.4.

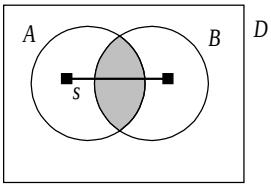


Figure 4.4

Finally, we consider strands and ties. Suppose two spiders are such that each has a foot in a zone z of the combined diagram D. Then z corresponds to zones $z_1 = \beta_1(z)$ and $z_2 = \beta_2(z)$ in $V(D_1)$ and $V(D_2)$, respectively. Again there are several cases to consider.

- If neither diagram $V(D_1)$ nor $V(D_2)$ contains both spiders, then they should be joined by a strand in z. In this case, one spider belongs to $V(D_1)$ and the other belongs to $V(D_2)$, so we have no information concerning their equality or otherwise if they belong to z; hence the spiders should be connected in the most general way.
- If exactly one of the diagrams, $V(D_i)$ say, contains both spiders, then they should be connected in z in the same manner as in z_i.
- If both diagrams contain both spiders then:
  - they are connected by a tie in z if they are joined by a tie in one of the regions z_1, z_2 and a tie or strand in the other region;
  - they are not connected in z if they are not connected one of the regions z_1, z_2 and are either not connected or connected by a strand in the other region;
  - otherwise they are connected by a strand in z.

Example 4.1. Consider the diagrams given in figure 4.5. Since $C(D_1) = C(D_2)$, it follows that each of the correspondence mappings α_1, β_1, α_2, β_2 defined above is the identity mapping. Since there are also no shaded regions, it follows that

$$\eta(t, D) = \eta(t, D_1) \cap \eta(t, D_2).$$

The habitat of the spider s is equal in all three diagrams.

We need to consider separately each zone in D which contains feet of both spiders. For $B - A$, the spiders are connected by a tie in one diagram (D_2) and a strand in the other; hence in the combined diagram, they are connected by a tie. For $A \cap B$, the spiders are separated in one diagram (D_1) and joined by a strand in the other; hence the spiders should be separated in D.

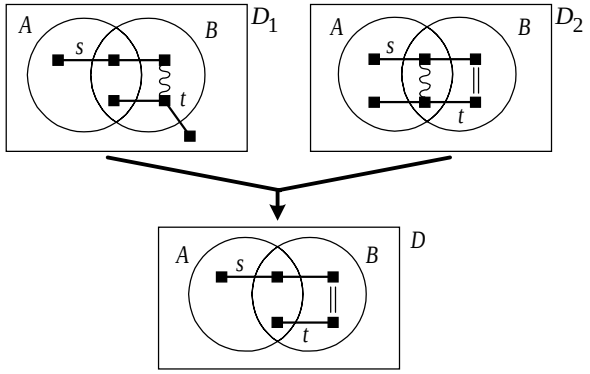


Figure 4.5

Example 4.2. This example illustrates that it is possible for two spiders, s and t, to be separated in z_1 and be joined by a tie in z_2. As before, z_1 and z_2 denote zones in D_1 and D_2, respectively, which correspond to the zone z in the combined diagram D containing feet of both s and t.

Consider the zone $z = A \cap B \cap C$ in the composite diagram D shown in figure 4.6 below. This zone contains feet both of s and of t. The zone z together with the corresponding zones $z_1 = \beta_1(z)$ and $z_2 = \beta_2(z)$ are illustrated with thickened borders. Note that s and t are separated in z_1 but are tied in z_2.

Although it is not possible for the elements corresponding to s and t *both* to belong to $A \cap B \cap C$, this information is not captured in D. Thus it could be argued that it is immaterial how s and t are connected in z. We have chosen to connect their feet with a strand so that each pair of diagrams, D_1, D and D_2, D, is consistent.

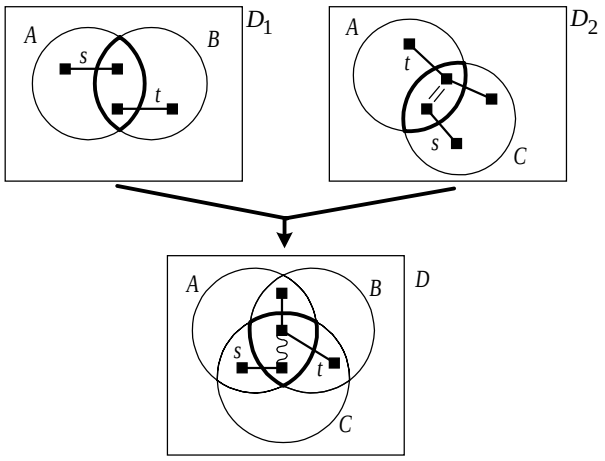


Figure 4.6

5. Conclusion

In this paper, we have given the main syntax and semantics of spider diagrams together with four inference rules and a rule governing the equivalence of the Venn and Euler forms of spider diagrams. In some cases, the inference rules do not give as strong an inference as possible. For example, in figure 5.1, diagram D' can be inferred from diagram D. However, removing spider s from D by rule 2 would require a strand between spiders t and u in the resulting diagram, a weaker result. Some of these stronger inference rules, and the cases in which they apply, have been identified and are discussed in (Howse, Molina, Taylor, 1999). Investigation of these rules is ongoing.

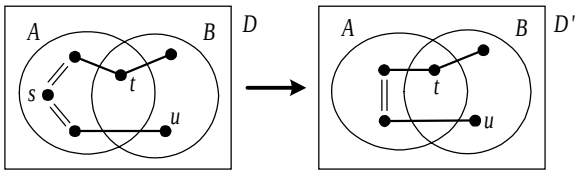


Figure 5.1

We have also given rules for combining spider diagrams to produce a single diagram which retains as much of their combined information as possible. However, we would like to combine diagrams so that no information is lost. Here is one possible solution. We could give figure 5.2 as the combined diagram for the diagrams D_1 and D_2 in example 4.2 (see figure 4.6).

Figure 5.2 is a disjunctive spider diagram. Either the left-hand component holds or the right-hand component holds (see (Shin, 1994) where disjunctive Venn diagrams are used). It is not possible for the elements corresponding to s and t *both* to belong to $A \cap B \cap C$. In D, either s can be in $A \cap B \cap C$ and t not, as in the left-hand component, or vice versa, as in the right-hand component. All the semantic information of D_1 and D_2 is captured in disjunctive diagram D. Investigation of composing spider diagrams (i.e., combining two spider diagrams into one with no loss of information) is ongoing.

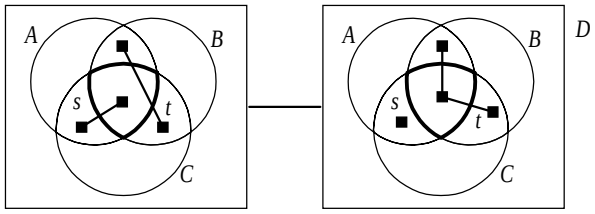


Figure 5.2

Acknowledgements

We gratefully acknowledge Richard Bosworth, Paul Courtney, Yossi Gil, Richard Mitchell, Dan Simpson and the referees for comments on earlier versions of this paper. The authors acknowledge support of the UK EPSRC grant numbers GR/K67304, GR/M02606.

References

Euler, L (1761) *Lettres a Une Princesse d'Allemagne*. Vol. 2, Letters No. 102-108.

Gil, Y., Howse, J., Kent, S. (1999) Formalizing Spider Diagrams, *Proceedings of IEEE Symposium on Visual Languages* (VL99), IEEE Press.

Hammer, E.M. (1995) *Logic and Visual Information*, CSLI Publications.

Howse, J., Molina, F., Taylor, J., (1999) Inference rules for Spider Diagrams, TR UBC99-2, University of Brighton.

Kent, S. (1997) Constraint Diagrams: Visualising Invariants in Object Oriented Models. *Proceedings of OOPSLA 97*

Peirce, C (1933) *Collected Papers*. Vol. 4. Harvard University Press.

Rumbaugh, J., Jacobson, I., Booch, G. (1999) *Unified Modeling Language Reference Manual.* Addison-Wesley

Shin, S-J (1994) *The Logical Status of Diagrams*. CUP.

Venn, J (1880) On the Diagrammatic and Mechanical Representation of Propositions and Reasonings, *Phil. Mag.* 123.

Warmer, J. and Kleppe, A. (1998) *The Object Constraint Language: Precise Modeling with UML*, Addison-Wesley.