A Semiotic-Conceptual Analysis of Euler and Hasse Diagrams

Uta Priss

Zentrum für erfolgreiches Lehren und Lernen, Ostfalia University
www.upriss.org.uk

Abstract. Semiotic-Conceptual Analysis (SCA) considers diagrams (and in general any signs) as consisting of representamens, denotations and interpretations which supports investigating these three components individually and jointly. A core notion for diagram research is "observability" which refers to logically valid statements that can be visually extracted from diagrams. This notion is included into the SCA vocabulary and discussed with respect to Euler and Hasse diagrams.

1 Introduction

Semiotic-Conceptual Analysis (SCA) is inspired by the theory of semiotics of the American philosopher Charles S. Peirce and uses some of his terminology [2]. SCA notions, however, are mathematically defined and thus, in some sense, more abstract than their philosophical counterparts. The purpose of SCA is to investigate questions of what and how something is represented and why certain representations have advantages over others under some circumstances. As an example, SCA is applied to Euler and Hasse diagrams in this short paper. Def. 1 summarises the core definitions of SCA. Further detail cannot be provided here and is presented by Priss [2].

Definition 1. For a set R (called *representamens*), a set D (called *denotations*) and a set I of partial functions $i : R \nrightarrow D$ (called *interpretations*) a *semiotic relation* S is a relation $S \subseteq I \times R \times D$. A triple $(i, r, d) \in S$ with $i(r) = d$ is called a *sign*.
For a semiotic relation S with a tolerance relation $\sim_D$, a tolerance relation $\sim_{D\cap R}$, an equivalence relation $\approx_R$ and a partial function $f : R \nrightarrow D$:
• (i_1, r_1, d_1) and (i_2, r_2, d_2) are *synonyms* $\Longleftrightarrow d_1 \sim_D d_2$;
• (i_1, r_1, d_1) and (i_2, r_2, d_2) are *polysemous* $\Longleftrightarrow r_1 \approx_R r_2$ and $d_1 \sim_D d_2$;
• $(i, r, d) \in S$ is an *icon* $\Longleftrightarrow r \sim_{D\cap R} d$ (i.e., describable by a unary relation)
• $(i, r, d) \in S$ is an *index* $\Longleftrightarrow f(r) = d$ (i.e., describable by a binary relation)
• $(i, r, d) \in S$ is a *symbol* $\Leftrightarrow (i, r, d)$ is neither icon nor index.

Representamens are physical representations of signs. Denotations are meanings of signs and in SCA presented as formalised concepts. Interpretations usually encode a context (time and place) of when a sign is used and possibly further information about a sign producer. A tolerance relation is a mathematical expression of similarity. An example of f would be an algorithm for calculating d from r, instead of a relationship between representamens and denotations that changes with every interpretation.

Several (partial ordering) relations can be defined for signs, for example, *implications* (based on logical implications amongst denotations) and *observations* (derived from compound representamens). For a sign a to be observable from a sign b, the representamen of a has to be derivable from the representamen of b by using some kind of visual algorithm or visual moves. Observability was motivated by Stapleton et al.'s definition [4]. Ideally, only logically true statements should be observed, thus if a sign a is an observation from a sign b then $b \implies a$ should hold. *Translations* amongst signs are morphisms that should preserve meaning in some form. They can lead to *translational loss* or *gain* because, for example, denotations can be modelled using different conceptual models and signs with equivalent denotations can produce different observations.

SCA starts with a qualitative framework (as in Section 2) that roughly characterises how these notions apply to an example. It then continues with more detailed formal analyses, in particular with respect to observations and translations. Both of which are only sketched in this short paper.

2 Applying the SCA Framework to Venn and Euler diagrams

Venn and Euler diagrams are a means for graphically representing sets and their intersections and unions. A more detailed introduction is, for example, provided by Rodgers [3]. Venn diagrams show all possible intersections for a set of sets. Euler diagrams are similar to Venn diagrams but exclude zones which are known to be empty. The following terminology applies to Venn and Euler diagrams in this paper: Venn and Euler diagrams consist of closed *curves* which have *labels*. *Minimal regions* are the smallest areas in a diagram which are surrounded by edges and are not divided further. *Regions* are sets of minimal regions. *Zones* are maximal regions that are within a set of curves and outwith the remaining curves. *Existential import* means that zones must correspond to non-empty sets.

The reason for distinguishing minimal regions and zones is that zones are the smallest set-theoretically meaningful areas in a diagram whereas minimal regions are the smallest visible areas in a diagram. In a *well-formed* Euler diagram, zones correspond to minimal regions. Further conditions for well-formed Euler diagrams are, for example, prohibiting more than 2 curves to cross in a point and curves to intersect themselves. Formalising and characterising well-formed Euler diagrams is not trivial. Flower, Fish & Howse [1] present an algorithm for well-formed Euler diagrams and provide a formalisation as *dual graphs* (with zones as sets of labels and edges between adjacent zones) and *superdual graphs* (with edges between any two sets of labels that differ by a single element).

Applying (a very brief) SCA Framework yields the following initial analysis:

Interpretations: relevant choices for types of interpretations are whether existential import is required (X+) or not (X-) and whether the names of the labels are important (L+) or the labels can be renamed arbitrarily (L-). Other interpretations are possible, for example, non-standard interpretations if someone misreads a diagram or sees a diagram but does not know what it is.

Denotations: a general conceptual model is presented by standard mathematical set theory and anything that is potentially known about it. A more concrete model for Euler diagrams might be a mathematical characterisation of well-formed diagrams.

Representamens: a Venn or Euler diagram is a compound sign. Diagrams can be considered equivalent representamens if a reversible visual translation exists between them. In particular, a translation must preserve existential import conditions in the case of X+ and labels in the case of L+.

Synonymy: one can investigate whether one synonym is "better" than another because it provides more observations or is well-formed or calculate how many synonyms are possible under certain conditions.

Polysemy: one can investigate how diagrams are affected by changing the interpretation, for example, from X+ to X- or by assigning actual elements to the sets.

Icons: depend on personal preferences and historical, cultural background. Presumably, the containment and intersection of circles is considered similar to set operations.

Indices: for example, dual graphs can be considered closely indexically related to Venn and Euler diagrams because they can be easily algorithmically determined.

Translations: Many translations are possible, for example using set-theoretic expressions with labels and $\{\cap, \cup, \subseteq, =\}$; dual or superdual graphs, partially ordered sets of zones, or conjunctive normal forms.

Translational loss and gain: for example, the actual positions and shapes of the curves are not considered relevant and omitted in translations. Different translations invoke different conceptual models which may add background information.

3 Observability of Euler and Hasse diagrams

Stapleton et al. [4] argue that while many expressions may be implied by a set-theoretic expression, only the expression itself is observable from a set-theoretic expression. They conclude that Euler diagrams have a maximal observational advantage over set-theoretic expressions because all logically valid statements can be observed from them. An example is presented by D1 in Fig. 1 which shows that $A \cap B = \emptyset \Longrightarrow C \cap B = \emptyset$. An alternative to Euler diagrams is provided by Hasse diagrams[1] of partially ordered sets, such as D2 in Fig. 1. The filled nodes in D2 correspond to the zones in D1 (including the top node which corresponds to the outer zone). The empty nodes correspond to empty sets. In many cases the Hasse diagram without the empty nodes and ignoring the ordering is isomorphic to the superdual graph of an Euler diagram. They are not isomorphic if the Hasse diagram contains edges between nodes that differ in more than one label, which implies a non-well-formed Euler diagram.

In D2, the highest shared node below a set of nodes presents an intersection (e.g. $A \cap B$). Containment amongst sets corresponds to following upwards edges (e.g. $C \subseteq A$) whereas implications amongst empty sets corresponds to following downwards edges (e.g. $A \cap B = \emptyset \Longrightarrow C \cap B = \emptyset$). One can argue that D2 has an observational advantage over D1 because one can additionally count how many implications are possible. Users, however, will most likely find D1 more intuitive to read, and more iconic for set

[1] SCA normally uses Hasse diagrams of lattices in the sense of Formal Concept Analysis but because of space limitations only partially ordered sets are discussed here.

containment, than D2. D1 also conveys a feeling of understanding of why an implication exists: it is physically impossible in a 2-dimensional space for C to get out of its container A and anywhere near B. Because of the physical constraints, changing D1 so that $(A \setminus C) \cap B = \emptyset$ but $C \cap B \neq \emptyset$ cannot be presented as a well-formed Euler diagram. In D2, however, it would be possible to fill in the bottom node. In that case empty nodes would only indicate that the set at that node is empty. Thus, Hasse diagrams can express any constellation of sets (the same as Venn diagrams with shading) whereas well-formed Euler diagrams with existential import cannot.

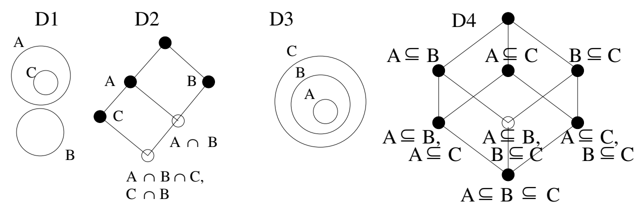

Fig. 1. Euler and Hasse Diagrams

We are proposing that even some mathematical expressions can lead to more than one observation. For example, we would argue that the mathematical expression $A \subseteq B \subseteq C$ allows the same observations as D3 if one knows the convention of abbreviating transitive operations in that manner. D4 displays logical statements and their conjunctions instead of sets and intersections as in D3. It contains an empty node because of $A \subseteq B, B \subseteq C \Longleftrightarrow A \subseteq B \subseteq C$. We would argue that while these two statements are logically equivalent, with respect to observations they are different.

The purpose of SCA is to provide a vocabulary that facilitates, for example, an investigation of why and how mathematically equivalent signs provide different observations. Apart from the basic definitions of SCA, it is not intended to develop new formalisms but, instead, to incorporate existing ones and combine them with a semiotic perspective. SCA is not restricted to mathematical applications because it can also be used for analysing natural or other formal languages [2]. This paper gives rise to questions about further relationships between well-formed Euler diagrams and partially ordered sets (or lattices) which will be addressed in a future publication.

References

1. Flower, J.; Fish, A.; & Howse, J. (2008). *Euler diagram generation.* Journal of Visual Languages & Computing, 19, 6, p. 675-694.
2. Priss, U. (2017). *Semiotic-Conceptual Analysis: A Proposal.* International Journal of General Systems, Vol. 46, 5, p. 569-585.
3. Rodgers, P. (2014). *A survey of Euler diagrams.* Journal of Visual Languages & Computing, 25, 3, p. 134-155.
4. Stapleton, G., Jamnik, M., & Shimojima, A. (2017). *What makes an effective representation of information: a formal account of observational advantages.* Journal of Logic, Language and Information, 26, 2, p. 143-177.