Observing the Truth: Diagrams, Sets and Free Rides

Gem Stapleton

University of Brighton,
Brighton, UK
g.e.stapleton@brighton.ac.uk

There are many different notations that can define relationships between sets, some of which are diagrammatic and others symbolic. Even when a notation is selected, there are choices to be made between semantically equivalent, yet syntactically different, statements. Syntactic choices include variations in both abstract syntax and concrete syntax where graphical and topological properties can differ. Whilst it is clearly important to understand the relative benefits of choices in all senses, the focus here is on the choice of notation and, within that, abstract syntax choices[1].

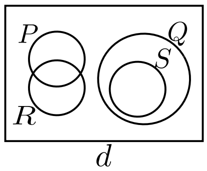

Fig. 1: Illustrating free rides.

This importance of notation choice has been explored previously, through the study of *free rides* [4]. To illustrate, suppose we are given the set-theory statements $P \cap Q = \emptyset$, $R \cap Q = \emptyset$ and $S \subseteq Q$. We can translate these three statements into the Euler diagram in figure 1. As a consequence, the diagram automatically expresses $P \cap S = \emptyset$ and $R \cap S = \emptyset$: we can see these two statements are true 'for free'. By contrast, one cannot read off $P \cap S = \emptyset$ and $R \cap S = \emptyset$ from $P \cap Q = \emptyset$, $R \cap Q = \emptyset$ and $S \subseteq Q$; instead, the first two statements must be inferred. Therefore, the diagram has a perceived advantage over the set-theory statements, with $P \cap S = \emptyset$ and $R \cap S = \emptyset$ being examples of free rides.

Free rides are thought to be a major reason why diagrammatic notations can outperform symbolic notations in reasoning tasks. In particular, free rides indicate why we may wish to choose Euler diagrams over symbolic set-theory as a notation for representing relationships between sets. A key requirement that underpins the theory of free rides is the existence of a translation from one set of statements, such as symbolic statements about sets, into another set of

[1] Readers interested in concrete syntax choices for Euler diagrams or linear diagrams are referred to [1–3].

statements, such as a set of (perhaps a single) Euler diagrams. Such a translation essentially chooses a representation, at the abstract syntax level, of the originally defined statements in the other notation. Here, this idea is generalized to the notion of an *observational advantage* that does not require the existence of a translation between notations. Instead, all that is required is for the two sets of statements to be semantically equivalent.

To illustrate, suppose we have the two set-theory statements $P = Q$ and $P \cap R = \emptyset$. The two Euler diagrams d_1 and d_2 in figure 2 are, between them, semantically equivalent to $P = Q$ and $P \cap R = \emptyset$:

$$\mathcal{D} = \{d_1, d_2\} \equiv_{\models} \mathcal{S} = \{P = Q, Q \cap R = \emptyset\}.$$

We can observe $P \cap R = \emptyset$ from $\mathcal{S}$ because it is written explicitly in $\mathcal{S}$. However, we cannot observe $P \cap R = \emptyset$ from either d_1 and d_2. In this case, from $\mathcal{D}$ we must infer d, which directly expresses $P \cap R = \emptyset$. Therefore, $\mathcal{S}$ has an observational advantage over $\mathcal{D}$. Moreover, $\mathcal{D}$ also has an observational advantage over $\mathcal{S}$, since we can observe $Q \cap R = \emptyset$ from $\mathcal{D}$ but not $\mathcal{S}$.

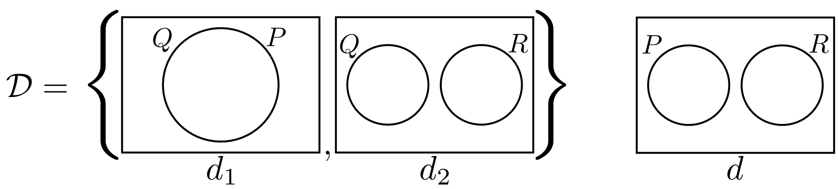

Fig. 2: A more complex example.

Through defining what it means to be able to observe one statement from a set of statements, we can precisely describe the observational advantages of one set of statements over another. In particular, in the case of Euler diagrams and set theory we can prove that Euler diagrams are *observationally complete*. That is given a finite set of set-theory statements, $\mathcal{S}$, there exists a *single* Euler diagram, d, that is semantically equivalent to $\mathcal{S}$ such that any set-theory statement that can be inferred from $\mathcal{S}$ can be observed from d [5][2]. Thus, it is possible to characterise, formally, the observational advantages of Euler diagrams over set-theory. Importantly, the theory of observational advantages allows us to explain why diagrams are advantageous representations of sets in a more general way than free rides.

Acknowledgement This extended abstract is based on research conducted collaboratively with Mateja Jamnik and Atsushi Shimojima [5, 6].

[2] We note that d must explicitly represent all sets occurring in $\mathcal{S}$ and that the inferred set-theory statements must only represent sets occurring in $\mathcal{S}$; we refer the reader to [5, 6] for more details.

References

1. Blake, A., Stapleton, G., Rodgers, P., Howse, J.: The impact of topological and graphical choices on the perception of Euler diagrams. Information Sciences 330, 455–482 (2015)
2. Rodgers, P., , Zhang, L., Purchase, H.: Wellformedness properties in Euler diagrams: Which should be used? IEEE Transactions on Visualization and Computer Graphics 18(7), 1089–1100 (2012)
3. Rodgers, P., Stapleton, G., Chapman, P.: Visualizing sets with linear diagrams. ACM Transactions on Computer-Human Interaction 22, 27:1–27:39 (2015)
4. Shimojima, A.: Semantic Properties of Diagrams and Their Cognitive Potentials. CSLI Publications, Stanford, CA (2015)
5. Stapleton, G., Jamnik, M., Shimojima, A.: What makes an effective representation of information: A formal account of inferential advantages. Journal of Logic, Language and Information (under review)
6. Stapleton, G., Jamnik, M., Shimojima, A.: Effective representation of information: Generalizing free rides. In: Diagrams 2016. Springer (2016)