



Image Schemas and Conceptual Blending in Diagrammatic Reasoning: The Case of Hasse Diagrams

Dimitra Bourou^{1,2}(✉), Marco Schorlemmer^{1,2}, and Enric Plaza¹

¹ Artificial Intelligence Research Institute, IIIA-CSIC,
Bellaterra, Barcelona, Catalonia, Spain
dbourou@iia.csic.es

² Dept. Ciències de la Computació, Universitat Autònoma de Barcelona,
Bellaterra, Barcelona, Catalonia, Spain

Abstract. In this work, we propose a formal, computational model of the sense-making of diagrams by using the theories of image schemas and conceptual blending, stemming from cognitive linguistics. We illustrate our model here for the case of a Hasse diagram, using typed first-order logic to formalise the image schemas and to represent the geometry of a diagram. The latter additionally requires the use of some qualitative spatial reasoning formalisms. We show that, by blending image schemas with the geometrical configuration of a diagram, we can formally describe the way our cognition structures the understanding of, and the reasoning with, diagrams. In addition to a theoretical interest for diagrammatic reasoning, we also briefly discuss the cognitive underpinnings of good practice in diagram design, which are important for fields such as human-computer interaction and data visualization.

Keywords: Diagrammatic reasoning · Image schema · Conceptual blending · Hasse diagram · Formal specification · Sense-making

1 Introduction

Diagrams are often advocated as an effective way to represent and reason with information—‘a picture is worth a thousand words’ goes the aphorism used in English and other languages—and their advantages over purely textual representations have been studied extensively [24, 35, 38, 40].

An important formal contribution to better understand the relative advantage of one representation formalism over another has been the theory of observational advantage put forward by Stapleton et al. [38], which stems from Shimojima’s early work on the efficacy of representations [35]. This theory characterizes the advantage that particular representations have for visualizing certain semantics, because their structure makes some information directly observable. In contrast, other representations require some transformation steps to provide

the same information. Therefore, this advantage of making some information directly observable can be a criterion for the suitability of a representational formalism for a particular reasoning task.

Nonetheless, the theory of Stapleton et al. is based on an abstract characterisation of ‘observation,’ defined in terms of translations to and from an abstract syntax for diagrams, not taking into account its actual geometry, and the active role of the observer in the interpretation [26,40]. Such approaches have been fruitful when applied to the study of reasoning with diagrammatic representations, once the interpretation of their syntax is clearly defined. However, we believe they do not fully capture the way we make sense of geometrical configurations in the first place, and the reason they afford certain interpretations and reasoning tasks, and not others.

In this paper we propose one way to formally and computationally model this sense-making process, drawing from the theory of image schemas and conceptual blending originating in cognitive linguistics [12,13,19,22]. We take ‘sense-making’ to be the process by which humans structure percepts into meaningful constructs [41]. We hereby model the sense-making of diagrams in particular as conceptual blends of image schemas with the geometric configuration that constitutes a diagram. To the best of our knowledge, modeling the sense-making of diagrams in this manner is novel, and we believe it could be of value for shedding further light into the efficacy of diagrammatic representations and their utility for fields pertaining to human-human or human-machine communication.

To illustrate our approach, we will use the particular example of a Hasse diagram (Fig. 1; left). Its geometry comprises a configuration of several points, some of which intersect pairwise with lines. The points are also positioned in specific locations relative to each other. Two of the possible ways for an observer to make sense of, for instance, points e , b and a , and the lines eb and ba that connect them, in Fig. 1 are that:

1. point e with b , and b with a , form two pairs of entities that are linked by lines eb and ba , respectively
2. points e , b , and a are increasing grades on a scale, with direction from e , to b and then towards a .

This understanding of the geometric configuration allows for the emergence of inferences such as the following: since a , b and e represent some quantities such that b is more than e and a is more than b , then a must be more than e . According to Stapleton et al. [38], such interpretations are ‘direct’ in the sense that they require zero transformation steps on the geometric configuration. Moreover, different conclusions can be drawn depending on whether the ‘scale’ or the ‘link’ conceptualisation is at play; the former imbues the sense of quantity, while the latter, the sense of symmetric association. In general, diagrams, taken as geometric configurations, do not bring up a unique way of making sense of them, that is, they do not have a one-to-one mapping with semantics. Therefore, both the geometric configuration and the semantics of a diagram are distinct from each other and from the diagram as we make sense of it.

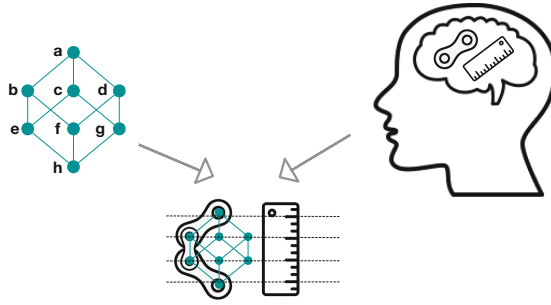


Fig. 1. Visual overview of our model. The geometry of the Hasse diagram (left), and the interpreted diagram (bottom), are distinct. The latter emerges only when schemas (right) are integrated with the geometry, giving rise to the interpreted diagram as a blend (bottom).

Our modeling view of the processes underlying the above scenario is that, although the inferences appear to arise directly from some geometric configurations, we consider that they emerge within a conceptual blend of certain image schemas with these configurations. Image schemas, like *LINK* and *SCALE*, are mental structures acquired by all humans at a very early age [19, 22]. According to the homonymous theory, humans can make sense of stimuli in their environment by unconsciously integrating image schemas with them. The conceptual blending theory examines in detail the principles under which this integration takes place [13].

Hence, in our model, perceptual stimuli (i.e., the geometric configuration of a Hasse diagram) become meaningful because they prompt the conceptual blending of image schemas with them. More precisely, we describe this unconscious process as constituted by the activation of those image schemas that are useful for inference in the current context, and their subsequent integration with the stimuli, by way of establishing suitable correspondences with it. Given these correspondences, a conceptual blend can be constructed, whereby the geometric elements are structured into a coherent, integrated unit through image schemas, and give rise to the diagram as made sense of by the observer (Fig. 1; bottom). We implement our model by formalizing the geometric configuration, the internal structure of the image schemas, the correspondences between the two, and, ultimately, by computing their blend. We further show that our model can account for several direct inferences afforded by Hasse diagrams.

The remainder of this paper is organised as follows: Sect. 2 introduces the key ideas directly related to this work. Section 3 presents our blending model and the inferences resulting therein. Section 4 reviews existing frameworks in diagrammatic reasoning, and existing formalisations of image schemas and conceptual blending. Section 5 explains how our work complements the existing approaches to diagrammatic reasoning, and how it could be developed and applied in the future.

2 Background

In this section we present the theoretical background upon which our computational model is based.

The literature of diagrammatic reasoning has been very valuable for formally studying the informational content, and the efficacy of diagrams for inference. To that end, an one-to-one and total mapping between the syntax (geometric configuration) and the semantics of the diagram is typically assumed [26]. However, as our example in Sect. 1 shows, a certain configuration does not have one possible abstract interpretation. Relatedly, many researchers have suggested that the interpretation of diagrams entails a constructive and imaginative process on the part of the observer [7, 26]. This is in agreement with the claims of enactive cognition.

The enactive cognition paradigm posits that cognition is the sense-making of self-sustaining agents who bring their own original meaning upon their environment [41]. Therefore, in our case study, an enactive cognition approach would posit that no geometric configuration is meaningful in itself, but it prompts the observer to unconsciously structure it into a meaningful diagram by activating suitable frames (in our case, image schemas), and integrating them appropriately with the configuration.

Image schemas fit the role of such frames because they are mental structures formed early in life, constituting structural contours of repeated sensorimotor contingencies, such as CONTAINER, SUPPORT, VERTICALITY and BALANCE [19, 22]. These mental structures are acquired by experiencing (for instance) our bodies being balanced, trying to maintain our balance, supporting an object, etc. Repeated experiences of the same kind lead to the formation of a mental structure reflecting what is invariant among them. This mental structure, called image schema, is a gestalt; it consists of components, in a specific relational structure, which can be systematically integrated with other domains, structure them, and enable conceptual meaning to arise in the mind of the observer. This is related to the phenomenon of mental visualization, i.e., seeing something in our ‘mind’s eye’, such as visualizing a generic chair when hearing the word ‘chair’ [17, 18]. Mental visualization is necessary for inference and prediction, and image schemas have been proposed to enable such visualization [25, pp. 513, 519–520].

Sense-making as integration of image schemas with other domains can be described through the theory of conceptual blending. The central claim of this theory is that a systematic process of building correspondences between different mental spaces underlies diverse instances of sense-making. Mental spaces are “small conceptual packets constructed as we think and talk, for purposes of local understanding and action.” [13, p. 40] They comprise coherent and integrated chunks of information, containing entities, and relations or properties that characterise them. To construct a blend, some pairs of elements from two mental spaces (called input spaces) must be put in correspondence with each other, and merged into the same entity in a new mental space (called blended space). This process allows properties of both corresponding elements to come together in the blend, leading to the emergence of novel structure and thus novel meaning.

3 Approach

As explained above, image schemas can lead to inferences as a result of their internal structure. We capture the structure of each schema formally with a typed first-order logic (FOL) theory, following existing conceptual descriptions of image schemas, or experimental work, when available. The geometry of the diagrams is captured in the same way, additionally using some existing Qualitative Spatial Reasoning (QSR) formalisms to represent topological and geometrical aspects in a manner compatible with human cognition.

We hereby present a case study of sense-making of diagrams, modeling it as a conceptual blending of image schemas with the corresponding geometric configuration. The integration of the image-schematic space with the geometric space follows the principles of conceptual blending, i.e., establishing a cross-space correspondence between these two spaces. Formalising these correspondences allows us to compute the conceptual blend that characterises the diagram as the combination of image schemas with the geometric configuration based on category-theoretic colimits [33].

3.1 Diagrammatic Syntax and Its Formalisation

The geometric configuration of Fig. 1 follows the convention of Hasse diagrams, representing the transitive reduction of a partially ordered set (poset). Typically Hasse diagrams are two-dimensional but this is not a requirement. They consist of edges and vertices, drawn as points and lines. Each point represents one element of the poset. Assuming elements x , y and z of the poset, ordered by the ' $<$ ' relation, then the lines between points are drawn according to the following syntactic rules:

- If $x < y$ then x is shown in a lower position than y in the configuration;
- x and y are connected by a line in the diagram iff $x < y$ or $y < x$, and there is no element z such that $x < z$ and $z < y$;
- lines may intersect with each other, but each one intersects with exactly two points

Therefore, the vertical position of the geometric elements in the configuration of a Hasse diagram has a proper syntactic role, representing the direction of ordering [8]. Consequently, the minimal and the maximal element are always visualised as the lowest and highest points respectively. A poset can be graded, or have ranks, when all maximal chains have the same finite length [37, p. 99]. This intuitively means that there exist groups of incomparable elements that are the same number of steps away from the minimum element. To emphasise this structure, the Hasse diagrams of these posets can be—optionally—drawn with the elements of the same rank as horizontal hyperplanes, as is the case in the Hasse geometric configuration of Fig. 1.

In order to describe the geometric configurations at hand, we draw from some formal systems developed in the QSR literature. In particular, we require

a logical formalism that can capture topological relations and relative positions of the elements in a configuration. Some suitable formalisms are [9] and [16] respectively. Geometric entities can be characterised as being of type point, line or region, and we can describe and reason about their precise topological configuration [9]. The relative position of two-dimensional objects, of any shape, with respect to each other, can also be formalised [16]. This is done by denoting the position (right, right-front, front, left-front, etc.) of an object relative to another.

3.2 A Formal Model of Sense-Making

In this subsection, we present the formalization of the geometric configuration, of the image schemas to be integrated with it, the correspondences between the two, and, finally, their blends. We present here the theories for image schemas LINK, PATH, VERTICALITY, and SCALE [19, 22].¹ In our model, we have described the structure of each image schema and of the geometric configuration of the diagram with a typed FOL theory. Our formalisations were guided by the existing literature, mainly [19, 22]. For the formalisation of the geometric configuration, the aforementioned QSR formalisms are also needed. This way, we can declare instances of geometric types (points and lines) and describe the topological relations and relative position between all pairs of these instances. In Appendix A we show some details of our formalizations. In the remainder of this section, however, we describe our model in a more intuitive and informal manner.² The category theoretical colimit is an abstract operation that can be applied on any kind of mathematical object. In our case, it is applied to logical theories. Having specified some correspondences between elements of these theories, the computation of their colimit yields a new theory where all counterpart elements (types, predicates or functions) are merged into the same element, and the remaining elements and axioms of both theories are also included (see also Appendix A). This mathematical framework is apt to model conceptual blending [33].

LINK. The prototypical LINK schema consists of two distinct linked entities, and a link connecting them. Being linked constrains two entities with respect to each other, i.e., they are bound in some way, due to being in the same relation. More concretely, being linked is a symmetric and irreflexive property. Our formalisation reflects this structure.

PATH. The PATH schema consists of a source, a goal, and a path. The path consists of a series of adjacent locations that connect the source with the goal. By the structure of the schema, it is obvious that, if someone is on a certain

¹ The image schemas, and the correspondences selected, are those that lead to a blend consistent with the set theoretical semantics of the Hasse diagram. This blend serves as a proof of concept, but it is not the only possible blend to model the sense-making of this diagram.

² The complete executable specifications are available in https://drive.google.com/drive/folders/1jcQdJT0qbnAua3uXIgTEW8zV3kF_2R14?usp=sharing and the sense-making of more diagrams is modeled in [5].

location of the path, then they have already traversed all prior locations, and that contiguous locations serially lead from the source to the goal without branching. Therefore, the PATH schema is axiomatised as a total order; a collection of serially neighboring locations with the source and goal as the terminal locations in this series.

VERTICALITY. This schema reflects the structure of our experience of standing upright with our bodies resisting to gravity, or of perceiving upright objects like trees. Thus, VERTICALITY involves a simple distinction between up and down. The VERTICALITY schema comprises an axis and a base, or the ground, as a reference point [34]. Therefore, we model VERTICALITY as a unique vertical axis with its base.

SCALE. The SCALE schema comprises an ordered set of several grades. Unlike VERTICALITY however, it does not imply a particular geometric orientation. SCALE has a cumulative property (if someone has 15 euros, they also have 10); consequently, we formalise it as a total order on grades.

Hasse Configuration. The Hasse configuration of Fig. 1 has eight points (a to h) and twelve lines (ba , ca , etc.). Each line intersects with a pair of points. The logical theory modeling this configuration states the topology and orientation relations among all entities of the configuration with predicates such as *intersects* [9], and *right_back* [16], respectively.

Overall Blend Network. The sense-making of the Hasse configuration is modeled as the conceptual blending of image schemas with it (Fig. 2). Some image schemas form blends among them, and the elements of these blends are subsequently put in correspondence and blended with the geometric configuration.

Specifically, linked entities of LINK are put in correspondence with contiguous locations of the PATH, giving rise to the CHAIN image-schematic blend, comprising a path of linked entities/locations.³ The *linked* and *contiguous* predicates are also put in correspondence (see Appendix A). Subsequently, CHAIN is put in correspondence with the geometric configuration in the following way: The blended entities/locations of the CHAIN are put in correspondence with points that intersect with the same line. The link of the LINK schema is put in correspondence with the line itself. This means that the sequence of points connected by lines in the Hasse configuration (e.g., points h , e , b , and a in Fig. 1), is in correspondence with an instance of the CHAIN image-schematic blend with contiguous entities/locations. Specifically, this image-schematic blend is put in correspondence with the geometry so that the source is the geometrically lowest point, and the goal is the geometrically highest one (*back*, and *front* of all other points, respectively, to use the terminology of [16]).

³ Here, we extend the convention of typesetting image schemas in small caps, to include also our own proposed image-schematic blends, CHAIN and VERTICAL-SCALE. Note that the resulting, blended type ‘entity/location’ now models the structure afforded by CHAIN, as reflected in the union of the axioms of LINK and PATH in the blend. See also Appendix A and [4] for more details on the construction of the blend.

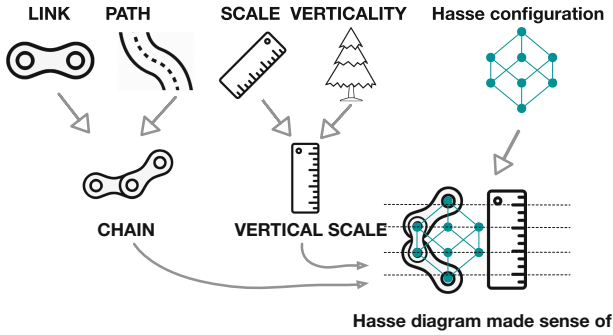


Fig. 2. Image schema blends modeling the sense-making of the Hasse configuration.

Regarding the SCALE and VERTICALITY schemas, they are blended into the VERTICAL-SCALE image-schematic blend. The correspondences between SCALE and VERTICALITY allow the construction of a blend which integrates quantitatively ordered grades of SCALE with vertically ordered marks of VERTICALITY, giving rise to blended levels (dashed horizontal lines in Fig. 2). As for the correspondences of the VERTICAL-SCALE blend with the geometric configuration, the levels are put in correspondence with points, with respect to their geometric ordering. For instance, the level that is immediately above the base is put in correspondence with points *e*, *f*, and *g*, resulting in their integration into the same level in the final blend.

Guided by all the aforementioned correspondences, the VERTICAL-SCALE and the CHAIN image-schematic blends, as well as the geometric configuration, are all blended into a final blend (Fig. 2; bottom right) which has the structure of an ordering that is schematic and geometric at the same time. In other words, this complex network of cross-space correspondences enables the computation of one final blend, whereby the Hasse configuration is structured into a single, coherent gestalt.⁴

Blended Structure and Inferences. The resulting blended space integrates geometric and image-schematic aspects, providing more meaningful structure to the geometric configuration. Within this integrated structure, a variety of inferences emerge. Blending the LINK schema with the geometric configuration of two points intersecting with the same line, gives rise to the interpretation that these points participate in some relationship, and are contingent upon one another in exactly the same way. The two points, together with the line, comprise a single whole; the *LinkSchema*. Blending CHAIN with the geometric configuration structures any set of serially linked shapes into an unitary configuration, i.e., a chain.

⁴ Viewing our model as a network of blends such as LINK-VERTICAL and SCALE-PATH is equally possible and mathematically equivalent to the blends presented, because the colimit operation is associative.

Ultimately, VERTICAL-SCALE and CHAIN, blended with the geometric configuration of the Hasse diagram, yield the Hasse diagram as an observer makes sense of it: as several chains of linked elements, arranged at several levels of generality along a down-up axis. It is important to clarify that the geometric configuration as a graded structure with directionality from point h to point a (reflected in the logical axioms of PATH, VERTICALITY, and SCALE, which jointly appear in the blend—see Appendix A), is neither geometric nor image-schematic. This graded structure emerges in a conceptual blend whose entities and properties are both geometric and image-schematic at the same time. This is only possible because the cognitive structure of the image schemas (reflected in their logical axioms) is blended with the geometric structure, and it yields a variety of inferences within the blend. First, points on the same horizontal plane (e.g., b , c , and d) are construed as being on the same level (dashed horizontal lines in Fig. 2). Second, some points are transitively above others, such as $above(a, e)$, $above(d, h)$ and so on. Notice that this predicate now models the structure afforded by VERTICAL-SCALE. Finally, through this blended $above$ predicate, and the CHAIN, the points that are serially and pairwise linked, form six maximal chains. All of these chains have points a and h , the geometrically uppermost and lowermost points, as their goal and source.

4 Related Work

In the literature of diagrammatic reasoning, it is often posited that the efficacy of diagrams lies in the sharing of structural properties between the geometric configuration and the semantics of a diagram [28,39]. These properties allow observers to make some inferences directly. Therefore, the more the properties of the geometry of a diagram match the properties of a given semantics, the more efficacious the diagram is to represent this semantics [35,38]. A similar framework, called Semiotic-Conceptual Analysis, is proposed by Priss [29]. This framework attempts to explain how meaning is represented with diagrams, language etc. and indeed accounts for various phenomena, such as polysemy, and whether a certain representational format is advantageous for some semantics.

Several research groups have worked on formalizing image schemas and relations among them. Rodriguez and Egenhofer provide a relational algebra based on the CONTAINER and SURFACE schema, used to model, and reason about, spatial relations of objects inside a room [31]. Kuhn formalised image schemas as ontology relations using functional programming, in a relatively abstract and general way [21]. Others concretise their formalisations more, using bigraphs [42], or QSR [15]. Such formalisms imbue topological and other properties into the schemas. The latter also formalises the interrelations of image schemas, as families of logical theories, constructed from combinations of primitive components. Embodied Construction Grammar formalises [2] and implements [6] language understanding by putting in correspondence the components of specific schemas (image schemas, and other kinds of schemas) with phonemes. The framework also incorporates an additional formalism (x -schema) allowing the modeling of inference.

Finally, regarding a formal view of conceptual blending, the blending process has been described through a general, mathematical theory [3, 10, 33]. This is done through amalgams, obtained by generalising the input spaces as much as necessary to find commonalities, and blending parts of them towards a consistent and novel output. This framework, together with image schemas, has been used to interpret an icon by blending a description of the schema with a QSR description of the icon [11]. This approach is a conceptual equivalent of the current computational model. In the same direction, other work formalises the blending of given mental spaces, in order to obtain inference as novelty. Related to our approach, Goguen [14] applied algebraic specifications and their category-theoretic operations for modeling the cognitive understanding of space and time when solving a riddle. Building on this work, Schorlemmer et al. [32] modeled the process of solving a riddle, using blending and typed FOL specifications of image schemas. The interrelations between amalgams, Goguen’s framework, and our current model of blending are discussed in [33]. All aforementioned work contributes valuable, useful formalisations of blending as a creative process.

5 Discussion

The predominant logical approach to diagrammatic reasoning requires a level of abstraction which does not allow for fully taking into account the spatial structure of the geometry, the embodiment of the observer, and the interaction of the two. We believe embodied experiences—whose invariants are crystallised in the form of image schemas—can provide additional insight into the process of understanding and reasoning with a diagram. We present a computational model of this perceptual structuring process, through the integration of image schemas with the geometry of a diagram. To the best of our knowledge, this approach is a novel and valuable theoretical contribution to the diagrammatic reasoning literature. Our work is also directly relevant for human computer interaction and data visualization because, as we explain below, it has the potential to unravel guiding principles towards more intuitive visualizations.

5.1 Diagrammatic Inference with Image-Schematic Blends

Given our modeling of diagram understanding as emerging from conceptual blends of image schemas with geometric configurations, in our Hasse diagram, the facts that: (a) point a is above point h (b) points h , e , b and a form a CHAIN and (c) points b , c and d are on the same level, all can be quickly inferred from the geometric configuration. To make inference (a), for instance, an observer may mentally visualise a physical path of linked locations, starting at location h , extending towards higher locations e and b , up to a , which lies above h and the rest of the locations traversed in the path. This mental visualisation facilitates the inference that $h < a$ directly from the Hasse diagram. Mental visualization is indeed necessary for inference, and image schemas are the mental structures that enable it [25, pp. 513, 519].

Eye-tracking experiments have shown that subjects can make inference (a) for one transitive step without physically manipulating the diagram they were shown [36], and this is interpreted using Shimojima's theory of direct inference [35].

Other inferences made possible through mental visualisation, modeled as the final blended space that integrates the structure of all four described image schemas with the geometry, are (Fig. 2): the transitive ordering of points in terms of their grade on the VERTICAL-SCALE, the inference that the point on the source of the CHAIN is ordered before all others (corresponding semantically to the minimal element), that the point on the goal is after all others (maximal element), and the existence of distinct instances of CHAIN (including all maximal chains).

In the present work, our main goal was to create a cognitively-inspired model of the sense-making of diagrams, not to make claims about human cognition. Consequently, we have not undertaken any psychological experiments. However, our claims about the cognitive structure of Hasse diagrams are consistent with experiments showing that being upright, as opposed to slanted, and explicitly showing levels, makes Hasse diagrams more efficacious (i.e., interpreted faster) [20]. Moreover, other work on diagrammatic reasoning also claims that Hasse diagrams prioritize visualizing the structure of the order they represent, through a vertical organization, and explicit visualization of levels [8]. Levels corresponding to elements with the same rank, i.e., same number of steps away from the minimum element, are geometrically orthogonal to the vertical axis. In fact, this axis is the one intended to be interpreted, and elements of the same rank are indeed not comparable semantically with respect to the ordering.

5.2 Efficacy of Diagrammatic Representations

According to the view of efficacy that we have discussed, some geometric configurations are more efficacious for representing a given semantics, than others. This phenomenon is attributed to some geometric configurations having more similar properties with certain semantics, than others do [35,38]. A Hasse diagram would then be considered very efficacious to represent a partial order, because the geometric arrangement of shapes along a vertical axis has a transitive and asymmetric property, as does a partial order. A diagram whose geometric configuration did not have these properties, or worse, had contradicting ones (e.g., symmetry), would be less efficacious to represent poset semantics. Euler diagrams, for example, have different properties. Representing that $Q \subseteq P$ and $P \cap R = \emptyset$ with the Euler diagram of Fig. 3, makes the inference that $Q \cap R = \emptyset$ directly observable [38]. A Hasse diagram can also represent this scenario, as well any possible constellation of sets. Then why are Hasse diagrams predominantly used to represent posets, and Euler diagrams for set membership?

According to our framework, the higher efficacy of Euler diagrams for set membership and inclusion, and of Hasse diagrams for poset semantics, can be explained as follows: The geometry of the Hasse diagram, comprising shapes that are one above another and grouped in parallel horizontal lines, as well as

the semantics of a poset, are easy to put in correspondence with the VERTICAL-SCALE schema. These correspondences enable constructing blends whereby the aforementioned inferences (transitivity, existence of maximal elements, minimal elements, and maximal chains) emerge. In contrast, the geometric configuration of an Euler diagram, comprising closed curves that are inside one another, is more compatible with the CONTAINER schema; the boundary, inside and outside of the CONTAINER schema can be put into correspondence with the boundary, interior and exterior of closed curves.

The above is also true for the semantics of set membership. Thus, having mentally structured the diagram as comprising physical containers, we can mentally visualize the impossibility of Q being inside P and inside R at the same time. Similarly, the CONTAINER schema also fits the semantics of set membership. This is in agreement with Priss's suggestion that observers find set membership and inclusion easier to read from Euler than Hasse diagrams, because the former enable mentally visualising the impossibility of Q exiting P and approaching R [30]. In fact, it has been proposed that our understanding of abstract set theoretical notions also rests on the same image schematic structures [23].

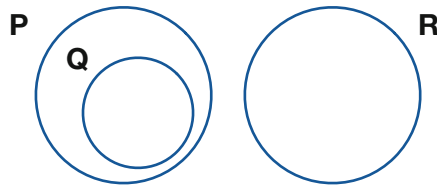


Fig. 3. A simple Euler diagram.

5.3 Conclusions and Future Work

In this paper we have provided a formal framework of sense-making of a diagram, as a creative, active process on the part of the observer, involving the conceptual blending of image schemas with the geometry of the diagram. Most previous computational work making reference to conceptual blending and image schemas only considered creative and problem solving tasks. However, the more fundamental process of sense-making is also a creative and active process that can be explained with conceptual blending. In our work, this view serves as the conceptual foundation.

In future investigations we would like to model a broader range of image schemas, enabling us to examine alternative—including erroneous—interpretations of diagrams. Moreover, we aim to characterise formally what it means for a diagram to be efficacious, in the context of our framework. To achieve both goals, we are currently expanding our framework with the role of the semantics, as well as by including some formal criteria for selecting well-integrated, consistent blends that are useful for reasoning with diagrams. This approach would be a cognitively plausible way to model possible interpretations of diagrams.

The outputs of our model could guide designers of various visual representations, so they can design them with the aim to match the properties of the geometric syntax with the intended meaning. For example, following up with the comparison we made in the previous subsection, if a designer wants to visually represent some ordinal values, a tool based on our framework might recommend the use of a vertical geometric configuration and not a horizontal one. This is because a VERTICAL-SCALE is likely to map to such a configuration and lead to a blend with the intended semantics. In contrast, if a designer wants to represent the notion of belonging in a group, a tool would recommend a configuration with topological containment, because its semantics are similar. Various such recommendations can be made precise thanks to our model and could contribute to new tools directed at designers.

Acknowledgments. The present research was supported by CORPORIS (PID2019-109677RB-I00) funded by Spain’s *Agencia Estatal de Investigación*; by DIVERSIS (201750E064) and DIAGRAFIS (202050E243) funded by CSIC; and by the *Ajuts a grups de recerca consolidats* (2017 SGR 172) funded by *AGAUR/Generalitat de Catalunya*.

A FOL Specifications of Image Schemas and Blends

In this appendix we present some more technical detail of our formalizations and their implementation. The complete formalisation using CASL [1] and HETS [27] can be found in https://drive.google.com/drive/folders/1jcQdJT0qbnAua3uXIgTEW8zV3kF_2R14?usp=sharing.

A.1 The Hasse Geometric Configuration

The specification of the Hasse geometric configuration defines sorts corresponding to geometric entities, and predicates that define their relations in terms of orientation and topology. Below, we present a fragment of this description, for points a , b , e , and h (here constants c_1 , c_2 , c_5 , and c_8) and lines r_{21} and r_{52} , which we discuss as an example throughout the paper.

$$\begin{aligned} &rightBackPP(c_1, c_2) \wedge rightBackPP(c_1, c_5) \wedge leftBackPP(c_5, c_8) \wedge backPP(c_1, c_8) \wedge backPP(c_2, c_5) \\ &intersectLP(r_{21}, c_1) \wedge intersectLP(r_{21}, c_2) \wedge intersectLP(r_{52}, c_2) \wedge intersectLP(r_{52}, c_5) \wedge \dots \end{aligned}$$

A.2 Image Schemas

To model image schemas we followed the principles of algebraic specification, declaring for each schema several types, defining operations and predicates on types, and modeling their behavior by stating axioms over these operations and predicates.

The LINK schema specification

$$\begin{aligned}
\forall s \in \text{LinkSchema} : \text{linked}(\text{anEnt}(s), \text{anotherEnt}(s)) \\
\forall l \in \text{Link} \exists! s \in \text{LinkSchema} : l = \text{link}(s) \\
\forall x, y \in \text{Entity} \forall s : \text{LinkSchema} : \text{linked}(x, y) \Leftrightarrow (\text{anEnt}(s) = x \wedge \text{anotherEnt}(s) = y) \\
\quad \vee (\text{anEnt}(s) = y \wedge \text{anotherEnt}(s) = x) \\
\forall x, y \in \text{Entity} : \text{linked}(x, y) \Leftrightarrow \text{linked}(y, x) \\
\forall x \in \text{Entity} : \neg \text{linked}(x, x)
\end{aligned}$$
The PATH schema specification

$$\begin{aligned}
\forall p \in \text{Path} \exists! s \in \text{SPGschema} : \text{path}(s) = p \\
\forall s \in \text{SPGschema} : \text{inPath}(\text{source}(s), \text{path}(s)) \\
\forall s \in \text{SPGschema} : \text{inPath}(\text{goal}(s), \text{path}(s)) \\
\forall s \in \text{SPGschema} : \text{source}(s) \neq \text{goal}(s) \\
\forall s \in \text{SPGschema} \forall l \in \text{Location} : \neg(\text{isFollowedBy}(l, \text{source}(s), \text{path}(s)) \\
\quad \vee \text{isFollowedBy}(\text{goal}(s), l, \text{path}(s))) \\
\forall s \in \text{SPGschema} \forall l \in \text{Location} : \text{inPath}(l, \text{path}(s)) \wedge l \neq \text{source}(s) \\
\quad \Rightarrow \exists! k \in \text{Location} : \text{inPath}(k, \text{path}(s)) \wedge \text{isFollowedBy}(k, l, \text{path}(s)) \\
\forall s \in \text{SPGschema} \forall l \in \text{Location} : \text{inPath}(l, \text{path}(s)) \wedge l \neq \text{goal}(s) \\
\quad \Rightarrow \exists! m \in \text{Location} : \text{inPath}(m, \text{path}(s)) \wedge \text{isFollowedBy}(l, m, \text{path}(s)) \\
\forall k, l \in \text{Location} \forall p \in \text{Path} : \text{isFollowedBy}(k, l, p) \vee \text{isFollowedBy}(l, k, p) \\
\quad \Rightarrow \text{inPath}(k, p) \wedge \text{inPath}(l, p) \\
\forall s \in \text{SPGschema} \exists! l \in \text{Location} : \text{inPath}(l, \text{path}(s)) \wedge \text{placed}(\text{trajector}(s), l) \\
\forall k, l \in \text{Location}; p \in \text{Path} : \text{contiguous}(k, l) \Leftrightarrow \text{isFollowedBy}(k, l, p) \\
\quad \vee \text{isFollowedBy}(l, k, p)
\end{aligned}$$
The VERTICALITY schema specification

$$\begin{aligned}
\forall s \in \text{VerticalitySchema} : \text{inAxis}(\text{base}(s), \text{axis}(s)) \\
\forall m \in \text{Mark} : \neg \text{above}(m, m) \\
\forall s \in \text{VerticalitySchema}; m \in \text{Mark} : \text{inAxis}(m, \text{axis}(s)) \wedge (m \neq \text{base}(s)) \Rightarrow \text{above}(m, \text{base}(s))
\end{aligned}$$
The SCALE schema specification

$$\begin{aligned}
\forall c \in \text{Scale} \exists! s \in \text{ScaleSchema} : \text{scale}(s) = c \\
\forall c \in \text{Scale}; x, y \in \text{Grade} : \text{inScale}(x, c) \wedge \text{inScale}(y, c) \wedge (x \neq y) \Rightarrow \text{more}(x, y) \vee \text{more}(y, x) \\
\forall c \in \text{Scale}; x, y \in \text{Grade} : \text{more}(x, y) \Rightarrow \exists! c \in \text{ScaleSchema} : \text{inScale}(x, \text{scale}(c)) \\
\quad \wedge \text{inScale}(y, \text{scale}(c)) \\
\forall x, y \in \text{Grade} : \text{less}(x, y) \Leftrightarrow \text{more}(y, x) \\
\forall x, y \in \text{Grade} : \text{more}(x, y) \Leftrightarrow \neg \text{more}(y, x) \\
\forall x \in \text{Grade} : \neg \text{more}(x, x) \\
\forall x, y, z \in \text{Grade} : \text{more}(x, y) \wedge \text{more}(y, z) \Rightarrow \text{more}(x, z)
\end{aligned}$$

A.3 The Hasse Blend

Blends are computed with HETS as category-theoretic colimits, once the correspondences between instances of image schemas and the Hasse configuration have been established by means of spans of morphisms. Below we show a fragment of the blend, namely the constants and axioms of the blend combining CHAIN with the geometric configuration, and some theorems deducible from the blend. As mentioned, this blend includes signatures and axioms of the LINK and PATH schemas, as well as those of the geometric configuration. We denote those coming from LINK in orange, those from PATH in purple, and those from the geometry in teal.

Declaration of constants:

$lsc_{21}, lsc_{31}, lsc_{41}, lsc_{52}, lsc_{53}, lsc_{62}, lsc_{64}, lsc_{73}, lsc_{74}, lsc_{85}, lsc_{86}, lsc_{87} \in LinkSchema$
 $p_1, p_2, p_3, p_4, p_5, p_6 \in Path$ $s_1, s_2, s_3, s_4, s_5, s_6 \in SPGschema$
 $c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8 \in Point$ $r_{21}, r_{31}, r_{41}, r_{52}, r_{53}, r_{62}, r_{64}, r_{73}, r_{74}, r_{85}, r_{86}, r_{87} \in Line$

Axioms:

$rightBackPP(c_1, c_2) \wedge rightBackPP(c_1, c_5) \wedge leftBackPP(c_5, c_8) \wedge backPP(c_1, c_8)$
 $\wedge backPP(c_2, c_5)$
 $intersectLP(r_{21}, c_1) \wedge intersectLP(r_{21}, c_2) \wedge intersectLP(r_{52}, c_2) \wedge intersectLP(r_{52}, c_5) \wedge \dots$

$\forall x, y \in Point \forall s \in LinkSchema : linked(x, y)$
 $\Leftrightarrow (anEnt(s) = x \wedge anotherEnt(s) = y) \vee (anEnt(s) = y \wedge anotherEnt(s) = x)$
 $\forall s \in LinkSchema : linked(anEnt(s), anotherEnt(s))$
 $\forall l \in Line : exists!s \in LinkSchema : link(s) = l$
 $\forall x, y \in Point : linked(x, y) \Leftrightarrow linked(y, x)$
 $\forall x \in Point : \neg linked(x, x)$
 $anEnt(lsc_{21}) = c_2 \wedge anotherEnt(lsc_{21}) = c_1 \wedge link(lsc_{21}) = r_{21}$
 $anEnt(lsc_{52}) = c_5 \wedge anotherEnt(lsc_{52}) = c_2 \wedge link(lsc_{52}) = r_{52} \wedge \dots$
 $\forall s \in SPGschema : source(s) \neq goal(s)$
 $\forall s \in SPGschema \forall l \in Point : \neg(isFollowedBy(l, source(s), path(s))$
 $\quad \vee isFollowedBy(goal(s), l, path(s)))$
 $\forall s \in SPGschema \forall l \in Point : inPath(l, path(s)) \wedge l \neq source(s)$
 $\quad \Rightarrow \exists!k \in Point : inPath(k, path(s)) \wedge isFollowedBy(k, l, path(s))$
 $\forall s \in SPGschema \forall l \in Point : inPath(l, path(s)) \wedge l \neq goal(s)$
 $\quad \Rightarrow \exists!m \in Point : inPath(m, path(s)) \wedge isFollowedBy(l, m, path(s))$
 $\forall k, l \in Point \forall p \in Path : isFollowedBy(k, l, p) \vee isFollowedBy(l, k, p)$
 $\quad \Rightarrow inPath(k, p) \wedge inPath(l, p)$
 $\forall p \in Path \exists!s \in SPGschema : path(s) = p$
 $\forall s \in SPGschema : inPath(source(s), path(s))$
 $\forall s \in SPGschema : inPath(goal(s), path(s))$
 $\forall s \in SPGschema \exists!l \in Point : inPath(l, path(s)) \wedge placed(trajector(s), l)$
 $\forall k, l \in Point \forall p \in Path : linked(k, l) \Leftrightarrow isFollowedBy(k, l, p) \vee isFollowedBy(l, k, p)$

$$\begin{aligned}
& \text{path}(s_1) = p_1 \wedge \text{source}(s_1) = c_8 \wedge \text{goal}(s_1) = c_1 \\
& \text{path}(s_2) = p_2 \wedge \text{source}(s_2) = c_8 \wedge \text{goal}(s_2) = c_1 \wedge \dots \\
& \text{isFollowedBy}(c_2, c_1, p_1) \wedge \text{isFollowedBy}(c_5, c_2, p_1) \wedge \text{isFollowedBy}(c_8, c_5, p_1) \\
& \text{isFollowedBy}(c_3, c_1, p_2) \wedge \text{isFollowedBy}(c_5, c_3, p_2) \wedge \text{isFollowedBy}(c_8, c_5, p_2) \wedge \dots
\end{aligned}$$

Theorems deducible from the blend (about points linked by lines and located in paths):

$$\begin{aligned}
& \text{linked}(c_1, c_2) \wedge \text{linked}(c_2, c_5) \wedge \text{linked}(c_5, c_8) \wedge \dots \\
& \text{inPath}(c_1, p_1) \wedge \text{inPath}(c_2, p_1) \wedge \text{inPath}(c_5, p_1) \wedge \text{inPath}(c_8, p_1) \\
& \text{inPath}(c_1, p_2) \wedge \text{inPath}(c_3, p_2) \wedge \text{inPath}(c_5, p_2) \wedge \text{inPath}(c_8, p_2) \wedge \dots
\end{aligned}$$

References

1. Astesiano, E., et al.: CASL: the common algebraic specification language. *Theor. Comput. Sci.* **286**(2), 153–196 (2002)
2. Bergen, B., Chang, N.: Embodied construction grammar in simulation-based language understanding. In: *Construction Grammars: Cognitive Grounding and Theoretical Extensions*, vol. 3, pp. 147–190. John Benjamins (2005)
3. Bou, F., Plaza, E., Schorlemmer, M.: Amalgams, colimits, and conceptual blending. In: *Concept Invention. CSCS*, pp. 3–29. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-65602-1_1
4. Bourou, D., Schorlemmer, M., Plaza, E.: A cognitively-inspired model for making sense of Hasse diagrams. In: *Proceedings of International Conference of the Catalan Association for Artificial Intelligence* (2021)
5. Bourou, D., Schorlemmer, M., Plaza, E.: Modelling the sense-making of diagrams using image schemas. In: *Proceedings of the Annual Meeting of the Cognitive Science Society* (2021)
6. Bryant, J.E.: Best-fit constructional analysis. Ph.D. thesis, EECS Department, University of California, Berkeley, August 2008. <http://www2.eecs.berkeley.edu/Pubs/TechRpts/2008/EECS-2008-100.html>
7. Cheng, P.C.H., Lowe, R.K., Scaife, M.: Cognitive science approaches to understanding diagrammatic representations. In: Blackwell, A.F. (ed.) *Thinking with Diagrams*, pp. 79–94. Springer, Dordrecht (2001). https://doi.org/10.1007/978-94-017-3524-7_5
8. Demey, L., Smessaert, H.: The relationship between Aristotelian and Hasse diagrams. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) *Diagrams 2014. LNCS (LNAI)*, vol. 8578, pp. 213–227. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-44043-8_23
9. Egenhofer, M.J., Herring, J.R.: Categorizing binary topological relations between regions, lines, and points in geographic databases. Technical report, Department of Surveying Engineering, University of Maine (1991)
10. Eppe, M., Maclean, E., Confalonieri, R., Kutz, O., Schorlemmer, M., Plaza, E., Kühnberger, K.U.: A computational framework for conceptual blending. *Artif. Intell.* **256**, 105–129 (2018)
11. Falomir, Z., Plaza, E.: Towards a model of creative understanding: deconstructing and recreating conceptual blends using image schemas and qualitative spatial descriptors. *Ann. Math. Artif. Intell.* **88**, 457–477 (2019)
12. Fauconnier, G., Turner, M.: Conceptual integration networks. *Cogn. Sci.* **22**(2), 133–187 (1998)
13. Fauconnier, G., Turner, M.: *The Way We Think*. Basic Books, New York (2002)

14. Goguen, J.: Mathematical models of cognitive space and time. In: Reasoning and Cognition, pp. 125–128. Keio University Press (2006)
15. Hedblom, M.M.: Image Schemas and Concept Invention: Cognitive, Logical, and Linguistic Investigations. Springer, Cham (2020). <https://doi.org/10.1007/978-3-030-47329-7>
16. Hernández, D.: Relative representation of spatial knowledge: the 2-D case. In: Mark, D.M., Frank, A.U. (eds.) Cognitive and Linguistic Aspects of Geographic Space, pp. 373–385. Springer, Dordrecht (1991). https://doi.org/10.1007/978-94-011-2606-9_21
17. Jackendoff, R.: Semantics and Cognition, vol. 8. MIT Press, Cambridge (1983)
18. Jackendoff, R.: Consciousness and the Computational Mind. MIT Press, Cambridge (1987)
19. Johnson, M.: The Body in the Mind. University of Chicago Press, Chicago (1987)
20. Körner, C., Albert, D.: Comprehension efficiency of graphically presented ordered sets. In: Kallus, K., Posthumus, N., Jimenez, P. (eds.) Current Psychological Research in Austria. Proceedings of 4th Scientific Conference Austrian Psychological Society, pp. 179–182. Akademische Druck - u. Verla, Graz (2001)
21. Kuhn, W.: An image-schematic account of spatial categories. In: Winter, S., Duckham, M., Kulik, L., Kuipers, B. (eds.) COSIT 2007. LNCS, vol. 4736, pp. 152–168. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-74788-8_10
22. Lakoff, G.: Women, Fire, and Dangerous Things. University of Chicago Press, Chicago (1987)
23. Lakoff, G., Núñez, R.E.: Where mathematics comes from: how the embodied mind brings mathematics into being. *AMC* **10**(12), 720–733 (2000)
24. Larkin, J.H., Simon, H.A.: Why a diagram is (sometimes) worth ten thousand words. *Cogn. Sci.* **11**(1), 65–100 (1987)
25. Mandler, J.M., Cánovas, C.P.: On defining image schemas. *Lang. Cogn.* **6**(4), 510–532 (2014)
26. May, M.: Diagrammatic reasoning and levels of schematization. In: Iconicity. A Fundamental Problem in Semiotics, pp. 175–194. NSU Press, Copenhagen (1999)
27. Mossakowski, T., Maeder, C., Lüttich, K.: The heterogeneous tool set, HETS. In: Grumberg, O., Huth, M. (eds.) TACAS 2007. LNCS, vol. 4424, pp. 519–522. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-71209-1_40
28. Palmer, S.E.: Fundamental aspects of cognitive representation. In: Cognition and Categorization, pp. 259–302. Erlbaum, Hillsdale (1978)
29. Priss, U.: Semiotic-conceptual analysis: a proposal. *Int. J. Gen. Syst.* **46**(5), 569–585 (2017)
30. Priss, U.: A semiotic-conceptual analysis of Euler and Hasse diagrams. In: Pietarinen, A.-V., Chapman, P., Bosveld-de Smet, L., Giardino, V., Corter, J., Linker, S. (eds.) Diagrams 2020. LNCS (LNAI), vol. 12169, pp. 515–519. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-54249-8_47
31. Rodríguez, M.A., Egenhofer, M.J.: A comparison of inferences about containers and surfaces in small-scale and large-scale spaces. *J. Vis. Lang. Comput.* **11**(6), 639–662 (2000)
32. Schorlemmer, M., Confalonieri, R., Plaza, E.: The Yoneda path to the Buddhist monk blend. In: Proceedings of the Joint Ontology Workshops 2016. CEUR Workshop Proceedings, CEUR-WS.org (2016)
33. Schorlemmer, M., Plaza, E.: A uniform model of computational conceptual blending. *Cogn. Syst. Res.* **65**, 118–137 (2021)

34. Serra Borneto, C.: Liegen and stehen in German: a study in horizontality and verticality. In: *Cognitive Linguistics in the Redwoods*, pp. 459–506. *Cog Linguist*, Mouton de Gruyter (1996)
35. Shimojima, A.: On the efficacy of representation. Ph.D. thesis, Indiana University (1996)
36. Shimojima, A., Katagiri, Y.: An eye-tracking study of exploitations of spatial constraints in diagrammatic reasoning. *Cogn. Sci.* **37**(2), 211–254 (2013)
37. Stanley, R.P.: *Enumerative Combinatorics*. Cambridge University Press, Cambridge (2011)
38. Stapleton, G., Jamnik, M., Shimojima, A.: What makes an effective representation of information: a formal account of observational advantages. *J. Logic. Lang. Inf.* **26**(2), 143–177 (2017)
39. Stenning, K.: *Seeing Reason: Image and Language in Learning to Think*. Oxford University Press, Oxford (2002)
40. Stenning, K., Lemon, O.: Aligning logical and psychological perspectives on diagrammatic reasoning. *Artif. Intell. Rev.* **15**(1–2), 29–62 (2001)
41. Varela, F.J.: Organism: a meshwork of selfless selves. In: Tauber, A.I. (ed.) *Organism and the Origins of Self*, pp. 79–107. Springer, Dordrecht (1991). https://doi.org/10.1007/978-94-011-3406-4_5
42. Walton, L., Worboys, M.: An algebraic approach to image schemas for geographic space. In: Hornsby, K.S., Claramunt, C., Denis, M., Ligozat, G. (eds.) *COSIT 2009*. LNCS, vol. 5756, pp. 357–370. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-03832-7_22

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

