

Defining Visual Narratives for Mathematics Declaratively

Max Krieger 

Carnegie Mellon University, USA

maxkrieger@cmu.edu

Wode Ni

Institute for Software Research, Carnegie Mellon University, USA

woden@cs.cmu.edu

Joshua Sunshine 

Institute for Software Research, Carnegie Mellon University, USA

sunshine@cs.cmu.edu

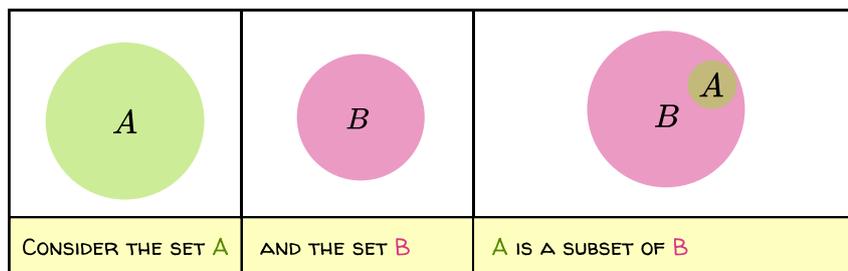
Abstract

Mathematical diagrams are becoming easier to create, with some systems only requiring input of the original mathematical notation. This opens up opportunities to create new kinds of visual media. We define a visual explanatory medium for mathematics, the Visual Mathematical Narrative, derived from the broader medium of Visual Narratives (“comics”). We propose extending the declarative description language of the PENROSE system with constructs that map to the visual primitives we envision in this medium. Finally, we explore potential correspondences between existing textually-oriented mathematical media and visual mathematical narratives.

2012 ACM Subject Classification Software and its engineering → Domain specific languages

Keywords and phrases Comics, Declarative Programming, Visualization, Narrative, Mathematics Education

Acknowledgements Special thanks to Eli Parra and Neil Cohn for their thoughtful suggestions.



```
{  
  Set A, B  
  Panel([A], "Consider the set [A]")  
  Panel([B], "and the set [B]")  
  
  IsSubset(A, B)  
  Panel([A, B], "[A] is a subset of [B]")  
}
```

■ **Figure 1** A simple subset construction comic, described with our proposed patterns.

1 Introduction

Today, there are several systems available to describe and render a mathematical diagram. To describe a diagram using these systems, the end-user typically writes in a domain-specific language that could resemble any level of abstraction from low-level shape primitives and layout parameters to the mathematical content itself. As the description language becomes more high-level, resembling the latter, more qualitative attributes like the former are defined automatically at runtime during the rendering phase.

As computationally generated diagrams become easier to make, new kinds of mathematical media integrating such diagrams can emerge. One such medium is what we refer to as the *visual mathematical narrative*. A visual mathematical narrative is simply a display of several diagrams with a story that binds them together. Such a display can be varied and complex enough that we consider it a medium in its own right.

1.1 Visual Narratives

To further characterize visual mathematical narratives, and the constructs they may contain, we look to the broader space of *visual narratives*. Visual narratives famously appear as some of the oldest known human-created media, from the Lascaux Caves to Egyptian tombs [9]. They also appear in the present as *comics* in the United States, or *manga* in Japan. This broader medium has been studied more extensively than its mathematical variant, and much prior work on it has informed the math-specific patterns we define below.

Why are visual mathematical narratives an important medium to explore? We identify three of its qualities that may augment the capabilities existing mathematical media, from proofs to lectures to textual explanations. These qualities are derived from prior literature on visual narratives. We further hypothesize that maximizing these three qualities make any medium more amenable to human understanding.

1.1.1 Decomposability

If a concept or process can be broken up into parts, it is *decomposable*. According to Clark and Mayer’s “Segmenting Principle” a complex lesson can be simplified by breaking its constituent concepts and interactions down into “manageable segments...presented one at a time” [7]. This is a relevant principle for constructing mathematical narratives since all narratives are fundamentally about representations in sequence.

1.1.2 Local Structure

Arnheim writes that “to see means to see in relation” [5]. It follows that a useful visual representation (or “depiction” [20]) has enough structure that a reader’s perceptual system can internalize how objects relate to each other within it. Larkin and Simon observe in their study that diagrams afford “localities”; places where information can be “accessed and processed simultaneously” in an efficient way. Further, localities that are connected “adjacent[ly]” aid inference and problem-solving ability by conveying relations on the 2-d plane [22]. Kosslyn et al argue that while propositional (“descriptive”) representations “make explicit and accessible semantic interpretations”, depictive representations “make explicit and accessible all... spatial relations” [20]. Both representational types have strengths and weaknesses, but media that express concepts in only one may obscure an entire class of relations.

1.1.3 Global Structure

Local structure allows for comparison *within* representations, but comparison *between* representations can facilitate understanding as well [18]. A well-defined *global structure* is amenable to *deixis* (pointing), *saccades* (glancing), and self-guided pacing forwards and backwards within a narrative, which are cognitive-physiological mechanisms that we will see can aid comprehension. In other words, high-level structure provides an interface for finding relations between the frames of a narrative.

Note that we define a narrative as having global structure only if it displays local structures adjacently in space so that they can be readily compared. Therefore, animated videos lack such a quality, since they display frames one at a time.

1.2 The Description Problem

We now return to the scope of existing diagram-creation systems. Though these systems can produce individual diagrams, the problem of describing an *inter-diagrammatic narrative* using high-level descriptions has largely gone unaddressed. A particular problem many of these systems face is *visual continuity*: if objects reappear in multiple diagrams, the system should make these objects easily recognizable across diagrams.

Though visual continuity may appear to concern only the rendering side of diagram creation, it is also relevant to the design of the description language. What defines the boundary between separate diagrams if some content is reused between them? There are opportunities for additional high-level narrative constructs that build on top of this continuity principle; for instance, affordances that guide the reader's attention to a particular object or transformation in a diagram.

1.3 Penrose

PENROSE is a mathematical diagram-creation system that uses a high-level description language called SUBSTANCE. SUBSTANCE aims to resemble the original mathematical notation that corresponds to the objects being visualized. This resemblance comes at the expense of omitting visual descriptions, such as shape primitives, colors, sizes, and layout. These visual descriptions are written at another abstraction layer through the STYLE language. The benefit of the SUBSTANCE-STYLE distinction is that STYLE programs are reusable and interchangeable for the same mathematical content: a SUBSTANCE program of nested sets can be represented through a venn diagram or a tree STYLE at the user's discretion.

PENROSE ultimately enables creation of ad-hoc visual constructions to explain a singular frame, or *panel*, of mathematics with little stylistic effort required of the user [28]. However, when defining several diagrams in sequence, *semantically similar* descriptions in PENROSE may not necessarily render *visually similar* diagrams due to the system's sampling-based approach. Visual continuity is therefore lost, as the system assumes diagrams only exist as isolated panels.

The central assumption of the PENROSE design is that diagrams can be universally described through a formal stylistic grammar corresponding to mathematical objects. Similarly, this paper's central assumption is that diagrammatic narratives can be described through a formal *narrative grammar*. In this paper, we suggest how one might define a visual mathematical narrative using a declarative description language based on the PENROSE system. For such a language, we define several high-level visual constructs, or *patterns*, that might appear in the medium. We also explore the design implications of the medium:

how might a visual mathematical narrative correspond to a textual one? Our suggestions constitute a synthesis of prior literature on visual narratives and mathematics education.

2 Extending Related Work

We draw upon several existing designs and principles from visual narratives, declarative languages for such, and mathematical narratives.

Several language designs inspired by drama scripts have significantly influenced our design: *THAPL*, “A Theatrical DSL” [12], and the esoteric *SPL*, “The Shakespeare Programming Language” [16]. We paid especially close attention to a declarative construct both works call the *dramatis personae*, which defines the “variables” or “objects” that are in scope for a particular scene. Consider mathematical objects to be the *dramatis personae* of this work. In addition, *ComicsML* [27] and a comprehensive derivative [3] provide a foundational structure for a declarative semantics of comics.

Sloan’s *Visual Narrative Styles in Mathematics and Computer Science* [34] provides a comprehensive set of cognitive design principles for creating visual mathematical narratives. Bach et al’s “Graph Comics” work derives a comprehensive set of concrete patterns for creating visual narratives around graphs [6].

Cohn’s work around the *Visual Narrative Grammar* theory [9] [10] [11] [8] applies grammatical and cognitive methods to the visual narrative medium, which have informed how we derived our patterns and syntax.

Lastly, Martens et al’s *Discourse-driven Comic Generation* [25] formalizes many of Cohn’s and McCloud’s principles into a grammar of panel transitions, and implements this grammar as a system for generating abstract visual narratives. This work recognizes a need to encode visual and narrative continuity in the narrative’s semantics, a concept the authors refer to as *relatedness* (via Saraceni, [32]).

3 Patterns

We are interested in enumerating the high-level inter-diagram relationships (“*patterns*”) that appear in mathematical visual narratives. The patterns adhere to our principle of visual continuity (which Latour calls *optical consistency* [23], and Saraceni, *relatedness* [32]). We describe the principle through the question: *what are the valid state transitions for these objects which preserve visual continuity between frames?*

These patterns are derived from cognitive design principles [34] and motifs in existing visual mathematical media. Though we present concrete ways of visually expressing these patterns, they may manifest through other visual representations, or *STYLES*, depending on the design of the host medium.

3.1 Conjunction

Cohn observes that different panel displays in sequential media provide different “windows” on the same scene. To model this, he borrows an idea from linguistics: *conjunction*, a pattern that conjoins narrative primitives into a hierarchy [10]. This mechanic enables narrative flexibility, separating underlying content from its sequential representation. It is sensible, then, to provide the ability to recombine objects into arbitrary panel delineations (while maintaining their optical consistency). Authors can first introduce objects separately to the reader, and compose the ensemble as a conjunction panel. Or, objects can reappear in different environments, with further mutations applied independently of each other.

The Conjunction pattern demonstrates an important distinction between what we refer to as a *scene* and a *panel*. A scene is the invisible scope that defines the objects that can be depicted in a space; the underlying content. Scenes are delimited in our syntax as curly brackets, { }, within which are the objects in the scene. Scenes are also implicitly composed through newlines: if one line defines a **Set A**, then the next line defines a **Set B**, the resulting (desugared) scene hierarchy is:

```
{
  Set A
  {
    Set B
  }
}
```

Panels, on the other hand, are the concrete visual units which may or may not show the entire scene. These are the frames the reader sees. We represent panels as function calls: a panel showing just **Set A** would be called within its scene `Panel([A])`. A panel showing both A and B, along with a caption, could be called within their composed scene (as well as a caption) as `Panel([A, B], "Here are [A] and [B]")`.

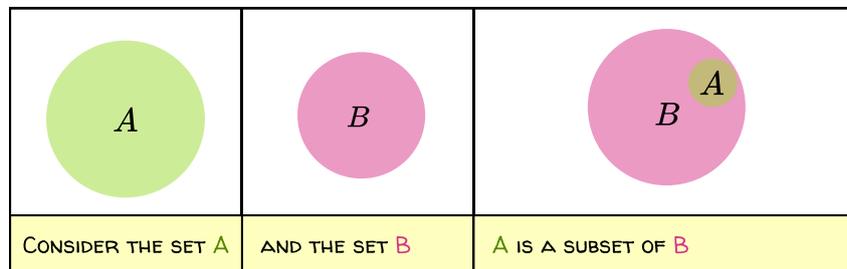
By separating the scene’s object definitions (dramatis personae) from their inclusion in panel definitions, a conjunction pattern emerges in our language. Authors can *selectively include* references to objects defined in the background. For a usage example, compare the scenes with two sets in Figures 2a and 2b.

3.2 Multiples

Tufte dedicates an entire chapter in his 1997 work to the concept of “multiples” [37, p. 105]. A “multiples” display shows multiple images riffing off of the same underlying idea. This allows for comparison, pattern recognition, and reinforcement; processes that are essential for internalizing the nature of an object. Perceptual studies have found that comparisons between different pairs of objects can lead to different outcomes and underlying conceptions [13]. In particular, Kellman et al found that comparative lineups of mathematical representations with varying structures improved students’ speed in solving both visual and symbolic problems [18].

In mathematics literature, a plurality of concrete examples is often foregone in favor of presenting “arbitrary” instances that generalize to any member of a set. What if generality and concreteness could be reconciled by making it easier to provide a surplus of examples? A large number of handpicked examples can potentially show diverse slices of an otherwise vast space. In particular, displaying an array of examples allows for “special cases” to emerge, exposing the reader to salient points in the concept space. However, producing these examples has historically been painstaking work. The PENROSE system constructs diagrams through an inherently sampling-based process, making this task trivial.

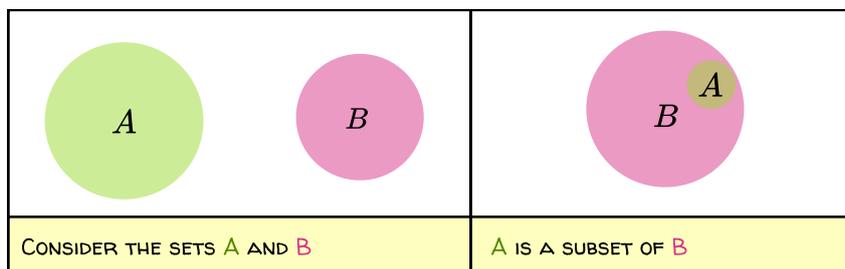
We extend the `Panel` function defined previously with a new function, `SeedPanel`, which takes in a seed integer as well as the existing arguments. PENROSE samples diagrams pseudo-randomly from a seed, so one approach to ensure we get a diverse Multiples display is to pin these samplings explicitly by specifying a seed. In Figure 3, we take a figure from *Dynamics* [1] and interpret it as such a display.



```
{
  Set A, B
  Panel([A], "Consider the set [A]")
  Panel([B], "and the set [B]")

  IsSubset(A, B)
  Panel([A, B], "[A] is a subset of [B]")
}
```

(a) Two sets are depicted separately before their subset predicate.

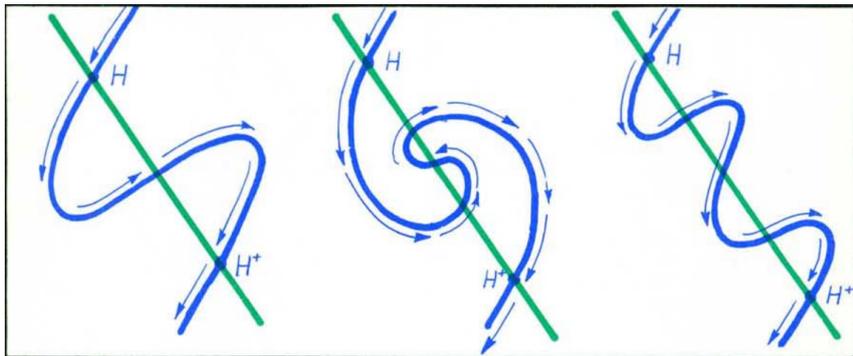


```
{
  Set A, B
  Panel([A, B], "Consider the sets [A] and [B]")

  IsSubset(A, B)
  Panel([A, B], "[A] is a subset of [B]")
}
```

(b) Two sets are depicted together before their subset predicate: a conjoined variation of Figure 2a.

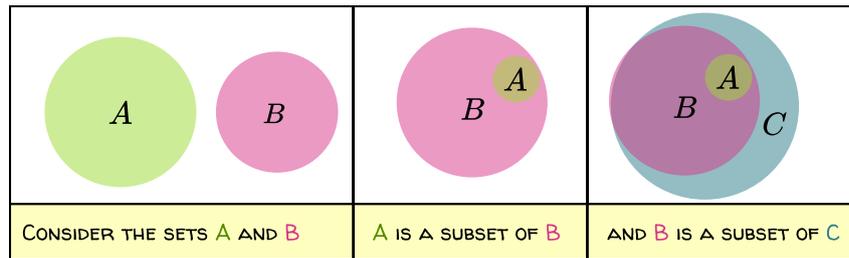
■ **Figure 2** A mockup of a set theory narrative and its syntax for the conjunction pattern.



```
{
  Trajectory T
  Inset I
  Outset O
  Point H, H+ ∈ I
  Intersection(I, O)

  SeedPanel(5340958503, [T, I, O, H, H+])
  SeedPanel(3536319804, [T, I, O, H, H+])
  SeedPanel(4501177698, [T, I, O, H, H+])
}
```

■ **Figure 3** Figure 13.5.3 from *Dynamics* [1, p. 401] which “shows three possible connections for the outset curves”, interpreted as a panel construction in our language with arbitrary seeds.



```

{
  Set A, B
  Panel([A, B], "Consider the sets [A] and [B]")

  IsSubset(A, B)
  Panel([A, B], "[A] is a subset of [B]")

  Set C
  IsSubset(B, C)
  Panel([A,B,C], "and [B] is a subset of [C]")
}

```

■ **Figure 4** Extending Figure 2b, we depict a three-layered construction.

3.3 Predicates

Predicate functions are fundamental to PENROSE’s expressive ability, and mathematical descriptions in general. A *Predicate* defines specific properties that an object can have, both mathematically and, in PENROSE’s case, perceptually. It constrains highly general, faceless objects into acutely specific instances through boolean rules.

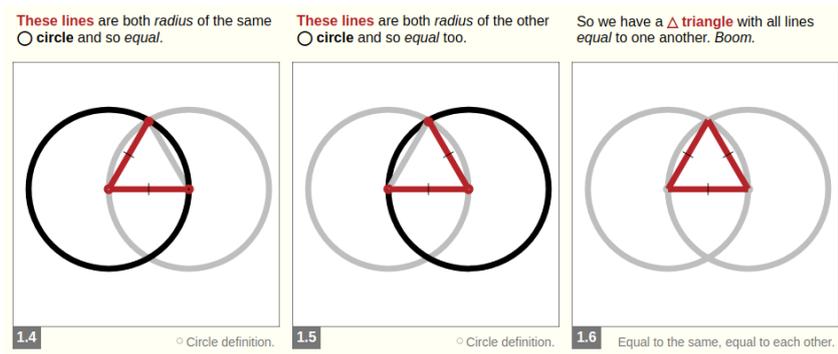
In many ways, Predicates are the refined, non-stochastic variants of Multiples. By iteratively applying Predicates, an author can construct a highly complex object through narrative. Or, an object can assume several parallel Predicate forms, exposing alternatives. This iterative approach naturally decomposes the object into its constituent parts. Each predicate application step is amenable to visual comparison with the next, a process whose cognitive benefits have been described in the Multiples section.

Extending the earlier narrative in the Conjunction section (Figure 2b), we emphasize the repeated application of PENROSE’s `IsSubset` predicate in Figure 4. This effectively decomposes a three-layered set construction into three panels, with subset relationships defined one at a time.

3.4 Foci

Attention as a finite resource is an important finding in cognitive science that is highly applicable to visual narrative design. Recent ideas such as biased competition theory suggest that objects in a scene “compete” in the brain for attention, creating a zero-sum game of limited processing ability [33]. As Tufte remarks, “when *everything* is emphasized, *nothing* is emphasized” [37, p. 74]. Using visual cues, we can direct the learner’s attention to important features of a panel and mute distracting ones.

There are several visual mechanics we have found in existing mathematical media that attempt to focus the viewer’s attention. Anderson et al observe lecturers’ extensive use of



■ **Figure 5** Eli Parra’s comics rendition of Euclid’s *Elements* [29]. The primary foci of the panels, lines, are highlighted in red. Secondary foci, circles, are highlighted in black.

“attentional marks” on their slides that add a secondary layer on top of existing figures to non-destructively call attention to pieces of content in a scene [4]. Greiffenhagen’s video analysis of blackboard lectures finds use of deixis, pointing at “these” and “those” pieces of notational objects, to guide the student’s attention through the space of an otherwise complex scene [14].

In other media, authors mutate the visual properties of existing objects, rather than introducing new properties. This ranges from a simple fill color change to an opacity adjustment or layering change. For example, Figure 5 focuses on different lines and circles in each panel to emphasize different features of the same scene; namely, the equality of circle radii and triangle edges. Bach et al denote focal elements as “main characters”, which are distinguished by changing visual properties, clipping minor characters out of the panel, or explicit labelling [6].

The *DataToon* system infers several of these focal mechanisms automatically to interpolate between frames and maintain continuity. In particular, *DataToon* recognizes if an author intends to “zoom” on an object, show a particular period of time, or “filter” out classes of data [19].

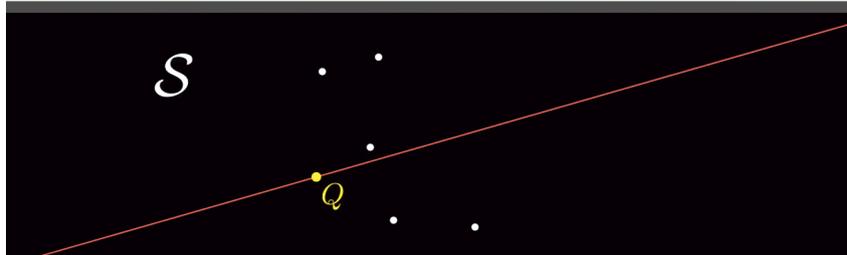
It is important to note that Foci don’t just appear on existing objects. Foci can be applied to objects that are introduced into a scene, and their bright color, figure-ground distinction, and opacity choices draw attention and mute existing objects that were present in prior instances.

Foci can also draw attention to correspondences between disparate structures. Again, there are no changes to the content at hand needed when directing the reader’s attention. Corresponding colors and other attributes create connections in a display when there otherwise would be none. They can also clarify the meaning of conventional representations, like notation and text, through juxtaposition and correspondence with visual representations (Figure 6).

When the objects depicted have features at several scales, it is useful to focus on particular localities through “zooming” This appears frequently in representations of continuous structures, such as real-valued functions, where smaller intervals convey information that informs the reader of properties in larger intervals, and vice versa (e.g. see figures 7.1.1 and 7.1.2 in *Dynamics* [1, p. 235-236]). Local-global shifts can provide a holistic view of multiscale objects in a visual setting.

The encoding of foci is not much different from predicates, as foci are essentially functions being applied to objects or other predicates. They could be as generic and STYLE-specific as a simple invocation to `Focus(A)` on a particular object A, or as concrete as a `Zoom([A,B,C])`

Problem 2. Let S be a finite set of at least two points in the plane. Assume that no three points of S are collinear. A *windmill* is a process that starts with a line ℓ going through a single point $P \in S$. The line rotates clockwise about the *pivot* P until the first time that the line meets some other point belonging to S . This point, Q , takes over as the new pivot, and the line now rotates clockwise about Q , until it next meets a point of S . This process continues indefinitely. Show that we can choose a point P in S and a line ℓ going through P such that the resulting windmill uses each point of S as a pivot infinitely many times.



■ **Figure 6** Several Foci at time 2:20 in a *3blue1brown* video [31]. Notation and text is highlighted to show what is being depicted beneath it. The object Q in the depiction, a focal element in the description above it, has been tinted yellow for its emphasis.

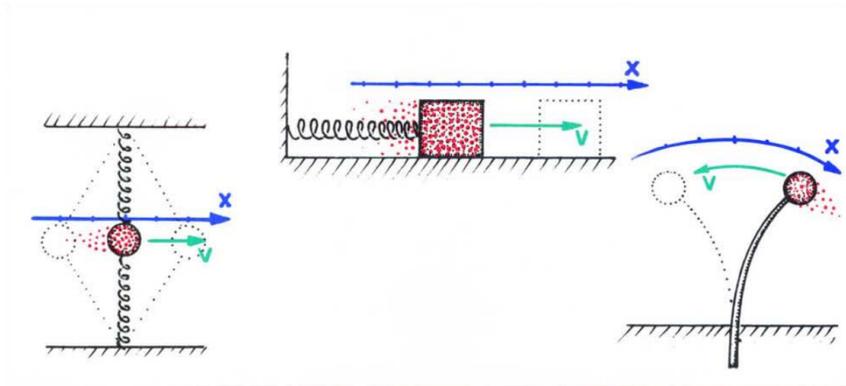
call on multiple objects A , B , C . Different objects in a scene can be denoted as different foci through isolated sub-scene declarations:

```
{
  Set A
  Set B
  IsSubset(B, A)
  {
    Focus(A)
    Panel([A,B], "Here is [A]")
  }
  {
    Focus(B)
    Panel([A,B], "and here is [B]")
  }
}
```

3.5 Emanata

When given the chance, effective mathematical instruction expresses itself through qualitative analogies and gestures that suggest lifelike, embodied processes. Lakoff and Núñez suggest that the process of explaining and doing mathematics is, in fact, intrinsically embodied [21]. An analysis of a blackboard lesson on the Dutch Book argument finds moments where the lecturer compares two quantities through elaborate gestures and crouching [14]. Qualitative properties are also in animations like *3blue1brown*'s linear algebra series, where the instructor describes an embodied mental model for linear transformations. He conveys that matrix multiplication “gives the feeling of squishing and morphing the space itself”, a piece of intuition that reoccurs in his series [30]. Indeed, former Disney animators famously described their craft as a process to create an “Illusion of Life” [36].

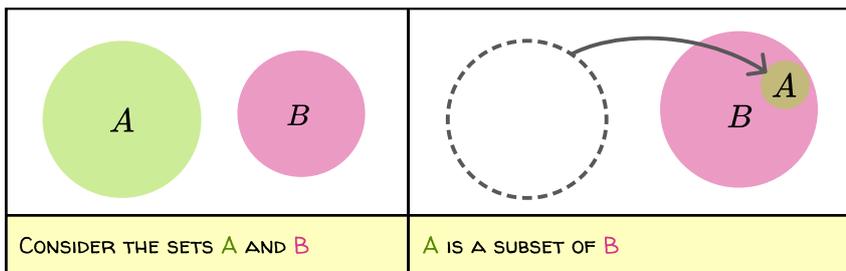
An author could enhance the reader’s understanding of mathematical actions, then, if given affordances that supplement the material with embodied qualitative annotations, known as “emanata” among comics theorists [38].



■ **Figure 7** Figure 2.3.2 of *Dynamics* [1, p. 72], showing several emanata in use to depict the motion of some mechanical oscillators. Several “visual morphological relations” that Cohn [11] enumerates appear. The ball and box have outlined shadows (“suppletion”) that suggest bounds to their oscillation, not to mention the arrows themselves that trace their paths. There are also red particles trailing the bodies (“affixation”) that suggest rapid motion.

What are the visual primitives that constitute emanata? The answers are highly domain-specific, depending on the *STYLE* of the visualizations being depicted. Cohn enumerates several recurring “visual morphological relations” in comics [11], some of which appear in Figure 7.

One common representation for the emanatum of motion, or a moving object, is through “sliced frames”. Translations and transformations can be represented through a sampled interpolation, with notable moments captured in particular panels. See Figures 3, 4, 9, and 10 in *To Dissect A Mockingbird* [17] for examples of this in an animated lambda calculus notation. As mentioned, the DataToon system affords a sliced frames effect through automatic interpolation [19].



```
{
  Set A, B
  Panel([A, B], "Consider the sets [A] and [B]")
  SuppletePanel({IsSubset(A,B)}, [A,B], "[A] is a subset of [B]")
}
```

■ **Figure 8** Modifying Figure 2b, we decorate the *IsSubset* predicate with a suppletion: In the second panel, set *A* is traced as if it is moving inside of *B* from where it originally was in the first panel.

Given their post-hoc nature, annotating operations that already occurred, we envision

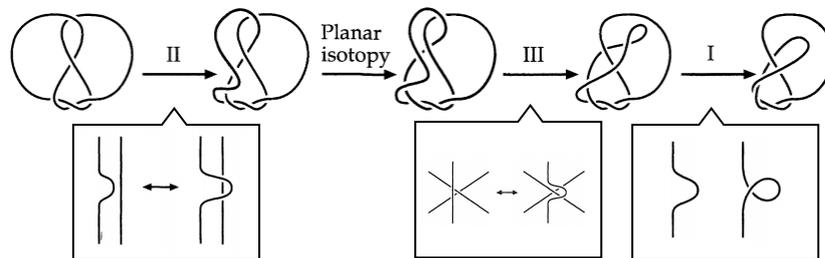


emanata as decorators on predicates. A simple example is the suppletion emanatum: in Figure 8, the `IsSubset` predicate is brought to life by decorating it as a suppletion.

It is worth noting that while Foci and Emanata appear to be similar, they play distinct, complementary roles in creating a smooth visual narrative. While both guide the reader’s attention, Emanata supplement mathematical “verbs” with qualitative properties, while Foci take the burden of defining the frame of local focus and direct the motion of the eye between mathematical “nouns”.

3.6 Nesting

So far we have discussed patterns that describe inter-panel relationships. The visual narrative medium also has a simple, powerful mechanic: placing panels *within* panels. Sousanis remarks on the medium: “[F]rom the forking paths, tangential (and parenthetical), layered and overlapping, intersecting, moments nested within moments, comics can hold the unflat ways in which thought unfolds.” [35] Tufte writes of “visual confections”: “an assembly of many visual events, selected... from various Streams of Story, then brought together and juxtaposed on the still flatland of paper.” [37, p. 121] These two descriptions suggest that there can be nonlinearity and hierarchy in visual narratives, brought out by composition, or nesting, of panels. Indeed, the visual narrative isn’t the only medium where nesting occurs. “Nested Art” can be found in paintings (sometimes referred to as “mise en abymes”), literature, poetry, and film [24]. Cohn argues that visually nested panels, “recursive carriers”, are powerful expressive features of the comics medium analogous to linguistic recursion in human language [8].



■ **Figure 9** A modified version of figures 1.22, 1.23, 1.24, 1.26 from *The Knot Book* [2]. Reidemeister moves (originally referenced as Roman numerals) have been annotated with their respective definitions as popout “thought bubbles”.

Textual mathematical proofs are structured with dense relations. Claims are backed by prior evidence in the form of numbered and named theorems and lemmas. If the “scope” or “proof environment”, is not sufficiently implanted in the reader’s mind, the evidence being cited is simply a dangling reference. In visual narrative variations of these proofs, we hypothesize that due to the inherent nonlinearity of visual media, providing the evidence inline for a juxtaposed proof step as a nested parenthetical, perhaps a “thought bubble”, would not significantly interrupt the reading experience. As one lecturer walks through a proof, Greiffenhagen observes, he uses a sliding blackboard’s panelling to store intermediate lemmas, and gestures towards them as needed [15]. A proof step citing a significant theorem may require a statement to the likes of, “Recall that, in lemma 4.2...” This could manifest visually as a picture-in-picture, a popout element that can easily be skipped over for readers familiar with the juxtaposition, as seen in Figure 9.

Nested panels can also provide a sense of parallelism. Temporally, a panel showing the “final product” early on provides anticipation of what’s to come. Or, several nested panels in a row can decompose a complex process summarized by the parent panel. Spatially, a nested panel that shows a local slice of a global representation provides a simultaneous perspective. McCloud calls the phenomenon of “observing the parts but perceiving the whole” “closure” [26, pg. 63]. A reader is capable of interpolating disparate panels, demarcated by borders (forming closure), into a coherent visual narrative.

The implementation of nesting is nearly as straightforward as any inter-panel relationship. The nesting function is a binary operation between the parent and child panels. One could bind the child and parent panels, e.g. p_1 and p_2 , anchor p_1 to p_2 ’s Set A and then present the panels after binding by calling a function: `NestPanels(p1, p2, A)`. The spatial parallelism described is a special usecase of the “Zoom” function in the Foci pattern. Temporal parallelism, showing interpolated processes, can be represented similarly to the motion emanatum described in the Emanata section: several sliced frames in sequence within a parent frame.

4 Representing Formalisms

If visual narratives are just another way of representing mathematical formalisms, could there be a mapping between them and textual mathematical narratives, such as textbooks and proofs? The following is a speculative, qualitative look at potential correspondences between the two media.

4.1 The Gutter as Logical Implication

A key component of McCloud’s comics theory is the role of the “gutter”. The gutter is the blank space between a comic’s panels, and it can represent any temporal, spatial, or conceptual boundary that glues the narrative together [26, c. 3]. We discussed the Nesting pattern in reference to this idea: it forms a gutter between two panels being presented hierarchically or in parallel. This is analogous to mechanics seen in textual mathematical narratives, for instance, figures, theorems, and lemmas. The typical gutter that forms between adjacent panels, however, can take an even more fundamental function: logical implication. From one panel to the next, a state transition occurs that presumably makes sense, or is valid, to the reader.

This correspondence is especially clear in Eli Parra’s comics rendition of Euclid’s Elements. Consider the gutters between the sequence of panels in Figure 5. There is a logical sequence that takes place here: equal radii implies equal lines implies an equilateral triangle.

Why is this correspondence important? It could inform the language design decisions in this space. Our language is declarative and semantics-driven, so mathematical notations and renderings represented in our language ought to be logically valid, especially when constructing visual proofs.

4.2 Parallelism as Counterfactual

A recurring meta-pattern in our patterns has been the idea of parallelism: Multiples, Predicates, and Nesting are all capable of conveying parallel scenarios. What arises from the use of this concept is a fundamental primitive of mathematical inquiry: the use of counterfactuals, or “what-ifs”. When mathematicians face a fork in the road for a proof (separate “cases”), they construct parallel timelines and walk down each. We can reify this

process by taking advantage of the affordances of the 2-d plane: parallel scenarios can be shown *in parallel*.

4.3 Additions as Augmentation

Several patterns don't necessarily have direct textual correspondences, but we believe this highlights the shortcomings of the latter medium. In particular, Multiples, Foci, and Emanata are patterns that reify abstraction in a way that textual mathematical narratives have failed to capture. Most importantly, the patterns leverage the reader's perceptual facilities in a variety of ways, making ideas digestible, comparable, and ready-at-hand.

5 Conclusion

As mathematical diagrams become easier to create, we must remember that single diagrams only tell part of the story. In this paper, we have suggested a future where diagrams tell complex stories with deep structure, attentional mechanisms, and recombinations. We are not new in suggesting such a medium, but rather, in identifying the medium's principles and assimilating them into a declarative language that intertwines mathematical objects and their narrative presentation.

5.1 Future Work

We envision our proposed language as an additional feature of PENROSE itself. To that end, the next step is to implement this feature. The continuity problem alone poses a significant task, as it requires a high degree of control over the optimizer's behavior: between panels, which values should be held invariant, and which are free to assimilate around new objects? This depends on the high-level question of what makes an object recognizable to the reader between each panel.

A significant segment of the visual narrative medium we have not addressed is *layout*. We have implicitly assumed that the `Panel` function creates standalone panels, and the `NestedPanels` function somehow places child panels by anchoring them to objects in the parent. Sequential panels might have different proportions on a printed page, or mixed vertical and horizontal orderings. Floating captions and text-only "speech bubbles" within panels also occur frequently in visual narratives, with a host of layout constraints of their own. This may be a design problem left for the diagram creator to deal with. Or, the PENROSE system can optimize local and global layouts like it does already for mathematical objects and labels.

Visual mathematical narratives as we have proposed here are still static pieces of content. We envision interactivity as an extension of this medium: prodirect manipulation, draggable global parameters, and *Hypercard*-like clickable targets are just a few paths of potential inquiry.

References

- 1 R.H. Abraham and C.D. Shaw. *Dynamics: The Geometry Of Behavior*. Addison-Wesley studies in nonlinearity. Basic Books, 1992.
- 2 C.C. Adams. *The Knot Book*. W.H. Freeman, 1994. URL: http://math.harvard.edu/~ctm/home/text/books/adams/knot_book/knot_book.pdf.

- 3 Jason Alderman. Generating comics narrative to summarize wearable computer data. *School of Literature, Media, and Communication Masters Projects*, 2006. URL: <https://smartech.gatech.edu/handle/1853/16822>.
- 4 Richard J. Anderson, Crystal Hoyer, Steven A. Wolfman, and Ruth Anderson. A study of digital ink in lecture presentation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, pages 567–574, New York, NY, USA, 2004. ACM. URL: <http://doi.acm.org/10.1145/985692.985764>, doi:10.1145/985692.985764.
- 5 Rudolf Arnheim. *Visual thinking*, page 54. Univ of California Press, 2004.
- 6 Benjamin Bach, Natalie Kerracher, Kyle Wm. Hall, Sheelagh Carpendale, Jessie Kennedy, and Nathalie Henry Riche. Telling stories about dynamic networks with graph comics. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 3670–3682, New York, NY, USA, 2016. ACM. URL: <http://doi.acm.org/10.1145/2858036.2858387>, doi:10.1145/2858036.2858387.
- 7 Ruth C Clark and Richard E Mayer. *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning*, page 206. John Wiley & Sons, 2016.
- 8 Neil Cohn. Beyond speech balloons and thought bubbles: The integration of text and image. *Semiotica*, 2013(197):35–63, 2013. doi:10.1515/sem-2013-0079.
- 9 Neil Cohn. Visual narrative structure. *Cognitive science*, 37(3):413–452, 2013.
- 10 Neil Cohn. Narrative conjunction’s junction function: The interface of narrative grammar and semantics in sequential images. *Journal of Pragmatics*, 88:105 – 132, 2015. URL: <http://www.sciencedirect.com/science/article/pii/S0378216615002507>, doi:<https://doi.org/10.1016/j.pragma.2015.09.001>.
- 11 Neil Cohn. *Combinatorial Morphology in Visual Languages*, pages 175–199. Springer International Publishing, Cham, 2018. URL: https://doi.org/10.1007/978-3-319-74394-3_7, doi:10.1007/978-3-319-74394-3_7.
- 12 Joseph (Yossi) Gil, David H. Lorenz, and Matan I. Peled. THAPL—a theatrical domain specific language. In *Domain-Specific Language Design and Implementation (DSLDI 2017)*, co-located with SPLASH, October 2017. URL: <https://2017.splashcon.org/details/dslidi-2017/3/Thap1-A-Theatrical-DSL>.
- 13 Robert L. Goldstone, Sam Day, and Ji Y. Son. *Comparison*, pages 103–121. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. URL: https://doi.org/10.1007/978-3-642-03129-8_7, doi:10.1007/978-3-642-03129-8_7.
- 14 Christian Greiffenhagen. Video analysis of mathematical practice? different attempts to “open up” mathematics for sociological investigation. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 9(3), 2008. URL: <http://www.qualitative-research.net/index.php/fqs/article/view/1172>, doi:10.17169/fqs-9.3.1172.
- 15 Christian Greiffenhagen. The materiality of mathematics: Presenting mathematics at the blackboard. *The British Journal of Sociology*, 65(3):502–528, 2014. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-4446.12037>, doi:10.1111/1468-4446.12037.
- 16 Karl Hasselström and Jon Åslund. The shakespeare programming language, 2001. URL: <http://shakespearelang.sourceforge.net/report/shakespeare/shakespeare.html>.
- 17 David C Keenan. To dissect a mockingbird: A graphical notation for the lambda calculus with animated reduction, August 1996. URL: <http://dkeenan.com/Lambda/>.
- 18 Philip J. Kellman, Christine M. Massey, and Ji Y. Son. Perceptual learning modules in mathematics: Enhancing students’ pattern recognition, structure extraction, and fluency. *Topics in Cognitive Science*, 2(2):285–305, 2010. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1756-8765.2009.01053.x>, doi:10.1111/j.1756-8765.2009.01053.x.
- 19 Nam Wook Kim, Nathalie Henry Riche, Benjamin Bach, Guanpang A. Xu, Matthew Brehmer, Ken Hinckley, Michel Pahud, Haijun Xia, Michael McGuffin, and Hanspeter Pfister. Datatoon:

- Drawing dynamic network comics with pen + touch interaction. In *CHI 2019*, pages 1–12. ACM, May 2019. URL: <https://www.microsoft.com/en-us/research/publication/datatoon-drawing-dynamic-network-comics-with-pen-touch-interaction/>.
- 20 Stephen M Kosslyn, William L Thompson, and Giorgio Ganis. *The case for mental imagery*, page 12. Oxford University Press, 2006.
 - 21 G. Lakoff and R.E. Núñez. *Where Mathematics Comes from: How the Embodied Mind Brings Mathematics Into Being*. Basic Books, 2000.
 - 22 Jill H. Larkin and Herbert A. Simon. Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11(1):65–100, 1987. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1551-6708.1987.tb00863.x>, doi:10.1111/j.1551-6708.1987.tb00863.x.
 - 23 Bruno Latour. Visualisation and cognition: Drawing things together. *Avant: Trends in Interdisciplinary Studies*, 3(T):207–260, 1986.
 - 24 Paisley Livingston. Nested art. *The Journal of Aesthetics and Art Criticism*, 61(3):233–245, 2003. URL: <http://www.jstor.org/stable/1559175>.
 - 25 Chris Martens and Rogelio Enrique Cardona-Rivera. Discourse-driven comic generation. 2016.
 - 26 Scott McCloud. *Understanding Comics*. HarperCollins, 1994.
 - 27 Jason McIntosh. Comicsml: A simple markup language for comics, 2001. URL: https://jmac.org/projects/comics_ml/about.html.
 - 28 Wode Ni, Katherine Ye, Joshua Sunshine, Jonathan Aldrich, and Keenan Crane. SUBSTANCE and STYLE: domain-specific languages for mathematical diagrams. In *Domain-Specific Language Design and Implementation (DSL/DI 2017)*, co-located with *SPLASH*, October 2017. URL: <https://2017.splashcon.org/details/dsl/di-2017/10/Substance-and-Style-domain-specific-languages-for-mathematical-diagrams>.
 - 29 Eli Parra. Points & lines: A reinterpretation into comics of book 1 of euclid’s elements., January 2018. URL: <http://elzr.com/pointsandlines/>.
 - 30 Grant Sanderson. Linear transformations and matrices | essence of linear algebra, chapter 3, August 2016. URL: <https://www.youtube.com/watch?v=kYB8IZa5AuE>.
 - 31 Grant Sanderson. This problem seems hard, then it doesn’t, but it really is, August 2019. URL: <https://www.youtube.com/watch?v=M64HUIJFTZM>.
 - 32 Mario Saraceni. *Relatedness: aspects of textual connectivity in comics*, pages 115–127. Bloomsbury Publishing Company, 1 2016.
 - 33 Paige Scalf, Ana Torralbo, Evelina Tapia, and Diane Beck. Competition explains limited attention and perceptual resources: implications for perceptual load and dilution theories. *Frontiers in Psychology*, 4:243, 2013. URL: <https://www.frontiersin.org/article/10.3389/fpsyg.2013.00243>, doi:10.3389/fpsyg.2013.00243.
 - 34 Joseph D. Sloan. Visual narrative styles in mathematics and computer science. In *Proceedings of the 2007 Visual and Computational Teaching and Learning Conference*, 2007. URL: <https://pdfs.semanticscholar.org/7f31/bee2db113a0f5f27ff0ec5657111e6effefe.pdf>.
 - 35 Nick Sousanis. *Unflattening*, page 66. Harvard University Press, 2015.
 - 36 F. Thomas and O. Johnston. *The illusion of life: Disney animation*. Hyperion, 1995.
 - 37 Edward R. Tufte. *Visual Explanations: Images and Quantities, Evidence and Narrative*. Graphics Press, Cheshire, CT, USA, 1997.
 - 38 M. Walker. *The Lexicon of Comicana*. iUniverse, 2000.