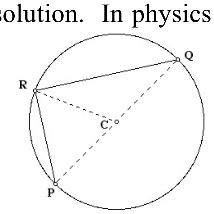
Why Illustrations aid understanding

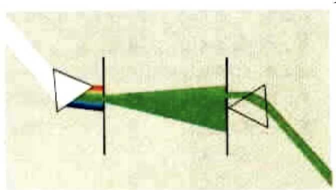
David Kirsh
Department of Cognitive Science
University of California, San Diego
La Jolla, CA 92093-0515
kirsh@ucsd.edu

Abstract:. A small collection of illustrations is provided to show some of the diverse ways illustration may aid understanding. The display of parts and assemblies often relies on techniques such as explosions and canonical views to communicate the global structure and relations of a system that may have hidden pieces. Book illustrations exemplify specific visions of described situations and allow readers to save memory and summarily review potentially complex descriptions. Visual proofs abstract from details and embody reliable metatheories that provide semantic guarantees for inferences. And conceptual illustrations when effective rely the logical method of universal generalization to help viewers grasp general ideas.

Introduction

Illustrations, when well designed, are a powerful aid to understanding non-quantitative information. In geometry they help students to understand Euclidean problems and reason in a constructive manner to a solution. In physics they help students grasp principles, which on the basis of discursive accounts alone, may be hard to comprehend. In narrative fields, such as children's stories, fairy tales or biographies, they typically enhance the reading (or listening) experience and facilitate comprehension. Much has been written on the psychology of text-image interaction, and typically illustrations are thought to fall within the purview of these discussions. For example, Levin has suggested five primary functions for pictures [Levin et al. 1987]: Decoration, Representation, Organization, Interpretation and Transformation,. Hunter and colleagues distinguish between: Embellish, Reinforce, Elaborate, Summarize and Compare. [Hunter et al 1987]. In Mishra [1999] the concern shifted to how scientific illustrations are seldom neutral but rather are theory laden representations that may bias the way readers understand fact and principles. My objective, here, is to explore a set of conceptual and foundational issues related to understanding illustration. I explore three main questions:





1. what is the native expressive power of illustration, and how can static visual techniques and conventions be used to increase that power;
2. why reasoning with illustrations can be seen as analogous to universal generalization in formal logic;
3. how can interactivity be used to increase the expressive power of illustrations.

What is an Illustration?

According to Miriam Webster an illustration is:

a: an example or instance that helps make something clear
b: a picture or diagram that helps make something clear or attractive

In appendix One there are several examples of illustrations. As is obvious from their diversity, illustrations can serve different functions, and rely on a wide range of visual techniques to be effective. Rather than examine all these functions let us begin our inquiry by looking at one familiar way illustrations function: to provide a visual depiction of written text.

For instance, in Figure 1, we see a children's book illustration where the drawing elaborates the text and gives a possible interpretation of it. There is no requirement to assume that the 'real' situation referred to in the text is identical to the depiction. The illustration may exaggerate features, it may violate certain naturalistic assumptions for comic or other effect, or it may hide from view details of real situations. At the same time, an illustration also typically adds details that are unmentioned in the text.

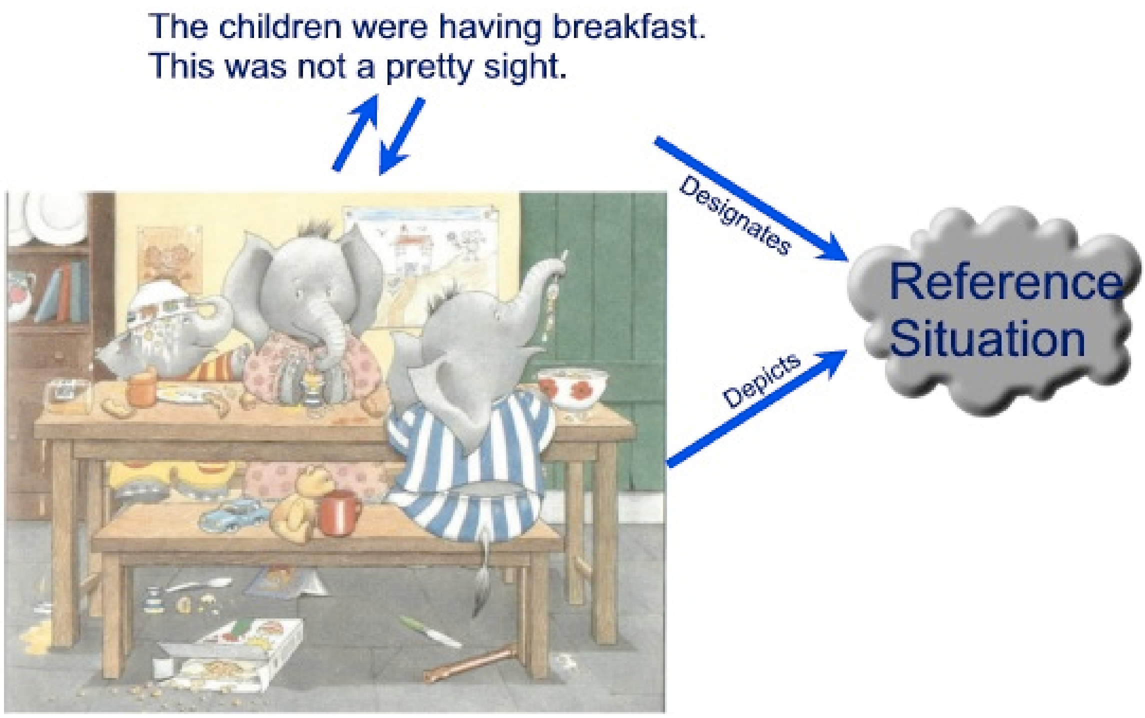
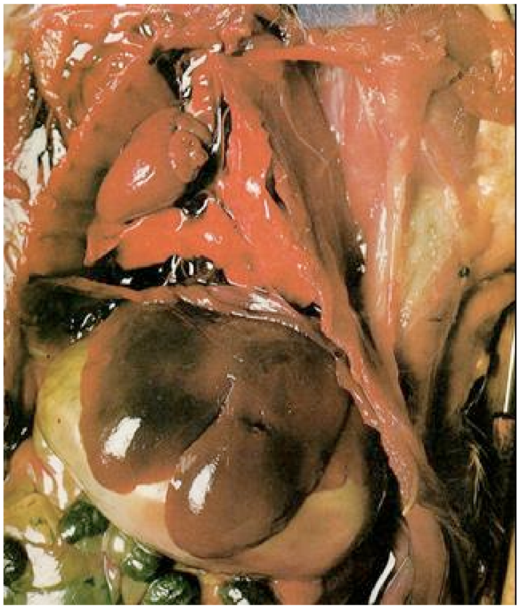


Figure 1. The semantics of text illustration here is that the text designates a 'fictional' reference situation which the illustration depicts. Unlike pictorial representations the constraints on depiction here are weak enough that the reference situation may be imagined in different ways. This illustration is simply one way of imagining three young elephants making a mess while eating breakfast.

Another feature of many illustrations, especially those illustrations which aid qualitative understanding, is that they may abstract from features in the referent situation in order to emphasize certain attributes or relations. For instance, in figure 2a and 2b there is an important difference between the picture of the cardio-vascular system and its illustration. The picture is cluttered and requires an experienced eye to identify parts. Even when major parts are labeled in the picture, it is still difficult to determine their relevant structure and their boundaries. Thus in 2b we see that the key structural entities are shown in outline and simplified. Moreover, parts that are important elements of the system but virtually impossible to see in real life are enlarged and inserted into the illustration to allow viewers to see the system as a functioning whole.



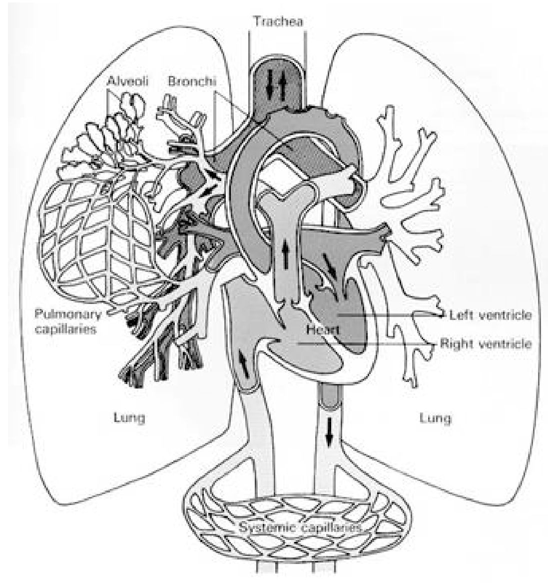
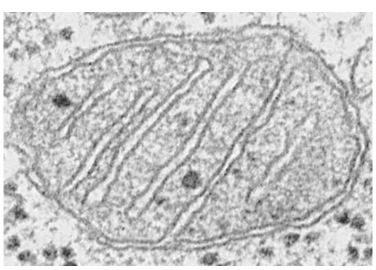


Figure 2a. **Figure 2b**

Here we see both a picture of a cardiovascular system and an illustration of it. Not only does the illustration abstract from many of the visible properties of the system seen in the picture, it distorts them too, since such things as 'the systemic capillaries', the network of vessels shown near the bottom of the picture, are in the wrong place spatially and much exaggerated in size.

In figure 3a and 3b we see another way natural images can be 'cleaned up'. The biological cell on display has been cut out in a way that would be impossible in real life. This makes it possible to exhibit structural parts of the cell from a perspective that maximizes display. The illustration depicts and partially defines a **canonical view** of the cell. Since these views are not built up by actual exposure to cell images they are canonical only in the sense that if we had a macro object of the same shape and structure we would show it off that way.



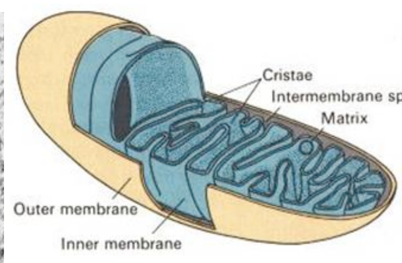
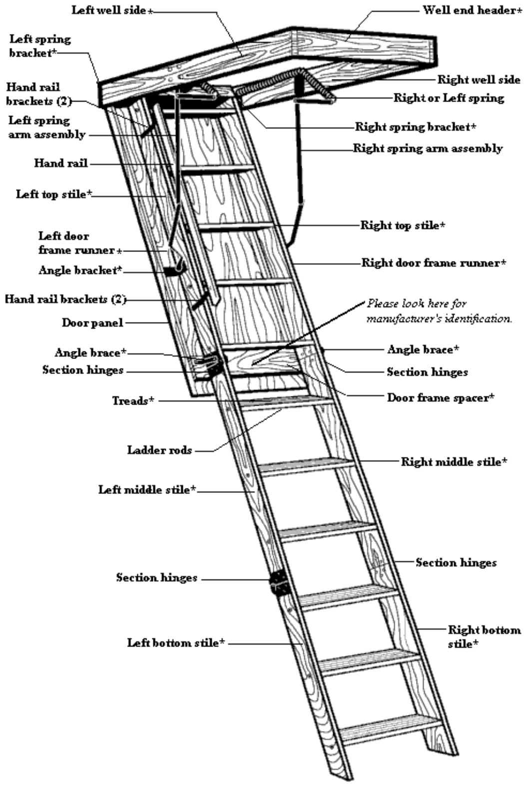


Figure 3a **Figure 3b** **Figure 3c**

An illustration offers the chance to emphasize certain elements of the real object which are hard to see, either because of 3D perspective, or because the natural object cannot, in real life, be cut away to best effect, or because its parts cannot be separated as cleanly as they can be in illustrations, where clutter can be removed. It is customary to try to illustrate a structure from a canonical view, which is both revealing and memorable.

In figure 4 a and 4b we see yet another example of illustrative techniques for structuring information. Illustrations often use a stylized format to display parts. In 4a the intact object is displayed in a naturalistic perspective that exposes as many parts as possible, each identified by an arrow and labeled. In 4b the object is exploded and the parts laid out in space in a way that allows readers to infer the location and assembly of each part in the whole system.



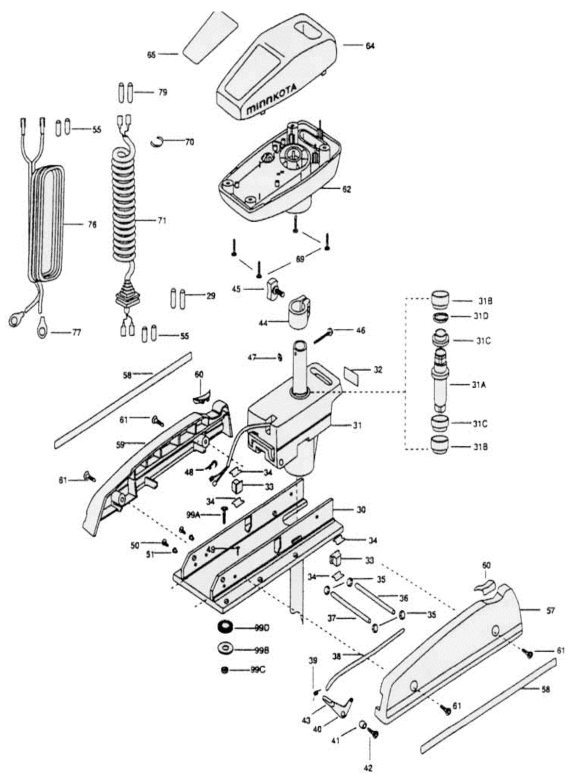


Figure 4a **Figure 4b**

In 4a all the parts of this drop down ladder are labeled by attaching connectors to a canonical view of the ladder. In 4b the power head of an outboard engine is shown in an exploded form to show off the small pieces and also show how they fit together.

Figures 2,3 and 4 all show conventions in illustrating parts or systems of functioning elements. They make essential use of visual-spatio layout to explain the physical structure of an object or system. But illustrations are also used to help explain hard concepts, and the reason they are successful in such case is more like the reason narrative illustrations are helpful: they save memory because they allow the viewer to look at all the facets of the situation at once rather than sequentially as in a paragraph of prose, they save the cognitive effort of visualization, and they allow for more ready understanding of the text, particularly when the correspondence between textual element and visual element is easy to map. For instance, children can answer questions about a passage better if it is accompanied by an illustration.

In figure 5 we see a version of a sketch by Newton's of his famous experiment demonstrating that a second refraction of a colored ray does not alter its color. In one sense a sketch is not necessary to explain that white light is made up of all the colors of the rainbow and that pure colors, such as green, are simple and cannot be broken into a rainbow. A prism pulls the colors out of white light, it does not magically create rainbows. Newton may have chosen to sketch his experiment because its visual depiction allows a viewer to linger on the image and think through the story and argument visually. It saves the reader from going back and reading the explanation and description sequentially.

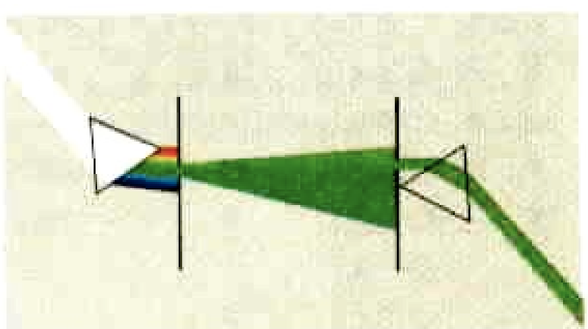


Figure 5.

One advantage of depicting an experiment visually is that readers can simultaneously see the experimental elements and relations. Illustrations save memory. Images are spatial, and in sense, non-temporal. They can store all the relations discussed in the text simultaneously. Text, by contrast, is read in temporal sequence and requires memory to keep all the parts in place.

The final type of illustration we will consider is one that is closely related to conceptual illustration. Visual proofs are illustrations in which the solution to an abstract proposition can be 'readily' grasped by looking at the labels and properties of the illustration. As can be seen in figure 6, for example, it is possible to prove the mathematical claim that $(A + B)^2 = A^2 + 2AB + B^2$ by visually noting patterns, properties or regularities right in the diagram. Although such patterns should be easy to recognize they still require the viewer to reason as well as note salient features. For instance in figure 6 it is necessary to notice that the length of the sides of the square are (A + B), it is necessary to remember that the area of a square is determined by squaring the length of a side, in this case $(A + B)^2$, and it is necessary to see that there are two instances of (A + B). Readers must also know that the sum of the area of all the parts of a surface equals the area of the whole surface. Thus there are several inferential steps involved in realizing that the illustration embodies a proof. But the illustration, in some sense, contains enough information, and in a readily accessible form, that what would otherwise be a highly abstract proof is nearly trivial here and also satisfying. It is a better way of representing the problem than the more familiar algebraic one which does not 'explain' the identity as well. [1]

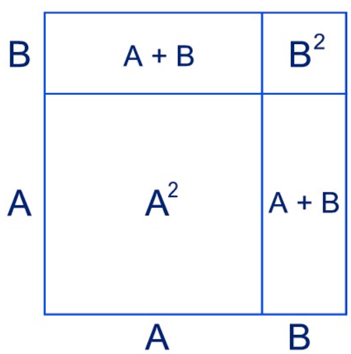


Figure 6a.

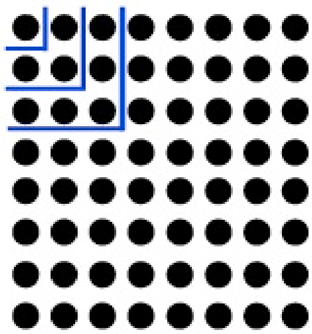


Figure 6b.

Figure 6a

The visual proof that $(A + B)^2 = A^2 + 2AB + B^2$ requires several steps of noticing important relations, remembering general principles and deriving their implication. But the result is a highly intuitive and satisfying understanding of the identity.

Figure 6b

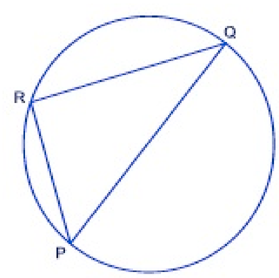
Prove $1 + 3 + 5 + \ldots + 2n\text{-}1 = n^2$

Let us look more closely at the reasoning process involved in interpreting conceptual illustrations.

Illustration as Universal Generalization

The idea I shall pursue here is that conceptual illustration helps readers understand a hard passage the way the method of universal generalization in logical reasoning helps reasoners to find a proof. This may sound obscure but the idea is simple. The rule in deductive reasoning called Universal Generalization says that if P(c) holds for any arbitrary element c of the universe, then we can conclude that $\forall x\ P(x)$. Accordingly, to prove $\forall x\ P(x)$, we take an arbitrary element c in the universe and prove P(c). Then using the rule of Universal Generalization we can conclude $\forall x\ P(x)$. The hard part of applying this powerful method of proof outside of formal logic is knowing that the reasoning one has been doing on a concrete instance does not depend on any idiosyncrasies of that instance, and so applies to all instances in that domain.

For instance, in geometric reasoning, if we are asked to prove that the angle PRQ is always 90°, if it lies on the circumference of a circle in which PQ is the diameter, it is almost impossible to proceed without first drawing a triangle to interpret the claim. Of course, the triangle we draw has special properties. It may be isosceles, or PR may be tiny, but in looking for a proof we must be certain that none of the properties we infer about this triangle depend on its being the particular triangle it is. In Figure 7 we give a brief proof.

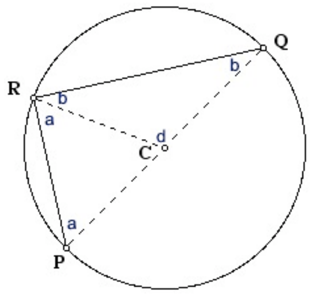


---

[1] To algebraically prove that $(A + B)^2 = A^2 + 2AB + B^2$ expand $(A + B)^2$ to (A + B)(A + B) and use the distributive property to derive $A^2 + AB + AB + B^2$. This proof is easy enough but hardly carries the intuitive force that comes from the geometric version.

Figure 7.

To prove that the angle PRQ is always 90° we begin by drawing a circle with PQ as its diameter and R on its circumference. Now draw a line from the midpoint C of the circle (which is also the midpoint of the base) to R. Because C is the center of the circle the lines PC, CQ and CR are all equal in length. Because the angles of any triangle sum to 180° we know that $2b + d = 180° = 2a + c$. We also know that $c + d = 180°$. From this it follows that $c + d = 2b + d$, hence $d = 2b$, and likewise $c = 2a$. Hence $c + d = 2(a + b) = 180°$, and so $a + b = 90°$.



The key step from a logical viewpoint is generalizing to the claim that this proof applies to all such triangles. We can convince ourselves that this follows because R could in principle be moved along any point on the circumference between P and Q and the resulting triangle would still have PC = CQ = CR.

Now looking at a conceptual illustration for a scientific claim we find a similar case of working with a specific instance being easier than working with the general case. In figure 8 we see Newton's illustration of why the moon revolves around the earth and how it is related to the falling of an apple. He begins by pointing out that a stone or an apple if dropped from a height falls straight to the earth. If it is thrown outward it doesn't go straight forever but is once again drawn by its weight to earth. The harder you throw an object the farther it travels straight before coming back to earth. If you throw it hard enough it will go completely around the earth and come back to your hand. Now imagine that instead of standing on the surface of the earth, or as in the illustration, instead of standing on a mountain and throwing the stone, you are many miles higher. If you throw the stone now it will rotate around the earth before coming to rest, or going off into space (if thrown really hard). The higher the object the longer the orbiting. There is even, presumably, an equilibrium point at which it orbits forever, neither coming to rest on the earth or veering off into space.

As with our geometric case, the illustration here is a special case of Newton's argument. Indeed it is filled with superfluous attributes, ranging from the image of the Earth to the height of the orbits, to the misleadingly exaggerated elliptical shape of the outermost orbit. Nonetheless, the illustration seems valid because figures on the earth are irrelevant details, unconnected to gravity, the choice of trajectories doesn't matter for conveying the main point about the regularity of moving further but finally returning to the surface.

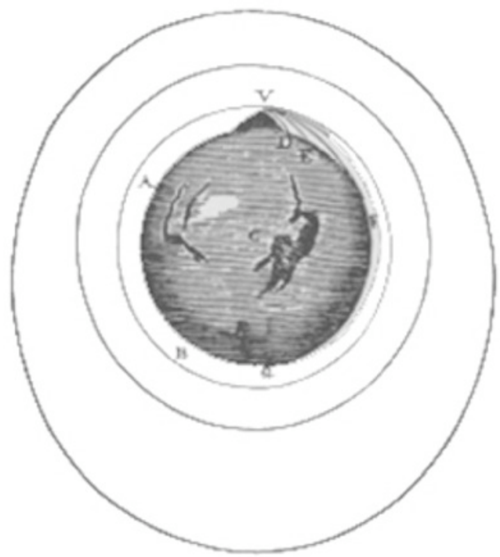


Figure 8.

Unfortunately there are no guarantees that reasoning from this sort of illustration are valid. That is not the case in all diagrams however.

Semantic Guarantees

In geometry the practice of construction and diagramming has evolved so that students are taught how to move back and forth between formal algebraic or Cartesian reasoning and purely geometric, visuospatial reasoning. This reduces the chances of generalizing special attributes invalidly. Students learn not to label angles or lengths as identical based on appearances alone. Labeling is a discipline which students need to learn. And though the particularities of an illustration may help a student to conjecture attributes that may be valid the actual marking of these attributes is supposed to follow certain rules of validation.

The more formal or mathematically based an illustration is the more likely it has a well defined metatheory about what is and is not a licensed inference. For instance, map projections are attempts to portray the surface of the earth or a portion of the earth on a flat surface. Some distortions of conformality, distance, direction, scale, and area always result from this process. Some projections minimize distortions in some of these properties at the expense of

maximizing errors in others. As a result a knowledgeable map user will keep several different projections available for answering different questions. If he wants to know the area of a region then he will use a map which preserves area, though distorting the apparent angles between points. If he wants to determine the direction to sail he will choose a map that portrays azimuths (angles from a point on a line to another point) correctly in all directions.

Let us call an illustration in which the rules of interpretation are well defined and valid an illustration with clear semantic guarantees. If you perform the authorized actions on the illustration – if you interact with it according to certain rules – you enjoy certain guarantees that your actions will deliver correct answers.

All too often, however, it is not possible to know what the semantic guarantees of an illustration are. It is undeniable that certain inferences are validly licensed but it can be difficult to know which.

Abstraction

Because of the dangers of invalid generalization a natural convention has developed that illustrations should abstract from irrelevant detail. Instead of showing cases which resemble real world situations with all their irrelevant color, texture, size, shape and so on, there is tendency to abstract from reality and move into a more stylized and stripped down display often with well defined visual symbols. See figure 9.

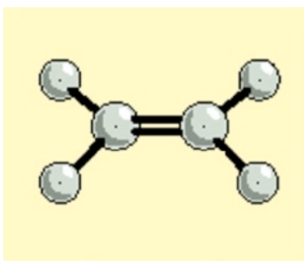


Figure 9a

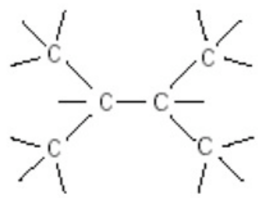


Figure 9b

Figure 9c

Figure 9.
Chemical bonds are shown here in figure 9a. with extra detail to help engage spatial and mechanical intuitions. In figure 9b the volume of each atom has been removed and replaced with a symbol representing it's atomic identity. In Figure 9c even this level of detail has been removed to allow for easier inclusion in complex molecules.

There is a tension in using abstraction for illustration. For on the one hand the addition of labels, visual symbols and interpretative conventions can help overcome the limited expressiveness of imagery. See appendix 2 for some examples of expressive limitations. But on the other, the more illustrations approach linguistic description the less special they are. Indeed one reason for using illustrations is to help readers explore a problem's structure. For instance, one obvious limitation of pure imagery is its difficulty in expressing incomplete knowledge. I may know that route 66 either runs through Washington DC or Baltimore but I don't know which. In a single illustration route 66 has to be placed going through the one town or the other. It must be given a definite location not an either here or there without showing either one. And yet the very reason I might wish to use an illustration is to explore what it would be like if Route 66 went to Baltimore. Reasoning from what I know I may be able to derive a contradiction and so reject this case.

Discussion

The topic of illustration and how it engages thought and understanding can benefit from foundational inquiries into the logical role it may play in the construction of proofs, refutations and conceptual exploration. Owing to the diversity of illustration it is unlikely that there will be a single unifying theory. But in each case it is helpful to determine the metatheory of the illustration type to understand if there are semantic guarantees it offers for inferences based on it.

References

Brown, R. B. (1991) *The Laboratory of the Mind: Thought Experiments in the Natural Sciences*. Routledge: New York.

Hunter, B.; Crismore, A.; and Pearson, P.D. (1987). Visual displays in basal readers and social studies textbooks. In *The Psychology of Illustration, Basic Research*, eds. D.M. Willows and H.A. Houghton, 2:116-135. New York: Springer.

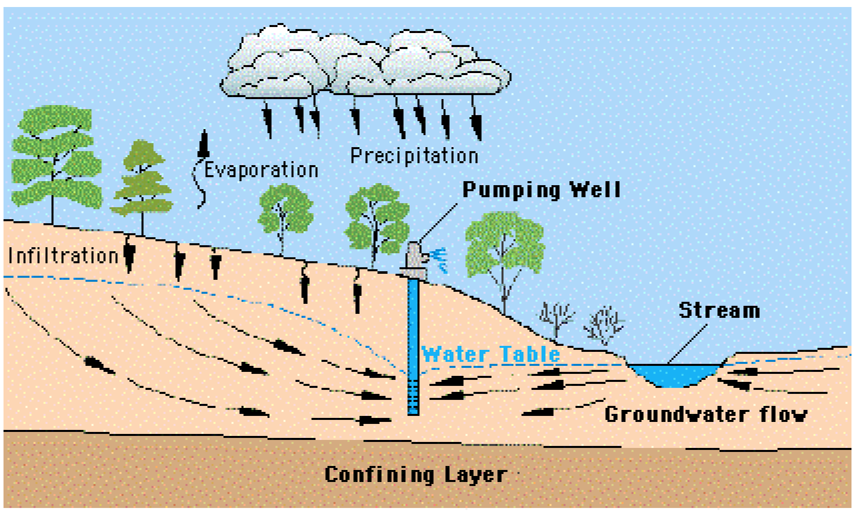
Levin, J.R.; Anglin, G.J,; and Carney, R.N. (1987). On Empirically validating functions of pictures in prose. . In *The Psychology of Illustration, Basic Research*, eds. D.M. Willows and H.A. Houghton, 2:51-85. New York: Springer.

Mishra, P. (1999). The role of abstraction in scientific illustration: Implications for pedagogy. *Journal of Visual Literacy*. 19, 2, 139-158

Appendix I

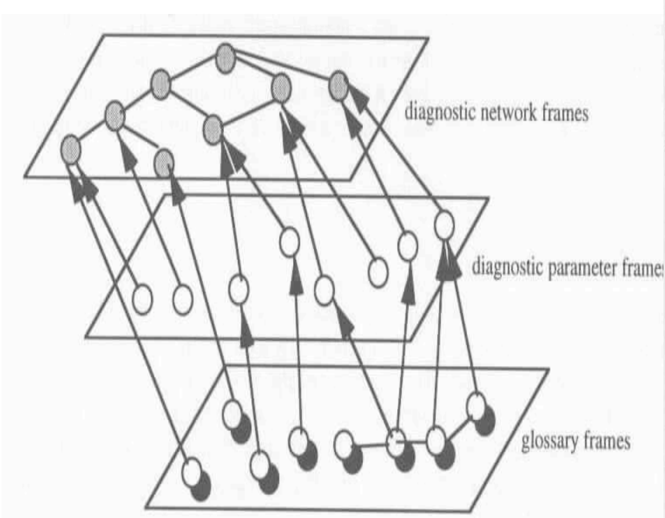
A few other illustrations worth reviewing.

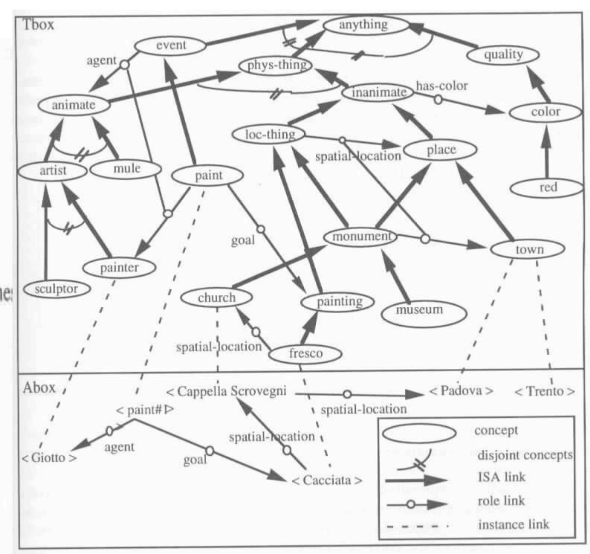
Visual Narratives



(Adapted from USGS)

Connectivity or network Diagrams (org charts)





Appendix II

Expressive power

It has been well noted that images and pictures – at least those lacking labels, annotations, icons, a conventional visual vocabulary and interactivity – have difficulty in expressing incomplete knowledge, modal relations and quantitative values. It is hard o r impossible to **show** many of the things we can **say**. Let us consider each of these expressive limitations, briefly.

Incomplete knowledge:

- Negation: $\forall x \neg Ax$, $\neg Aa$, there are no dragons (in this domain); a is not an apple. There is no natural way to graphically express negation, especially since in logic the scope of a negation symbol can vary. Visual conventions such as overlaid crosshairs do not tell us how comprehensive the negation is (does a dragon with an X through it mean there are no dragons in this frction of the domain but there are elsewhere, or does it mean that dragons are prohibited.) It is virtually impossible to capture the universal notions of everything is not P, which is means nothing is P.

- Disjunction: $A \vee B$, $\forall x(Ax \vee Bx)$, Everything is either Black or White in this domain (but I'm not sure which); the spray on this apple is either poisonous or makes you thirsty (but I can't remember which)

- Modality: $\Box P \rightarrow Q$, $\Diamond Px$, If you forget your number you will certainly have trouble; it might rain tomorrow. Buyt he tame token we cannot express counterfactuals.

- Nested quantifiers: this was a problem for semantic network languages.

- Existential possibility: $\forall x\ (Ax \rightarrow \exists y(In(Cy, Ax)))$; Every apple has at least one creature in it.

Although there are visual devices and conventions for overcoming expressive limitations of bare images none work in all contexts. For instance, to illustrate the proposition that every apple has a creature in it (though we are not sure whether it is a worm or some other thing), we can rely on vagueness, imprecision or size to show the presence of a creature without showing which creature it is.