In E. Lepore and Z. Pylyshyn (Eds.) *What is Cognitive Science?*, Basil Blackwell, 208–229, 1999.

# Does Vision Work? Towards a Semantics of Perception

Jacob Feldman

Dept. of Psychology
Center for Cognitive Science
Rutgers University

Running head: *Does Vision Work?*

Send correspondence to:

Jacob Feldman
Dept. of Psychology
Center for Cognitive Science
Rutgers University, Busch Campus
New Brunswick, New Jersey, 08903
e-mail: jacob@ruccs.rutgers.edu

**Abstract**

Vision routinely provides us with completely compelling—and as far as we know, true—beliefs about the world around us. But are perceptual beliefs really that reliable, and if so, why? Answering this question requires us to delve into the *semantics* of perception: an account of the *meaning* and *truth conditions* behind the heuristic rules whereby the visual system operates. This chapter makes a few proposals about how a "semantics of perception" might work.

## 1 Seeing is believing

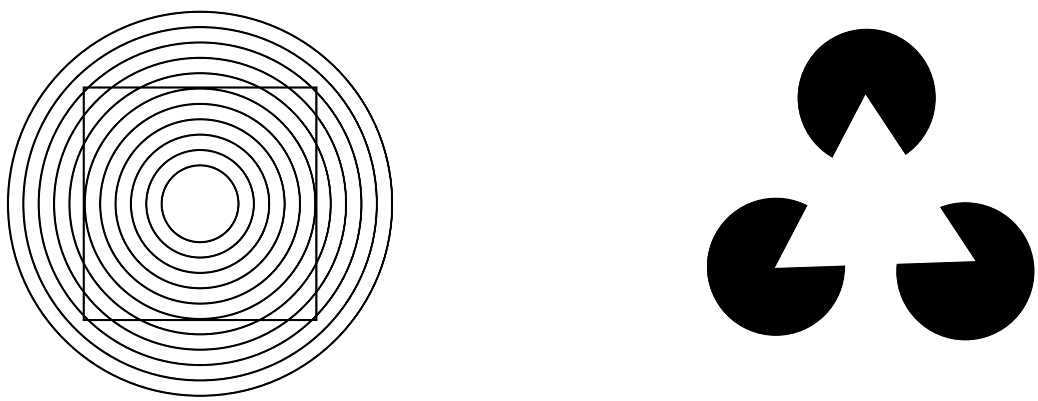
As the old saying goes, seeing is believing. To the average person—a Naive Perceiver—the perceptual experience is so compelling as to leave no room for doubt. What you see is—precisely, unambiguously, and quite literally—what is out there in the world. After all, our very lives depend on the reliability of vision. Every time we cross the street, step into a forest, or put a piece of food in our mouths, we are counting on vision to accurately report the state of the world—no car, no tiger, no hemlock.

But what justifies our faith in vision? To the Naive Perceiver, perceptual beliefs are justified simply because they are *true*—or, in the terminology of vision science, "veridical." How can one doubt what is before one's eyes? But there are a few simple facts about perception that ought to send a shudder of epistemological doubt through our hearts.

*Vision is fallible.* First of all, the visual system is vulnerable to being tricked by certain types of images and scenes, usually called *optical illusions*. (Fig 1 shows two famous examples.) Really, "perceptual illusions" would be a more appropriate term, because usually the explanation has nothing to do with optics, but rather with the way perceptual inference works. These displays have been carefully constructed in order to take advantage of the tricks the visual system exploits all the time (but usually don't lead to errors).

Traditionally textbooks emphasize the sometimes subtle way in which optical illusions reveal the internal operation of the visual system. But the more basic lesson of these illusions is simply that vision is capable of providing you with a completely compelling but *false* belief. This can hardly help but make one wonder about the justification for all the *other* beliefs provided by vision—the ones on whose reliability we daily pledge our lives. How do we know those don't have an equally shaky foundation? Indeed, how would one go about showing that ordinary, garden-variety beliefs generated by the visual system are well-founded? Indeed, what does it mean for a perceptual belief to be well-founded? Philosophers have pointed out that for a belief to be justified it has to be more than simply *true*—after all, that could simply be a fortunate accident, like a broken watch is correct twice a day. To be justified, a perceptual belief must be true in virtue of some reliable causal connection to the world around us. But what is the nature of this connection? This is one of the questions considered in this chapter.

*Vision is ambiguous.* A closely related point is that, as vision scientists never tire of repeating, the visual stimulus is inherently ambiguous. The visual image that falls on the retina contains only a limited amount of information, which is perfectly consistent with a multitude of different states of the world. In the lingo, the *proximal stimulus*—the pattern of stimulation on the retina, the light-sensitive surface of the eye—is consistent with many distinct interpretations of the *distal stimulus*—the thing



**(a)** **(b)**

Figure 1: Two famous perceptual illusions. In (a), the square appears "squished," but isn't. In (b) (due to Gaetano Kanizsa) most people see a triangle in front of three circles (rather than simply three "pac-men". In addition most people report that the triangle is a brighter white than the background, with a noticeable boundary between them. Actually there is no brightness difference, as can be confirmed simply by blocking out two pac-men with your fingers.

"out there" that you are actually trying to perceive.

The most notorious example is depth. The world is three-dimensional, the retinal image two-dimensional; the third dimension, depth, is lost when the world is projected onto our retinas. Fig. 2 gives an example showing how even a simple picture is consistent with any number of totally different 3 - D interpretations, some of them quite bizarre. The reconstruction is unconscious and automatic—you're not aware of having to work at it—but it is nonetheless complex and, in a certain way, intelligent.

We take it for granted that the interpretation we pick—e.g., (a) in the figure—is usually the objectively correct one. But given that all of the wrong interpretations are just as consistent with what we see, what makes us so sure that the one we see is the right one?

## 2 A semantics for perception

These questions of justification and epistemology in vision bear bears on what we might call the *semantics* of perception. This is a somewhat fuzzy term, used in different ways by philosophers, linguists, and computer scientists, and hardly used at all by psychologists, ever eager to affirm their empiricist bona fides. The common threads through all the uses are the twin ideas of *meaning* and *truth conditions.* In philosophy the term usually refers to attempts to clarify the meaning, and thus the content, of ordinary concepts—and hence to spell out what transforms an ordinary belief (an idea held by an agent) into a piece of knowledge (a *justified, true* belief). Philosophical semantics is thus about adding content and substance to otherwise dry syntactic (i.e. formal) constructs.

In this sense, vision seems like a ripe domain for inquiries about semantics, because the visual system is nothing other than a system for mechanically (i.e., computationally) producing beliefs—peculiarly compelling ones, in fact. Yet the philosopher Jerry Fodor has argued (Fodor 1980) that psychological theories, e.g. of perception, cannot have a semantic component, because such theories properly concern only internal mechanisms and representations, while the truth of the resulting beliefs depends on the state of the world outside. The result is that computational theories of perception concern the production of *belief* but not of *knowledge.* That is, nothing in the perceptual theory directly explains whether perceptual conclusions are true or justified, only how they are produced.

Yet it is commonplace for vision researchers to give casual, informal justifications for particular perceptual rules and algorithms, explaining why they perform with reasonable success in practice. And indeed such explanations must depend on characterizations of the way the world typically behaves, and hence take the theorist "outside the head." Such additional theoretical comments may be strictly extraneous

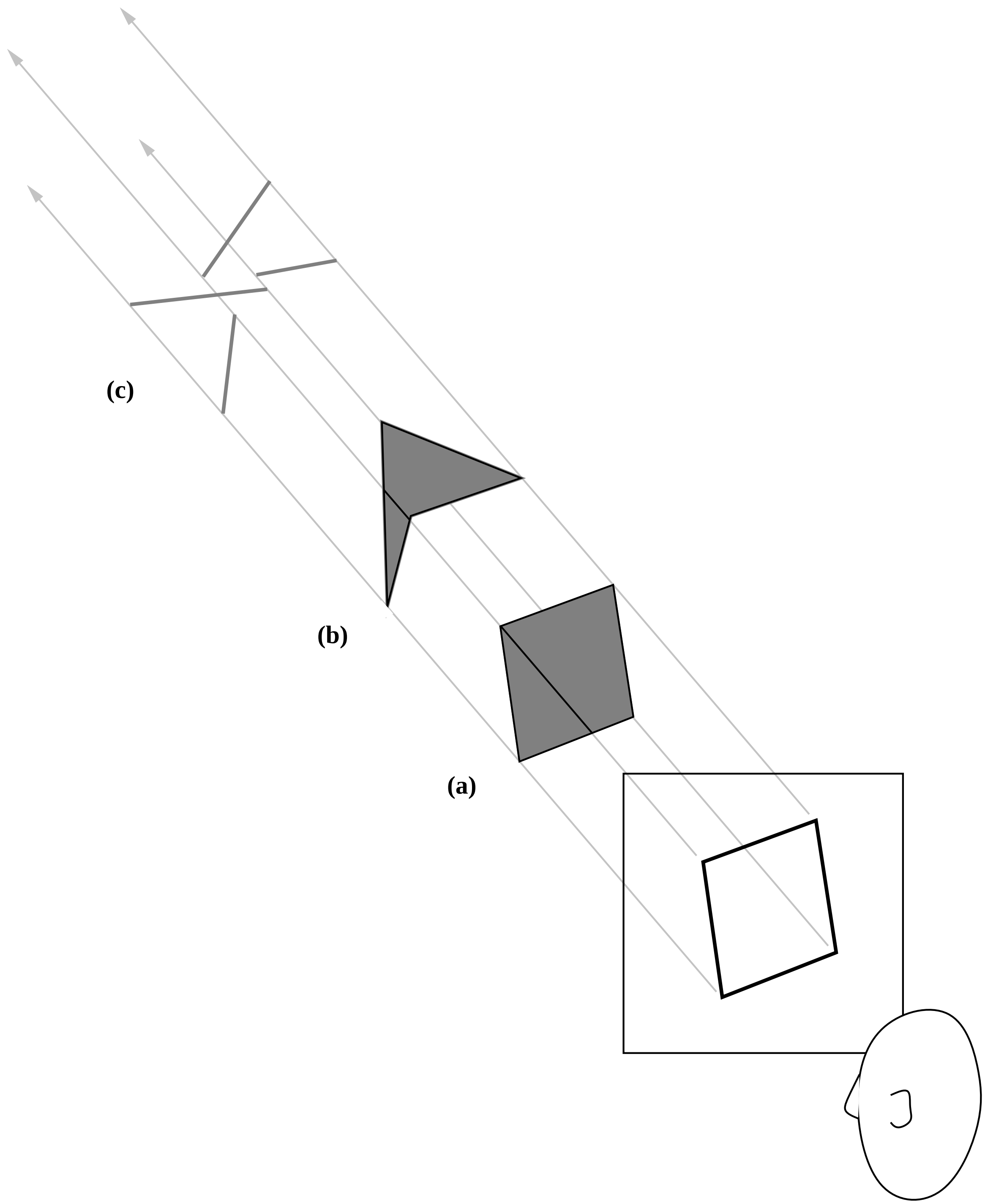



Figure 2: Ambiguity of three-dimensional structure. The image is consistent with a rectangular object (a), which is what we see. But it is also *equally* consistent with an irregularly-shaped object (b); or even with the disconnected collection of line segments (c).

to what psychologists usually call a "process model" (a theory of what mechanisms actually occur inside the system). But they certainly seem like part of a full account of the system—explaining for example why the system works, and why it might have evolved the way it did. In the case of vision this means explaining why the visual system tends to produce true beliefs.

Vision researchers have hardly ever touched on semantics in an explicit way. A notable exception is the work of Bennett, Hoffman, and Prakash (1989), who have proposed a single, unified mathematical framework for all canonical acts of perceptual inference. In this framework it is possible to express truth conditions for a perceptual hypothesis in a consistent and general way, in much the same spirit as the ideas presented here (although with a very different mathematical flavor).

In computer science and mathematical logic, semantics has a similar goal but superficially a very different form. Here the idea is to construct explicit truth conditions for the complex syntactic forms that are the main substance of formal theory. Typically the enterprise is to specify formal conditions that would make a logical formula—a purely syntactic object—true or false. An example is the construction of truth tables, which exhaustively enumerate the truth of a propositional formula as a function of the truth of its atomic propositions. In first- and second-order logic semantic theory gets more complicated and subtle, often centering around attempts to prove that a certain abstract syntactic theory actually has a "model," a concrete object that obeys all of the theory's formal characteristics. To non-logicians, such proofs often have a somewhat anticlimactic air, as the "concrete objects" actually turn out to themselves be symbolic objects (e.g. the so-called Herbrand universe). This can leave the newcomer grasping for some link to the real world. But still the results can lend a satisfying completeness to syntactic theories, justifying in a completely rigorous way that the theory does what it was intended to do, and under what assumptions. Spelling out those assumptions is especially critical. Tacit assumptions sometimes pop up unexpectedly, and different sets of assumptions tend to produce different conceptions of what is going on.

The aim of this chapter is to try to point the way to a semantics of perception, in a sense combining the conceptual goal of philosophical semantics (explaining why perceptual beliefs tend to be true) with the methodology of computational semantics (using formal methods). Although the goal is a formal theory, though, the treatment here will be mostly informal and tutorial.

# 3 Tricks and cheats

Let's start by trying to characterize the canonical logical form of an act of perceptual inference. As discussed above, vision is inherently a problem in which not enough information is given to choose the right answer—as vision scientists say, the image "under-determines" the interpretation. Another way of putting it is that vision is *inductive*—many conclusions are consistent with the bare facts of visual stimulation. All of vision in one way or another deals with inferring from limited information something that is hidden or not directly perceivable.

A very general way of stating the problem is that there are *observable variables* and *hidden variables.* The visual system has access to the value of the former but not the latter. But unfortunately it is usually the latter that are of interest. A canonic example is the dilemma facing our favorite caveman, Ugg, when he picks a piece of fruit. For a piece of fruit $x$ he would like to know whether the fruit is edible

$$\texttt{edible}(x)$$

or poisonous

$$\neg\texttt{edible}(x)$$

Unfortunately for Ugg, that value of the predicate `edible` is not in any way *directly* perceivable—he doesn't, for example, have "edibility" sensors on his retina. All he knows is whether the fruit is `blue`

$$\texttt{blue}(x)$$

or yellow

$$\neg\texttt{blue}(x)$$

(I'll assume throughout that fruits are always either blue or yellow, and that a fruit is edible if an only if it is not poisonous.) How can he infer edibility from color? Clearly, there is no *necessary* relationship between these two variables. But does that mean that there is no usable relationship at all? Indeed it does not.

Let's say that Ugg (or maybe Ugg's ancestors) has some experience with fruit. Specifically, he (or his ancestors) has learned the following "fruit rule:"

*All blue fruit are poisonous*

because, say, in the valley where Ugg lives there only exist blue, poisonous hemlock, and yellow, edible bananas. (In fact I realize that hemlock isn't a fruit, and as far as I know it isn't blue, but let's pretend.)

We can write this rule in logical notation using the operator $\Rightarrow$, usually read as "implies," and the so-called universal quantifier $\forall$, usually read as "for all:"

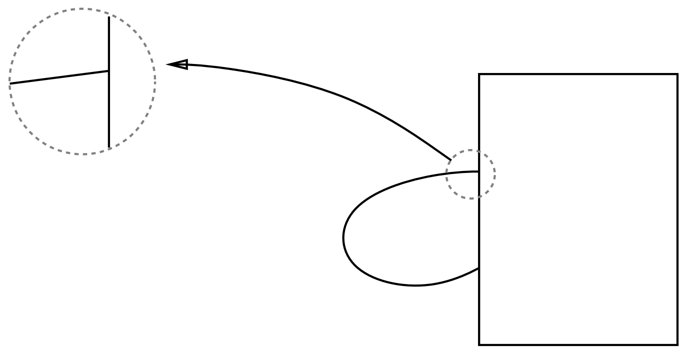


Figure 3: A T-junction. The heuristic rule used is: interpret a T-junction so that the top of the T is the contour of a nearer object, and the stem of the T is the contour of a more distant object. Note that in the setting shown, which is typical, this interpretation is correct.

$$\forall x \ \texttt{blue}(x) \Rightarrow \neg\texttt{edible}(x),$$

The predicate `blue`, unlike the predicate `edible`, is observable, in that inspection of (or some prior inference about) the proximal stimulus can uncover its value. This difference in the epistemological status of the two predicates is crucial.

I admit that this may seem like an odd example because it seems doubtful that knowledge of the relationship between shape and edibility would be hardwired into the visual system—and indeed, the example is completely artificial in that sense. But the point is that there is a crucial relationship between a rule that Ugg uses and a putative fact about the world Ugg lives in.

Such a rule is called a *heuristic*, a loose term meaning more or less that the rule usually tends to work in practice, but that there is no definite logical or a priori reason why it must work. Such rules are pervasive in the visual system, as indeed in any successful inductive inference system. Indeed the entire operation of the visual system can be thought of as an attempt to "cheat" by picking up *indirectly* information that is not available directly.

One example of a visual heuristic is the way the perceptual system treats a T-junction, a point in the image where one contour abuts against the middle of another (Fig. 3). The default interpretation here is that the top of the T is the edge of a nearer object, while the stem of the T is the edge of a more distant object. It is entirely possible for this interpretation to be false. For example the T could be a wire frame, with neither part actually an occluding edge of any object. However, as the picture suggests, this interpretation is usually right under ordinary circumstances.

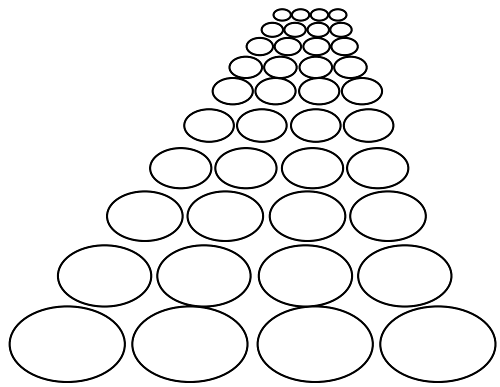


Figure 4: Another heuristic rule: a repetitive texture with a gradually changing scale is interpreted as a uniform texture slanted in depth.

Another example of a heuristic rule is illustrated by Fig. 4. Here a repetitive pattern of ellipses, growing smaller as one moves up the page, is interpreted as a uniform pattern on a surface that is slanted back in depth. (The technical term for such a uniform pattern is *isotropic*, meaning the "same in every direction".) Again, this interpretation may well be wrong. But it is compelling nonetheless, and one suspects that it is right most of the time.

The list of specific heuristic tricks proposed by vision researchers over the years is very large and varied, and, it sometimes seems, not very systematically organized. Even within a seemingly limited domain such as that of line drawings, it has been proposed that people assume that lines are parallel, that vertices tend to form right angles, that lines tend to align with gravity, that surfaces formed tend to be convex, that our viewing direction tends not to align with surfaces, and on and on for dozens more. And this is not to even mention the myriad tricks outside line drawings, involving shading, stereopsis, motion, color, texture, and so forth. This has sometimes led to the idea that vision is nothing more than a "bag of tricks"—a collection of unrelated heuristics with no underlying theory connecting them. Actually, though, many of the rules have a certain underlying formal similarity (see Chapter XX [Leyton] in this volume). Even beyond that, though, I will argue that the seeming chaos in the inventory of visual inference mechanisms is illusory. To see the argument, we must first ask what makes these heuristic rules tick.

The "fuzzy" nature of heuristics rules tends to make people casual about justifying why they usually work. But of course such rules are far from arbitrary. A crucial insight is the fact that if the world had a completely arbitrary structure—distinct properties were in no way causally related to one another—then we would not expect any heuristic rules to work. Rather the success of these rules follows from the way that they depend on things being orderly and well-structured. The rules are parasitic on regularity in the world. This point has been made in a number of thought-provoking articles by Horace Barlow, a pioneer in the modern study of vision (Barlow 1974; Barlow 1991; Barlow 1994; Barlow 1990). Barlow argued that vision depends on what he called "suspicious coincidences" or "significant associations" among what otherwise might be unrelated properties in the visual world. Roger Shepard, a pioneer in the study of learning, called these coincidences "regularities" (Shepard 1989), and argued that the ability of the visual system to accurately recover the structure of the world depends on its being sensitive to such regularities in the right kind of way. Richards and Bobick (1988) have termed this general idea the "Principle of natural modes." In a similar vein, Richards, Rubin, and Hoffman (1982) have taken an important step towards a formal characterization of the validity of perceptual inferences, by stating express mathematical conditions on the existence, uniqueness, and correctness of the solutions to perceptual problems.

The overarching idea in all this work is that behind every perceptual rule is some reliable relationship between some hidden property (e.g. the poisonousness of the fruit, the structure of the surface containing the texture) and some observable property (e.g. the color of the fruit, the texture on the surface). The fact that you can infer something about the one from the other depends on the fact that the relationship between them tends to be well-behaved.

# 4 Constraints

But what does it mean for a rule to be "based on" a fact about the world? Take the isotropic texture rule above as an example. The heuristic rule can be stated more or less as:

**(Isotropic Texture Heuristic)**

> *If a repetitive texture changes scale, then interpret the scale change as a depth change.*

As such, the rule is of the form: if a certain proximate situation obtains, infer a certain distal situation. But note that this rule is *entailed* by (though not equivalent to) a certain hypothesis about the world, namely:

**(Isotropic Texture Constraint)**

> *All repetitive textures are isotropic.*

If this is true in the world, then when one encounters a repetitive texture that does *not* appear isotropic, one can only conclude that it is really uniform but is slanted back in depth—so the heuristic rule holds water. In fact, if textures in the world are *always* isotropic, without exception, then the rule becomes deductively valid; its heuristic nature only stems from the fact that the world occasionally contains a non-isotropic texture. Hence there is a close relationship between the two ideas, but they are not ontologically equivalent. The one is a rule about what to think when a certain image appears; the other is a statement about the structure of the world.

The latter type is what David Marr, one of the founders of Computer Vision, called a *constraint* (Marr 1982). A constraint is an assertion that that the world is well-behaved—more well-behaved that it absolutely has to be—in a some particular way that a perceiver can benefit from. If textures in the world are random and idiosyncratic, then a perceiver cannot infer much about the world by analyzing any particular one. But if textures tend to be isotropic, then the perceiver gets some foothold onto distal structure. Hence a constraint really has two claims: (1) *if* the world behaves a certain way, then a certain kind of inference will be possible, and (2) the world really *does* behave that way, at least most of the time. Hence the success of the rule is directly dependent on the truth of claim (2).

One can see intuitively (without spelling it out completely formally) that the texture constraint entails the validity of the texture heuristic, but *not* vice versa. The heuristic is consistent with many different models of the world, only *one* of which precisely embodies the constraint—an important point we shall return to later.

Like the fruit rule, one can easily express the constraint as some sort of quantified logical statement about the world, e.g.

$$\forall x\ \texttt{texture}(x) \Rightarrow \texttt{isotropic}(x).$$

This type of logical formalism is very useful because it makes it feasible to express, and even prove, statements about the the relationship between specific inference rules used by a perceiver and the abstract universe in which the perceiver is embedded—just what we need for semantics. Hence mathematical logic is a convenient language for getting at the semantics of perception, a suggestion first made by Raymond Reiter and Alan Mackworth (Mackworth 1988; Reiter and Mackworth 1989).

In this connection the switch in the isotropy rule's ontological status—from a heuristic rule to a constraint—may seem subtle, but is important. Unlike the heuristic rule formulation, a constraint formulation such as the above is a (small part of a) *theory of the world*. Thus we have a chance to make an explicit connection between the rules that we use to see and the theories that we (tacitly, unconsciously) hold about the world.

In this light it begins to seem that there isn't really all that much of a distinction between the "bag of tricks" idea and the idea that perceptual inference is systematically organized. Each trick is entailed by a constraint on the world. Put all those constraints together and you have (a part of) a theory of the world. The bag of tricks may seem heterogeneous, but that's only because the underlying theory of the world is heterogeneous—because in fact the *world* is heterogeneous, and the theory is more or less faithful to the world it describes. As for justification, if the theory of the world is right, each of the rules entailed by one of its parts is right; and hence perceptual interpretations based on these rules are justified.

But what does it mean for the perceiver's underlying "theory of the world" to be right?

## 5 Having the right theory of the world

A common intuition about the justification of visual rules is that the system makes the right inferences when it has *the right theory of the world.* But what does it mean to have the right theory? First we must ask "what is a theory of the world"?

In mathematical logic, a sentence is a (possibly quantified) statement such as

$$\exists x \, \mathtt{duck}(x)$$

(the existential quantifier $\exists$ is usually read "there exists," hence "there exist ducks") and

$$\forall x \, \mathtt{duck}(x) \Rightarrow \mathtt{quack}(x)$$

(all ducks quack) and

$$\mathtt{duck}(\text{Ernest})$$

(Ernest is a duck). A *theory* is a set of sentences that is *closed under logical entailment,* meaning that it includes all sentences implied by sentences within it (see Genesereth and Nilsson 1987 for an introduction). For example the above set of sentences is not a theory, but would become one if you include the additional sentence

$$\mathtt{quack}(\text{Ernest}),$$

which is entailed by the previous three. A theory describes an artificial universe, and in a certain sense *creates* a universe by describing all of its properties. (Note that a set of sentences that is not closed under entailment describes some but not all of the properties of its universe.)

One can imagine a theory, in this technical sense, that describes the *actual* universe—the one we live in. Call this theory $\mathcal{T}_0$—"God's own theory of the world." Such a *universal theory* contains all scientific facts and laws, and also contains a myriad of trivial but true facts about what is where and so forth in the world as it stands. By hypothesis, a statement is empirically true if it is contained in $\mathcal{T}_0$. Such a theory is more or less what was envisioned by Logical Positivism, a school of thought about the epistemology of science that many feel more or less describes the attitude of the typical working scientist today. (Strictly speaking, what we usually think of as "science" is really a small subset of $\mathcal{T}_0$, containing only the most general and sweeping laws and omitting many details.) It may be difficult to accept that such a theory can exist in practice, and indeed even in principle there are well-known (e.g. Popperian) obstacles to completely discovering such a theory. But at least the idea has some foundation in conventional philosophy of science.

Now, imagine that we have some hypothetical perceptual inference rule, of the general form

$$A \Rightarrow \alpha,$$

where $A$ is some observable property or variable, and $\alpha$ is some hidden variable. (In general I'll use Latin characters for observable properties and Greek characters for hidden properties.) Here I'm taking the implies operator $\Rightarrow$ to mean more or less that from the antecedent one should infer the consequent.

Clearly, in order for this rule to be justified by reality, it must be the case that

$$\mathcal{T}_0 \Rightarrow [A \Rightarrow \alpha],$$

the universal theory entails the rule; or equivalently,

$$\mathcal{T}_0 \wedge A \Rightarrow \alpha,$$

($\wedge$ is the usual logical operator *and*). That is, if the theory of the world holds true and the observable variable holds true, then the implied hidden variable also holds true—very intuitive. However this formulation is far too weak, suggesting that the justification of a single little perceptual rule requires the support of an entire theory of the universe. This means that the only perceptual rules that are justified are the ones that are valid throughout the whole universe under all circumstances—certainly sufficient, but hardly necessary. We can do better.

What we need instead is a way of reflecting "local" conditions that, while not true in the universe in general, are true in the limited universe of the perceiver. The relationship between color and edibility in Ugg's valley is an example; there is no *general* relationship between color and edibility that holds universally, but there is a *contingent* relationship that holds in Ugg's environment, and can support valid inferences within its limited scope.

To build up a formal expression of specialized conditions, notice first that some theories imply others. We can denote this using the ordinary implication operator, as e.g.

$$\mathcal{T}_1 \Rightarrow \mathcal{T}_2,$$

meaning that if theory $\mathcal{T}_1$ holds, then so must $\mathcal{T}_2$. As we have defined it, this is equivalent to

$$\mathcal{T}_2 \subseteq \mathcal{T}_1,$$

because everything entailed by a theory is contained in it. Now consider augmenting a theory $\mathcal{T}$ by some additional knowledge, encoded in a set of sentences $S$. Denote by $\langle S \rangle$ the transitive closure of a set $S$ of sentences. When we add $S$ to some theory $\mathcal{T}$ and then transitively close their union, we get a new theory

$$\langle \mathcal{T} \cup S \rangle.$$

The additional set of sentences $S$ is called a *refinement* to $\mathcal{T}$, and the new theory $\langle \mathcal{T} \cup S \rangle$ is called the *refined theory*.

A famous example drawn from the history of science is the relationship between Newtonian and Einsteinian mechanics. In classical physics, momentum ($p$) is the product of mass and velocity,

$$p = mv.$$

In everyday physical situations, this expression is so nearly perfectly empirically correct that 19th century physicists could hardly blamed for taking it as gospel. However, using the Lorentz transformations, Einstein showed that momentum could more correctly be expressed by the more complex expression

$$p = \frac{mv}{\sqrt{1 - \frac{v^2}{c^2}}},$$

where $c$ is the speed of light. At first, this expression looks very different; how could Newton have been so wrong? But a closer examination reveals that the two expressions are not so different under ordinary conditions—namely, when velocity is much less than the speed of light. In everyday life, and in all the physical experiments on which classical mechanics was based, $v \ll c$. This means that the "Lorentz correction" approaches unity,

$$\sqrt{1 - \frac{v^2}{c^2}} \approx 1,$$

which means that Einstein's expression reduces to Newton's. Putting this in our terms (though speaking loosely of course) when the refinement $v \ll c$ is added to Einstein's theory, you get Newton's:

$$\langle \mathcal{T}_{\text{Einstein}} \cup [v \ll c] \rangle \approx \mathcal{T}_{\text{Newton}}.$$

This example is not as far from perception as it might seem. Certain aspects of motion perception, for example, make assumptions about the momentum of moving objects in the field of view. Clearly, the Newtonian expression is adequate in theories of motion perception, simply because in the environment in which the visual system operates, this expression is almost perfectly true. In other words, the refined theory assumed by visual system tacitly—but quite reasonably—assumes that objects don't move at near the speed of light.

For a more perceptual example, consider our caveman Ugg and his poisonous hemlock. Say that in the universe in general there is no relationship between color and edibility; but that inside Ugg's valley (the Boolean predicate `in_valley` set to true) blue fruits are poisonous and yellow fruits edible, while outside the valley the reverse is true. Hence by hypothesis the universal theory $\mathcal{T}_0$ contains the following set of sentences:

$$\begin{aligned}
\texttt{fruit}(x) \wedge \texttt{blue}(x) \wedge \texttt{in_valley} &\Rightarrow \neg\texttt{edible}(x),\\
\texttt{fruit}(x) \wedge \neg\texttt{blue}(x) \wedge \texttt{in_valley} &\Rightarrow \texttt{edible}(x).\\
\texttt{fruit}(x) \wedge \texttt{blue}(x) \wedge \neg\texttt{in_valley} &\Rightarrow \texttt{edible}(x),\\
\texttt{fruit}(x) \wedge \neg\texttt{blue}(x) \wedge \neg\texttt{in_valley} &\Rightarrow \neg\texttt{edible}(x).
\end{aligned}$$

Now, Ugg lives in the valley, so for him the predicate `in_valley` is true. Hence although the universal theory $\mathcal{T}_0$ does *not* entail the rule "blue fruit are poisonous,"

$$\mathcal{T}_0 \not\Rightarrow [\texttt{blue}(x) \wedge \texttt{fruit}(x) \Rightarrow \neg\texttt{edible}(x)],$$

Ugg's refined theory $\mathcal{T}_{\text{Ugg}} = \langle \mathcal{T}_0 \cup \texttt{in\_valley} \rangle$ *does* entail this rule:

$$\langle \mathcal{T}_0 \cup \texttt{in\_valley} \rangle \Rightarrow [\texttt{blue}(x) \wedge \texttt{fruit}(x) \Rightarrow \neg\texttt{edible}(x)].$$

That is, the refinement to Ugg's own environment justifies Ugg's inference rule.

Now we need to abstract a bit and see how multiple theories relate. When there exists some refinement $S$ that allows one theory $\mathcal{T}_1$ to entail another theory $\mathcal{T}_2$, we write

$$\mathcal{T}_1 \quad {}_S\, \mathcal{T}_2,$$

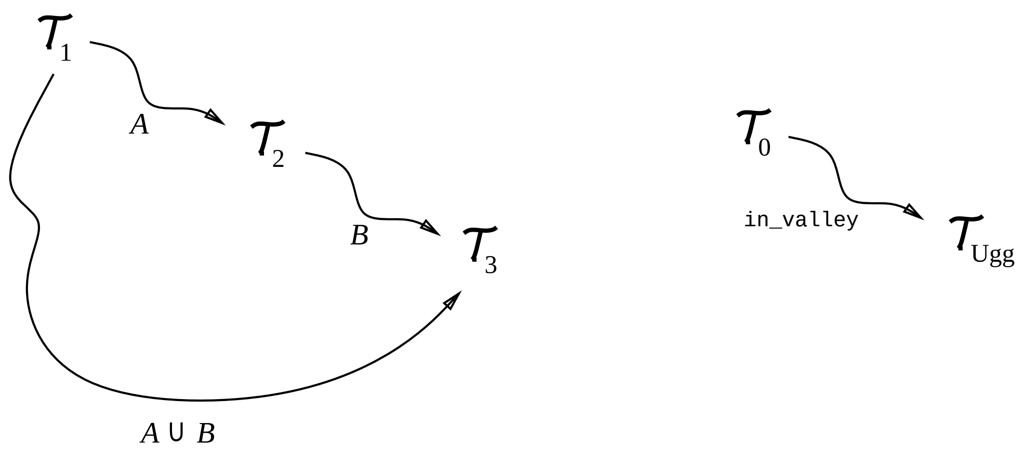



Figure 5: Schematic illustrating refinement. On the left, the general situation, showing how broad theories are refined into specific ones. On the right, Ugg's situation, showing how the universal theory $\mathcal{T}_0$, when augmented by the true predicate `in_valley`, leads to the Ugg's refined theory, which justifies his rule *blue implies poisonous*.

sometimes dropping the subscript when the refinement used is obvious. Clearly, the relation    is transitive, i.e.

$$(\mathcal{T}_1 \quad \mathcal{T}_2) \wedge (\mathcal{T}_2 \quad \mathcal{T}_3) \Rightarrow (\mathcal{T}_1 \quad \mathcal{T}_3).$$

In fact, one can collect refinements along the chain using union:

$$(\mathcal{T}_1 \quad {}_A\, \mathcal{T}_2) \wedge (\mathcal{T}_2 \quad {}_B\, \mathcal{T}_3) \Rightarrow (\mathcal{T}_1 \quad {}_{A \cup B}\, \mathcal{T}_3).$$

That is, refining a very broad theory to a much more specific one in many little steps is equivalent to refining it all the way in one big step which is the union of all the little steps. Fig. 5 illustrates the general situation, as well as Ugg's situation.

The important thing about this is that the refinement represents the additional, local knowledge—above and beyond universal laws—that is required to justify Ugg's perceptual rules. Exactly how general versus how specific this knowledge must be—or is, in practice—is not captured by the formalism, nor does it need to be. It might concern the environment only in the perceiver's own backyard, as in the Ugg example, but it might as well encode some physical fact which has been stable on Earth as long as the human cognitive system has been evolving. For the moment we are concerned only with the logical role this knowledge plays in justifying the perceiver's inferences, not its substance.

*When is a perceptual belief justified?* Having laid out the language of theories, we are at last in a position to state a justification—a truth condition—for an arbitrary perceptual rule. A perceptual rule $R$ of the form $A \Rightarrow \alpha$ is justified for a perceiver $P$ if, for some theory $\mathcal{T}$ and refinement $S$,

**(Justification Condition)**

**(i)** $\mathcal{T}_0 \quad {}_S\, \mathcal{T}$

**(ii)** $S$ holds for perceiver $P$

**(iii)** $\mathcal{T} \Rightarrow R$.

In this case, the refined theory of the world held by the perceiver is true, and this theory supports the rule in question; hence the rule is justified. The isotropic texture rule discussed earlier is justified for observers (such as ourselves, approximately) who live in a world with isotropic textures. Ugg's fruit rule is justified because he lives in a valley with poisonous hemlock. When he moves outside the the valley—and the predicate `in_valley` becomes false—the rule is no longer justified. Again notwithstanding the formalism this is all in accordance with common sense.

It is obvious that perceivers don't hold the full theory of the world $\mathcal{T}_0$ in their heads. Quite to the contrary, the mechanism in the head may consist entirely of a very impoverished implementation of the heuristic rules they use. Nevertheless the full theory $\mathcal{T}_0$ does play a causal role in establishing the *truth* of the perceiver's beliefs—and hence plays a central role in the semantics of those beliefs.

# 6 Minimal theories

We've seen how holding a certain theory justifies using a corresponding set of perceptual rules, supplying the "truth conditions" one wants as part of a semantics. But what about "meaning"? Given a perceptual rule, which let's assume to be a purely syntactic or formal construct: what is the meaning behind the rule?

This question is essentially the reverse of the first question. Above, we used a theory to justify a rule. Now we will see how a rule implies the assumption of a certain theory behind it. In a sense, the implied theory gives *meaning* to the rule—in the fairly literal sense of providing a pointer to the kind of world that would justify it. That is, the implied theory supplies the semantics behind the inference.

But *which* theory is implied by a given set of perceptual rules? It seems pretty clear that *many* theories would be consistent with any rules one might pick. For example, consider again Ugg's rule for fruit:

$$\forall x\ \texttt{blue}(x) \Rightarrow \neg\texttt{edible}(x),$$

"if it is blue then it is poisonous." What type of world would support such a rule? Clearly, one such world is one in which there exist poisonous blue hemlock and edible yellow bananas, and nothing else. Such a universe might be defined by the axiom set $\mathcal{A}_1$:

$\mathcal{A}_1 = \{$

$$\exists x \; \texttt{hemlock}(x)$$
$$\forall x \; \texttt{hemlock}(x) \Rightarrow \texttt{blue}(x)$$
$$\forall x \; \texttt{hemlock}(x) \Rightarrow \neg\texttt{edible}(x)$$

$\}$

(in part; for brevity I'll omit the part about bananas). This miniature universe is what logicians would call a *model* of the heuristic rule: a concrete object that satisfies it. Notice that simply postulating the existence of poisonous hemlock is *not* enough to justify Ugg's rule, since there might well be other types of blue plants that were *not* poisonous. In addition we need some version of what computational logic people call the *Closed World Assumption* (Reiter 1978; see Feldman 1997b for applications to perception). For our purposes this is simply the assumption that there are no entities in our theory other than those explicitly assumed—e.g., no other kinds of fruit other than hemlock and bananas. With this general assumption and the above axioms, it follows that all blue things are not edible, justifying Ugg's inferences.

But $\mathcal{A}_1$ is by no means the only set of axioms that justifies the rule. For example, consider the following universe $\mathcal{A}_2$:

$\mathcal{A}_2 = \{$

$$\exists x \; \texttt{hemlock}(x)$$
$$\forall x \; \texttt{hemlock}(x) \Rightarrow \texttt{blue}(x)$$
$$\forall x \; \texttt{hemlock}(x) \Rightarrow \texttt{flangible}(x)$$
$$\forall x \; \texttt{flangible}(x) \Rightarrow \neg\texttt{edible}(x)$$

$\}$

This theory invents a new (hidden) predicate, `flangible`, which intervenes between `hemlock` and ¬`edible`. Fig. 6 illustrates the situation.

The problem is that the second theory has additional structure over and above the first theory which adds little—an additional property with no new information. Clearly, by Occam's razor, it makes sense to prefer the simpler theory.

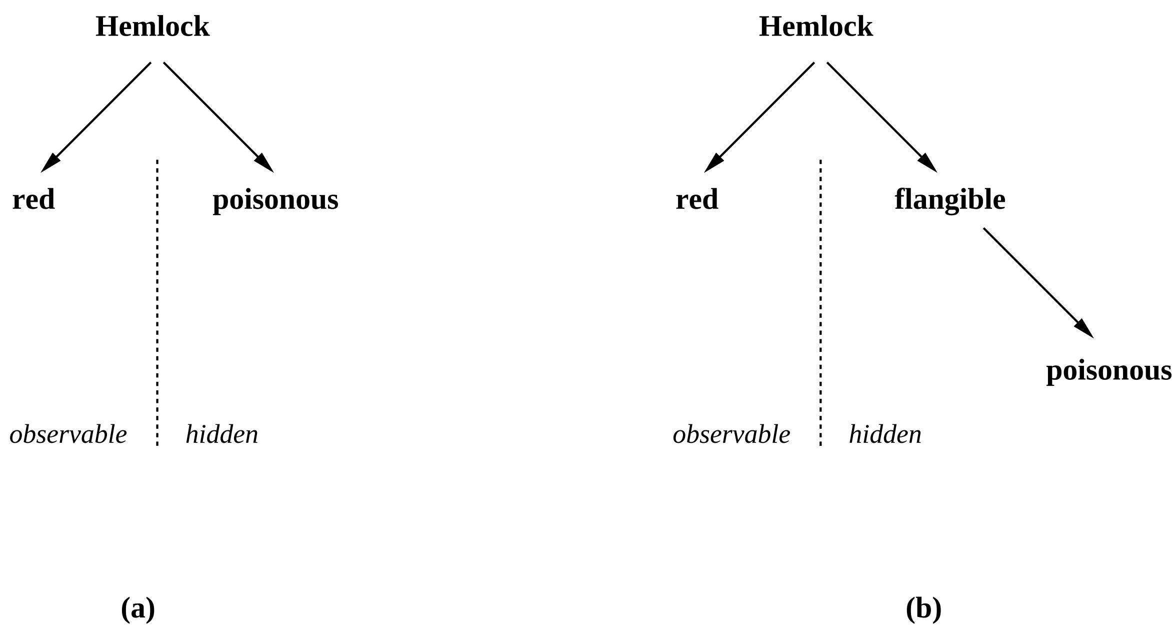



Figure 6: **(a)** A minimal theory of Ugg's edibility rule **(b)** A non-minimal theory.

However, there is a subtlety here. $\mathcal{A}_1$ is clearly simpler than $\mathcal{A}_2$ because it is a proper subset of it. But there is no way to prove that $\mathcal{A}_1$ is *absolutely* minimal, because there may well be other universes, perhaps involving other types of terms or statements, that are neither subsets nor supersets of $\mathcal{A}_1$. For example, it seems natural to postulate the existence of particular tangible entities that bear the properties that you observe, as is done in these two theories. But there may well be other, quite different ways of describing universes.

One solution is to restrict the language in which universes may be described. Unfortunately, there is no absolutely optimal way of doing this. Different choices about what kinds of universe underlie the observations simply lead to different conceptions of the "meaning" of those observations. I will refer to each such choice—each axiomatic description language for universes—as a *semantic language*. This allows us to state the minimality condition explicitly.

**(Minimality Condition)**

> Given a semantic language $\Sigma$, and a perceptual rule $R$, for all theories such that $\mathcal{T} \Rightarrow R$, use the one that is minimal under $\Sigma$.

Some examples are given in the next section.

## 7 The Entities-Properties semantics

A semantic language makes axiomatic assumptions about the structure and organization of the "real world" that underlies the observations a perceiver makes. For example the two theories $\mathcal{A}_1$ and $\mathcal{A}_2$ given above stipulate certain types of entities and their associated properties, and nothing else. Simple and spare as this is, it seems like a good starting place for a semantics.

**(Entities-Properties Semantics)**

Theories are constructed from the following two types of sentences:

(i) *Existential statements* of the form

$\exists x\ \lambda(x)$

(ii) *Property statements* of the form

$\forall x\ \lambda(x) \Rightarrow A(x)$ [observable properties]

$\forall x\ \lambda(x) \Rightarrow \alpha(x)$ [hidden properties]

To understand this conception of semantics, consider again the fruit rule "if it is blue then it is poisonous." Such a rule is purely *syntactic*; it expresses a relationship between one property and another. It does not actually say anything at all about the world per se. From a purely logical point of view, in fact, such a rule is valid in a world without *any* entities (since it is vacuously satisfied). Nevertheless it is very hard to think about this rule without in some sense inferring or assuming that there are blue, poisonous fruit lurking behind it. I propose that this conception is, in effect, the *meaning* behind the rule. But it is important to understand that this apparently implied reality is *not directly entailed* by the rule by itself. Once coupled with a semantic language *and* the minimality condition, however, the blue poisonous fruit are an immediate consequence. This, then, is the main idea: the "meaning" of a perceptual rule is the minimal theory that entails it, modulo a fixed conception of semantics.

In fact, we can state the relationship between a perceptual rule and its minimal theory in a stronger way. Because the theory entails the rule, and (assuming a fixed semantics *and* assuming minimality) the rule entails the theory, from a logical point of view the rule and the theory are *equivalent.* This is a surprisingly strong conclusion. An observer may only have certain reflexive reasoning rules explicitly encoded, but these rule may be logically equivalent to a substantive theory of the environment. This renders moot much discussion about whether the organism "really knows" about this or that aspect of the physical world; such knowledge may be equivalent to the inferential strategies the organism demonstrably uses. This is especially interesting when, as is the case with much of the visual system, the inferential system is thought to be largely hard-wired. Our hard-wired computational system is in essence equivalent to a particular theory of the world, so we can in effect think of this theory as itself hard-wired.

Let's take a more perceptual example. Consider the perception of *collinearity* (Fig. 7). It has long been known that the visual system pays special attention to elements of images that are collinear with another, i.e. that appear to fall along a line or curve. Much research has been devoted to the process whereby such patterns are processed (for example see Glass 1969; Foster 1979; Smits and Vos 1987; Pizlo, Salach-Golyska, and Rosenfeld 1997; Feldman 1997a). Normally such rules are purely syntactic: they indicate that when the geometry of the image elements has certain properties the elements should be grouped together. Now tacitly, we assume that such rules make sense because elements with that sort of geometry are likely to indicate a particular kind of entity out there in the world, namely an object contour or some other kind of coherent curvilinear object. In fact in computer vision, where researchers are more oriented toward real-world situations that require the use of perceptual grouping, this inference is sometimes spelled out explicitly (e.g. see Parent and Zucker 1989). But in psychological research the distal curve usually plays no formal role in grouping rules, although it is often conspicuous in *in*formal elucidations and justifications for various candidate rules. The goal of semantics is to bring this informal role out in the open.

The Entities-Properties semantics captures the essence of the inference in a canonical act of perceptual grouping. Loosely, the rule "group collinear items" entails (again, under this semantics and assuming minimality) a universe that contains curvilinear objects, elements of which will be nearly collinear with each other when observed in the image. This is the mental model underlying the grouping rule, and I suggest that this is the "meaning" of the raw perceptual agglomeration. Moreover, as argued above, in a certain sense the perceptual rule can be thought of as equivalent to the belief that there exist contours in the environment. The perception of collinearity is very basic in our visual system because the existence of contours is a very stable aspect of the physical world around us.

Again, though, it must be kept in mind that this relationship is strictly dependent on the choice of semantics and the assumption of minimality. Without minimality, or with a different semantic language, the same perceptual rules might entail a different theory of the environment.

*Causal semantics.* It is important to understand that the Entities-Properties Semantics is a very limited conception of the universe; at best it's a starting place. There is no dynamism, no relationship between different entities, no teleology—all things that one might conceive of applying to the world around us. One alternative is to center the semantics on the concept of *causality*, as has been advocated by Michael

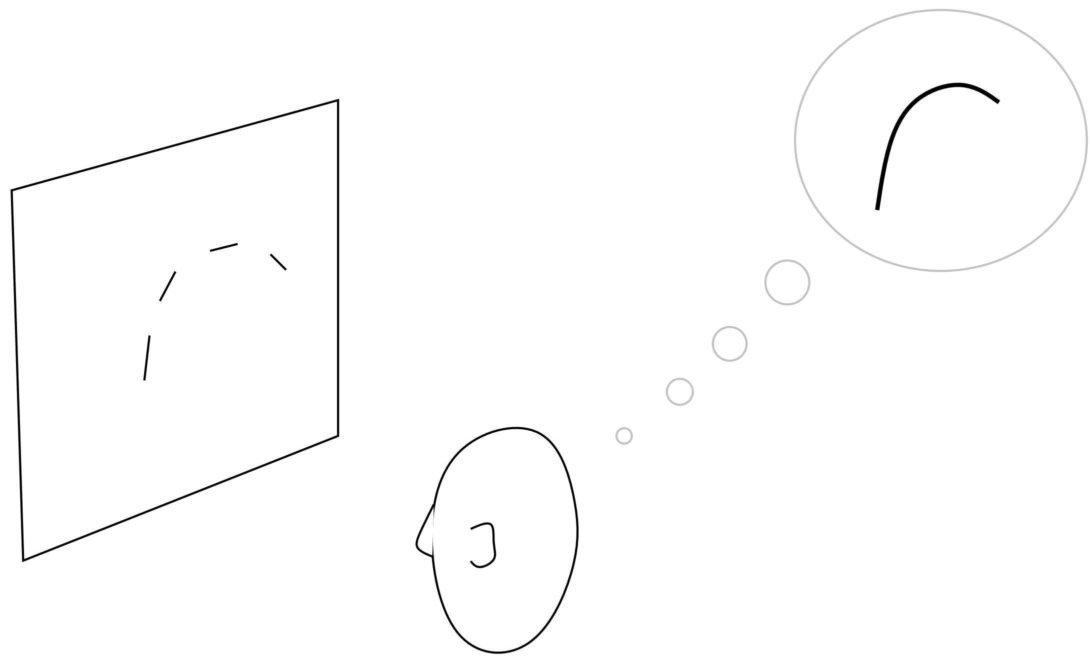


Figure 7: A series of collinear elements is interpreted as a distal curve or contour in the world.

Leyton (see Leyton 1992 and Ch.XX [Leyton] of this volume) as well as the author (Feldman 1997c).

In a causal semantics, special status is given to the historical sequence of transformations that brought the observed properties into existence. The assumption is that the world began in some initial state, and was then transformed by some operations into its current state, leading to the observed values of variables. Rather than just inferring a static world with certain properties, here the perceiver infers the relevant history of the world, thus placing additional meaning on observations in light of how they (apparently) came to achieve their current structure. The same perceptual rule implies a different universe when the world is conceived causally. Of course, other semantics are possible.

# 8 Does perception need semantics?

So what does all this mean for our Naive Perceiver? She might be surprised at the complexity of the machinery required simply to assure her that she can believe her eyes. But after all, if one wants to establish that that perceptual beliefs are well-founded, one must at some point make a connection between the state of the world and the content of those beliefs. The problem with doing this is that the world is made of stuff and things, but beliefs are made of representations and symbols—a type mismatch. To say in a rigorous way that a certain perceptual belief is true or false, we must relate it to the stuff "out there," but we can't, at least not directly. Instead what I've proposed in this chapter is to replace the world with an abstract theory of the world—something we *can* relate a propositional belief to explicitly. The circuit is closed entirely within the logical realm. The real world is only relevant to the extent it is well-described by theory.

But do we really need to establish that perceptual beliefs are well-founded? That is, do we need semantics? As I've suggested, vision researchers routinely include semantic characterizations in their discussions, but not usually as an explicit part of their models, but rather in a casual manner in order to suggest why a particular perceptual rule might be useful to the perceiver. Yet it is clear that such characterizations are a part of the complete story, if we would like a theory of perception which has explanatory as well as descriptive value. Certainly, if one is interested in the adaptive value of perceptual mechanisms, one needs to show how those mechanisms benefit the organism. Since presumably the benefit of perception is in yielding true beliefs, one cannot fully explain perception without showing why the beliefs it produces tend in fact to be true—no more than one would try to explain the digestive system without mentioning nutrition. And one cannot do this without bringing the world into the equation in a thorough and systematic fashion.

The truth-conditional part of this chapter, culminating in the Justification Condition, represent what is really a fairly straightforward attempt to do this. First I proposed a canonical logical form for perceptual rules, and then attempted to formalize the notion that a given rule is justified in a given universe if the validity of that rule is entailed by a theory that describes that universe. Carrying all this out in a formal rather than ad hoc manner may seem difficult, but it adds clarity. In the past, casual and inexplicit intuitions about what assumptions are actually required to justify particular perceptual inferences have occasionally led to frank errors, such as the fallacy of non-accidentalness (see Jepson and Richards 1992 for a discussion).

The other part of this chapter, concerning meaning and minimal theories, is somewhat more subtle. Here the goal is not simply to establish that the content of perceptual inferences tends to be true, but rather to *identify* that content. The key is simply to recognize that there are many underlying theories of the world that might lead one to use a given perceptual rule or mechanism. The minimum rule suggested here picks one of these theories out and gives it special status. I have argued that this minimal model is in effect the meaning of the rule, because it names the world that is pointed to by the rule—the reality behind the curtain, so to speak. The position I have implicitly taken here is that defining perceptual content is essentially a technical question, rather than a purely philosophical one: the content or meaning of a particular perceptual inference is a *model* of that inference in the logical sense, i.e. a well-defined mathematical object. This mathematical object seems to have a certain prima facie role in the formation of perceptual belief. Whether it has any deeper psychological validity remains a question to be explored.

**Author's notes**

I am grateful to Ernie Lepore, Zenon Pylyshyn, Whitman Richards, and Chuck Schmidt for many interesting discussions.

## References

Barlow, H. (1991). Vision tells you more than "what is where". In A. Gorea (Ed.), *Representations of vision: trends and tacit assumptions in vision research*. Cambridge: Cambridge University Press.

Barlow, H. (1994). What is the computational goal of the neocortex? In C. Koch and J. L. Davis (Eds.), *Large-scale neuronal theories of the brain*, pp. 1–22. Cambridge: M.I.T. Press.

Barlow, H. B. (1974). Inductive inference, coding, perception, and language. *Perception 3*, 123–134.

Barlow, H. B. (1990). Conditions for versatile learning, Helmholtz's unconscious inference, and the task of perception. *Vision Research 30*(11), 1561–1571.

Bennett, B. M., D. D. Hoffman, and C. Prakash (1989). *Observer mechanics: A formal theory of perception*. London: Academic Press.

Feldman, J. (1997a). Curvilinearity, covariance, and regularity in perceptual groups. *Vision Research 37*(20), 2835–2848.

Feldman, J. (1997b). Regularity-based perceptual grouping. *Computational Intelligence, 13*(4), 582–623.

Feldman, J. (1997c). The structure of perceptual categories. *Journal of Mathematical Psychology, 41*, 145–170.

Fodor, J. (1980). Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences 3*, 63–73. Reprinted in J. Fodor (1981), *Representations: philosophical essays on the foundations of cognitive science*, Cambridge, MA: M.I.T. Press.

Foster, D. H. (1979). Discrete internal pattern represetnations and visual detection of sall changes in pattern shape. *Perception and Psychophysics 26*(6), 459–468.

Genesereth, M. R. and N. J. Nilsson (1987). *Logical foundations of artificial intelligence*. Palo Alto, CA: Morgan Kauffman.

Glass, L. (1969). Moiré effects from random dots. *Nature 223*, 578–580.

Jepson, A. and W. A. Richards (1992). What makes a good feature? In L. Harris and M. Jenkin (Eds.), *Spatial Vision in Humans and Robots*. Cambridge University Press.

Leyton, M. (1992). *Symmetry, Causality, Mind*. Cambridge, MA: M.I.T. Press.

Mackworth, A. K. (1988). Adequacy criteria for visual knowledge representation. In Z. Pylyshyn (Ed.), *Computational processes in human vision: an interdisciplinary perspective*, pp. 462–474. Norwood, N.J.: Ablex.

Marr, D. (1982). *Vision: a computational investigation into the human representation and processing of visual information*. San Francisco: Freeman.

Parent, P. and S. W. Zucker (1989). Trace inference, curvature consistency, and curve detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence 11*(8), 823–839.

Pizlo, Z., M. Salach-Golyska, and A. Rosenfeld (1997). Curve detection in a noisy image. *Vision research 37*(9), 1217–1241.

Reiter, R. (1978). On closed world data bases. In H. Gallaire and J. Minker (Eds.), *Logic and data bases*. New York: Plenum Press.

Reiter, R. and A. K. Mackworth (1989). A logical framework for depiction and image interpretation. *Artificial Intelligence 41*, 125–155.

Richards, W. A. and A. Bobick (1988). Playing twenty questions with nature. In Z. Pylyshyn (Ed.), *Computational Processes in Human Vision: An Interdisciplinary Perspective*, pp. 3–26. Norwood, NJ: Ablex Publishing Corporation.

Richards, W. A., J. M. Rubin, and D. D. Hoffman (1982). Equation counting and the interpretation of sensory data. *Perception 11*, 557–576.

Shepard, R. N. (1989). Internal representation of universal regularities: a challenge for connectionism. In L. Nadel, L. A. Cooper, P. Culicover, and R. M. Harnish (Eds.), *Neural connections, mental computation*, pp. 104–134. M.I.T. Press.

Smits, J. T. and P. G. Vos (1987). The perception of continuous curves in dot stimuli. *Perception 16*(1), 121–131.