8

Action-Oriented Representation

Pete Mandik

Introduction

Often, sensory input underdetermines perception. One such example is the perception of illusory contours. In illusory contour perception, the content of the percept includes the presence of a contour that is absent from the informational content of the sensation. (By "sensation" I mean merely information-bearing events at the transducer level. I intend no further commitment, such as the identification of sensations with qualia.) I call instances of perception underdetermined by sensation "underdetermined perception."

The perception of illusory contours is just one kind of underdetermined perception (see Figure 8.1). The focus of this chapter is another kind of underdetermined perception: what I shall call "active perception." Active perception occurs in cases in which the percept, while underdetermined by sensation, is determined by a combination of sensation and action. The phenomenon of active perception has been used by several to argue against the positing of representations in explanations of sensory experience, either by arguing that no representations need be posited or that far fewer than previously thought need be posited. Such views include, but are not limited to, those of J. Gibson (1966, 1986), P. S. Churchland et al. (1994), T. Jarvilehto (1998), and J. O'Regan and A. Noë (2001). In this chapter, I argue for the contrary position that active perception is actually best accounted for by a representational theory of perception. Along the way, this will require a relatively novel conception of what to count as representations. In particular, I flesh out a novel account of *action-oriented*

FIGURE 8.1. Illusory contours. *Source*: Figure drawn by Pete Mandik.

representations: representations that include in their contents commands for certain behaviors.[1]

Examples of Active Perception

A somewhat famous and highly fascinating example of active perception is shown in the experiences of subjects trained in the use of P. Bach-y-Rita's (1972) Tactile Visual Sensory Substitution System (TVSS). The system consists of a head-mounted video camera that sends information to an array of tactile stimulators worn pressed against the subject's abdomen or back. The subjects can aim the camera at various objects by turning their heads and can adjust the zoom and focus of the camera with a handheld controller. Blindfolded and congenitally blind subjects can utilize the device to recognize faces and objects. Especially interesting are the ways in which the TVSS approximates natural vision. Subjects reported losing awareness of the tingles on their skin and instead saw through the tactile array, much in the same way that one loses awareness of the pixels on a television screen and instead sees through it to see actors and scenery. Bach-y-Rita reports an incident in which someone other than the subject wearing the device increased the camera's zoom. The subject ducked, since the zoom effect made objects seem as if they were heading toward the subject. Bach-y-Rita notes that these sorts of reports only occurred for subjects whose training with the TVSS involved the active control of the camera's direction, focus, and zoom. In conditions in which the subjects had no control over these features and instead only passively received the video-driven tactile information, the subjects never reported the phenomenon of seeing through the tingles on their skin to locate the perceived object in the external environment. For these reasons, then, experiences with the TVSS count as instances of active perception. Information provided at the skin by the tactile stimulators is insufficient to determine the perception of distal objects. The determination of the percept occurs only when certain contributions from action are combined with the tactile input.

---

[1] I did not coin the term "action-oriented representation," although I am unsure of what its first appearance in the literature was. See Clark (1997) and Colby (1999) for discussions of action-oriented representation.

An even simpler, "chemically pure," example of this sort of TVSS-based active perception is reported by C. Lenay et al. (1997) and S. Hanneton et al. (1999). Subjects use a tactile-based device to identify simple two-dimensional forms, such as broken lines and curves. The subjects wear a single tactile stimulator on a fingertip. The stimulator is driven by a magnetic pen used in conjunction with a graphic tablet. A virtual image in black-and-white pixels is displayed on a screen that only the experimenter is allowed to see. The subject scans the pen across the tablet and thus controls a cursor that moves across the virtual image. A stimulus is delivered to the fingertip only when the cursor is on pixels that make up the figure and not on background pixels. Subjects with control over the pen are able to identify the images. Subjects who merely passively receive the tactile information cannot.

One caveat should be stated concerning the proposal that both of these cases count as instances of underdetermined perception. If the sensory inputs are only the tactile inputs, then these are relatively clear cases of underdetermined perception. However, the contribution of action may be sensational if the contribution is exhausted by sensory feedback from the muscles. If this latter possibility obtains, then we have cases of relative, not absolute, underdetermined perception, since the percept would be underdetermined only relative to the tactile input. However, if the contribution of action is, say, an efference copy instead of sensory feedback from the muscles, then the cases are absolute cases of underdetermined perception. I postpone for now further discussion of the distinction between relative and absolute active perception.

A Challenge Posed to Representational Theories

What I am calling active perception has been alleged by others to undermine, either partially or totally, the representational theory of sensory perception. But how, exactly, is this undermining supposed to take place? Before answering this question we must first answer another: what is the representational theory of perception?

Many and various things have been written about the representational theory of perception – enough, perhaps, to render suspicious any claims that there is such a thing as *the* representational theory of perception.[2] However, the theory I sketch here will have sufficient detail to both serve

---

[2] The theory is also known in the philosophical literature as the causal theory of perception. See, for example, Grice (1961), Oakes (1978), and Hyman (1992).

the purposes of the current chapter and do justice to the main features common to typical explications of perception in representational terms. The representational theory of perception may be crudely characterized as the view that one has a perceptual experience of an *F* if and only if one mentally represents that an *F* is present and the current token mental representation of an *F* is causally triggered by the presence of an *F*.[3] There are thus two crucial components of this analysis of perception: the representational component and the causal component. The purpose of the representational component is to account for the similarity between perception, on the one hand, and imagery and illusion, on the other. As is oft noted at least since Descartes, from the first-person point of view accurate perceptions can be indistinguishable from dreams and illusions. This similarity is classically accounted for by the hypothesis that veridical mental states such as perceptions are representational. They are thus hypothesized to differ from their nonveridical counterparts (dreams and hallucinations) not in whether they are representations but in whether they are *accurate* representations.[4] The causality component in the account of perceptual experience is a further articulation of the idea that in spite of similarities, there are crucial differences between perceptions and other representational mental phenomena. It is thus part of the normal functioning of perceptions that they are caused by the things that they represent. Simply having a mental representation of, say, a bear is insufficient for perceiving the bear. The relevant mental representation must be currently caused by a bear to count as a percept of a bear.[5] Further, the causal component will have much to do with the specification of sensory modality. So, for example, if the causal processes intervening between the percept and the bear have largely to do with sound waves, then the perceptual event counts as hearing the bear, and if the causal

---

[3] More can be added to this analysis, of course. For example, if someone sneaks up behind me and hits me on the head with a hammer and this causes me to have a visual hallucination of a hammer, this wouldn't count as a visual perception of the hammer in spite of being a hammer-caused mental representation of a hammer. Additional criteria for perception would include, for example, specifications of the normal channels of causation, which were bypassed in the hammer example. However, attending to this level of detail in the analysis of perception is unnecessary for my present purposes.

[4] See J. Austin 1964 for a classic discussion and, of course, criticism of this line of thought oft referred to as the *argument from illusion*. I will not here review all of the various objections to this argument. My focus is instead to defend the representational theory of perception from attacks predicated on active perception.

[5] See Grice (1961) for an expanded discussion of the topic, including these sorts of causal conditions in the analysis of perception.

processes instead largely involve reflected light, then the perceptual event counts as seeing the bear.[6]

Typically, the notion of representation employed in the representation component is explicated in terms of the kinds of causal processes specified in the causal component. Certain causal relations that obtain between the percept and the thing that it is a percept *of* are thus brought in to explicate what it is for the percept to count as a representation. This is not to say that anyone believes in "The Crude Causal Theory" (Fodor 1987) that says a state represents *F*s if and only if it is caused by *F*s. It is instead to say that being caused by *F*s is going to be an important part of the story of what it is to represent *F*s. The typical kind of story of which the causal relation is a part is a kind of teleological story in which what it means to represent *F*s is to be in a state that is supposed to be caused by *F*s or has the function of being caused by *F*s or has been naturally selected to be caused by *F*s or is caused by *F*s in biologically optimal circumstances. (See, for example, Dretske 1995.)

This view of representation is perhaps the most widespread notion of representation used in the neurosciences. It underlies talk of detectors and instances in which something "codes for" a perceptible environmental feature. For example, there are claimed to be edge detectors in visual cortex (Hubel and Wiesel 1962) and face detectors in inferotemporal cortex (Perrett et al. 1989). Magnocellular activity codes for motion and parvocellular activity codes for color (Livingstone and Hubel 1988). Thus, from the neural point of view, being a representation of *F*s is being a bit of brain "lit up" as a causal consequence of the presence of such *F*s. The teleological element is brought on board to explain how *F*s can be represented even in situations in which no *F*s are present. The lighting up of the relevant brain bit represents *F*s because in certain normal or basic cases, *F*s would cause the lighting up of that bit of brain. This sort of view shows up in neuroscience in the popular account of imagery as being the off-line utilization of resources utilized on line during sensory perception. Thus, for example, the brain areas utilized in forming the mental image of an *F* overlap with the brain areas utilized in the perception of an *F*. S. Kosslyn et al. (2001) report that early visual cortex (area 17) is active in perception as well as imagery and that parahippocampal place area is active in both the perception and imagery of places. K. O'Craven and N. Kaniwisher (2000) report fusiform face area activation for both the

---

[6] For an extended discussion of the individuation of sensory modalities, see B. Keeley (2002).

imagery and perception of faces. In these sorts of cases, the main difference between imagery and perception of *F*s is that in imagery, unlike perception, no *F* need be present.

Another way in which this teleofunctional view of representation emerges in neuroscience is in explanations of underdetermined perception, such as the perception of illusory contours. Neuroimaging studies in humans show that illusory contours activate areas in striate and extrastriate visual cortex similar to areas also activated by real contours (Larsson et al. 1999). Additionally, orientation-selective neurons in monkey V2 also respond to illusory contours with the same orientation (von der Heydt et al. 1984, Peterhans and von der Heydt 1991).

The teleofunctional explanations of both imagery and illusory contours amount to what I shall term the nervous system's employment of a "recruitment strategy": Processes whose primary and original functions serve the perception of real contours get "recruited" to serve other functions. (S. Gould [1991] calls such recruitment "exaptation.") Viewing the nervous system as employing the recruitment strategy thus involves viewing it as conforming to the classical empiricist doctrine that nothing is in the mind that is not first in the senses. Further, it supplies an outline in neural terms of how that which is first in the senses can come to serve other cognitive processes.

The representational explanation of the perception of illusory contours helps to show that the representational theory has the resources to explain at least some cases of underdetermined perception. But the question arises of whether it has the resources to explain all cases of underdetermined perception, especially cases of active perception. Some theorists, such as Gibson (1966, 1986) and O'Regan and Noë (2001), have urged that it does not. O'Regan and Noë (2001) reject representational theories of vision: "Instead of assuming that vision consists in the creation of an internal representation of the outside world whose activation somehow generates visual experience, we propose to treat vision as an *exploratory activity*" (p. 940). According to O'Regan and Noë's alternative – their "sensorimotor contingency theory" – all visual perception is characterized as active perception, or, in their own words, "*vision is a mode of exploration of the world that is mediated by knowledge of what we call sensorimotor contingencies*" (p. 940, emphasis in original). I presume that O'Regan and Noë intend the *knowledge* of sensorimotor contingencies to not involve the *representation* of sensorimotor contingencies.

I will not here rehearse Gibson's or O'Regan and Noë's case against the representational theory, but instead sketch some general reasons why

viewing perception as active might be thought to pose a threat. In brief, the problem is that active perception highlights the importance of output while the representational story is told in terms of inputs. Recall that the notion of representation is explicated in terms of states that have the function of being caused by environmental events. Thus, the basic case of a representation is neural activation that occurs as a response to some sensory input. Active perception, however, is a kind of underdetermined perception – that is, perception underdetermined by sensory inputs. Further, what does determine the percept in active perception is a combination of the inputs with certain kinds of outputs. Since output seems to be bearing so much of the load, there seems to be little hope for a story told exclusively in terms of inputs. A further problem arises when we consider that it is not clear that the recruitment strategy is as readily available for active perception as it is for other kinds of underdetermined perception.

Illusory contour perception is subjectively similar to the perception of real contours. The reactivation of brain areas responsible for the perception of real contours gives rise to a subjective appearance similar to what is experienced when real contours are present. This is part of what it means to call illusory contour perception "illusory." If something similar were occurring in active perception, then we would expect an analogous tactile illusion. However, in the pen and tablet version of TVSS, the percept does not involve tactile illusion; that is, the subject doesn't feel the portions of the contours that are not currently being scanned. Given these sorts of considerations, the threat of active perception to the representational theory of perception seems to be two-pronged: The first prong criticizes the representational theory for being overly reliant on the contributions of input, and the second prong criticizes the representational theory for being overly reliant on the recruitment strategy.

Meeting the Challenge

Active perception poses an apparently serious threat to the representational theory of perception. However, this apparent seriousness should not be confused with hopelessness. On the contrary, a rather minor revision of the representational theory will suffice to ward off the threat. The revision concerns the conditions on being a representation and will include a role for output as well as input for determining representational contents.

To get the clearest possible grasp on this account of the representational basis of perception, it will be useful to consider the simplest possible

examples of a creature undergoing a fully determined visual perception. Imagine a creature that moves about a planar surface and utilizes a pair of light sensors – mounted on the creature's left and right, respectively – to orient toward sources of illumination. Sunlight is beneficial to various creatures in various ways, and thus positive phototaxis is a common example of an adaptive response to an environmental stimulus. In the two-sensor creature that we are imagining, activity in each sensor is a linear function of the local light intensity, and given a constant light source, degree of activation in the sensor represents proximity to the light source. Thus, the difference in the activity between the two sensors encodes the location of the light source in a two-dimensional egocentric space. Information encoded by the sensors can be relayed to and decoded by motor systems responsible for steering the creature. For example, left and right opposing muscles might have their activity directly modulated by contralateral sensors so that the greater contraction corresponds to the side with the greatest sensor activity, thus steering the creature toward the light. More complex uses of the sensory inputs would involve having them feed into a central processor that gives rise to a perceptual judgment that, say, the light is to the right. The example sketched so far constitutes an example of determined perception on the following grounds. If the perception is a state of the organism specifying the location in two-dimensional egocentric space of the light source, then this is a percept fully determined by the information encoded at the sensory transducers.

To see a simple example of underdetermined perception, in particular, an example of active perception, let us contrast the aforementioned case with a creature forced to make due with only a single light sensor. The single sensor only encodes information regarding proximity to the light source, and thus encodes information about only one dimension of egocentric location of the source. However, this does not prevent the creature from coming to know or coming to form a percept of the two-dimensional egocentric location of the distal stimulus. One way in which the creature might overcome the limitations of a single sensor is by scanning the sensor from left to right while keeping track of the direction in which it has moved the sensor. By comparing the reading of the sensor when moved to the right to the reading of the sensor when moved to the left, the creature thereby has access to information similar to the creature with two sensors. Here, two-dimensional location is encoded not in the difference between two sensors but, instead, in the difference between the activity occurring at two different times within the same sensor. In order to make use of this information, however, the

scanning creature needs some way of knowing when the sensor is in the left position and when the sensor is in the right position.

There are two general conditions in which the creature can accomplish this. In the first condition – the feedback condition – the creature receives sensory feedback regarding the states of its muscles. Thus, in the feedback condition, while the percept may be underdetermined by the input from the light sensor, it is not underdetermined by sensation altogether, since sensory input from the muscles, combined with the light sensor input, determines the percept. Thus, the feedback condition is only a case of relative, not absolute, underdetermined perception. In the second condition – the efference copy condition – the creature knows the position of the scanning organ by keeping track of what commands were sent to the scanning organ. Thus, in the efference copy condition, the percept is genuinely underdetermined by sensation, since what augments the sensory input from the light sensor is not some additional sensory input from the muscles, but instead a record of what the outputs were – that is, a copy of the efferent signal. If it is a requirement on active perception that it be underdetermined by sensation altogether, and not just underdetermined relative to some subset of the sensory inputs, then only the efference copy condition constitutes a genuine case of active perception. Thus, if so-called active perception is only relatively underdetermined, then it doesn't pose the kind of threat to the representational theory outlined earlier. There ultimately is adequate input information for the determination of the percept. However, as I will argue, even genuine (efference copy–based) active perception can be explained in terms of the representational theory of perception.

The representational theory of perception, although not defeated by active perception, will nonetheless require an adjustment. The adjustment required is to acknowledge that there are occasions in which outputs instead of inputs figure into the specification of the content of a representational state. I propose to model these output-oriented – that is, action-oriented – specifications along the lines utilized in the case of inputs. When focusing on input conditions, the schematic theory of representational content is the following: A state of an organism represents *F*s if that state has the teleological function of being caused by *F*s. I propose to add an additional set of conditions in which a state can come to represent *F*s by allowing that a reversed direction of causation can suffice. A state of an organism represents *F*s if that state has the teleological function of causing *F*s. Thus, in the single-sensor creatures described earlier, the motor command to scan the sensor to the left is as much

an adequate representation that something is happening to the left as is a sensory input caused by something happening to the left. Efference copies inherit their representational contents from the motor commands that they are copies of. Efference copies thus constitute an action-oriented version of the recruitment strategy. We are now in a position to define "action-oriented representation" as any representation whose content is determined, in whole or in part, by involving states whose teleofunction is to be the causal antecedents of actions. Another way to state the definition is that action-oriented representations are any representations that have, in whole or in part, imperative content.[7] Active perception thus does not threaten the representational theory of perception. Instead, it forces us to acknowledge that action-oriented representations can contribute to the representational content of perception, and further, that percepts themselves may sometimes be action-oriented representations.

I should note a point of contrast between the account of spatial content I articulate here and elsewhere (Mandik 1999, 2001, 2002, and 2003) and other action-involving accounts, such as G. Evans (1985) and especially R. Grush (2001, this volume). On Grush's "skill theory" of spatial content, certain behavioral dispositions are necessary for a mental state such as a percept to have spatial representational content. According to this view, it would thus be impossible for an organism to perceive a stimulus as being to the left without at the same time being able to orient toward that stimulus. On such a view, states at the input side of the cognitive system cannot by themselves carry spatial content; only states appropriately engaged with motor outputs count as genuinely representing spatial properties and relations. In contrast, though I grant that certain output-involving processes (such as motor commands and efference copies) are sufficient for spatial content, I reject the claim that they are thereby necessary. There are many varieties of spatial representation, only some of which significantly engage motor processes. (See Mandik 2003 for a longer discussion of these varieties of representation.)

Now that the representational account of active perception has been sketched, I devote the rest of the chapter to the following three questions. First, is the solution sketched feasible? That is, is it possible to employ it as an engineering solution to the problem of utilizing action to

[7] This contrasts with the way A. Clark (1997) defines action-oriented representations. For Clark, action-oriented representations always have both imperative and indicative content and are thus the same as what R. Millikan (1996) calls "Pushmi-Pullyu Representations." On my definition, even representations with only imperative content (e.g., motor commands) are action-oriented representations.

compensate for impoverished inputs? Second, is the solution sketched evolvable? Given the reliance on evolution in typical versions of the teleofunctional portion of the story sketched earlier, it remains a serious question whether the sort of incremental adaptations posited in most evolutionary scenarios could possibly give rise to such a solution. Third, even if feasible and evolvable, are such solutions actually instantiated in human nervous systems?

Is the Action-Oriented Solution Feasible? A Reply from Robotics

If the action-oriented representation solution described so far is indeed feasible, then it ought to be possible to construct a robotic model that employs such principles to exhibit perceptually guided adaptive behaviors. In Mandik (1999), I discuss a thought experiment about an imaginary robot named Tanky that traverses a planar surface by means of tank treads. Various patches of the surface are considered either nutritious or noxious to Tanky, and while poised over one of these patches, Tanky's chemoreceptors can indicate as much. However, the chemoreceptors are alone insufficient to give Tanky much information about the spatial arrangement of the various proximal and distal chemical patches in its environment. Tanky's perceptual contact with the spatial features of its environment is mediated through the tank treads, thus implementing a form of odometry. There are two general ways in which this odometry might be accomplished to give Tanky knowledge of, for instance, the distance between the chemical patch it currently perceives and the last patch it visited. The first way is akin to the feedback solution described earlier, whereas the second way is akin to the efference copy solution. On the feedback solution, distance estimates measured in numbers of tank tread revolutions are updated in virtue of information from a sensor that counts actual tank tread revolutions. In contrast, the efference copy solution forgoes sensor information and instead involves the counting of the number of commands sent to revolve the tank treads.

In Mandik (1999), I hypothesized that both solutions would be equally adequate to provide Tanky with a perception of the spatial arrangement of chemical patches in its environment. Since 1999, however, I have had many occasions to experiment with real robots and discovered, among other things, that odometry and tank tread locomotion don't mix very well due to the high degree of slippage where "rubber meets the road" necessary to effect steering in a treaded vehicle. Nonetheless,
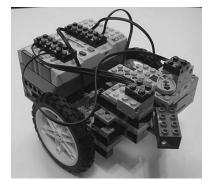
FIGURE 8.2. The robot "Tanky Jr."

real robots offer ample opportunities to demonstrate the viabilities of the feedback and efference copy solutions to spatial perceptual underdetermination. I constructed Tanky Jr. (depicted in Figure 8.2) using the LEGO® MINDSTORMS™ robot kit and programmed the robot using David Baum's (2002) third-party programming language, NQC (Not Quite C).[8] Tanky Jr. is an experimental platform for implementing strategies of positive phototaxis utilizing a single light sensor combined with the kinds of scanning strategies described earlier. Tanky Jr. has three motors: Two drive the left and right wheels, respectively, and the third is utilized to scan Tanky Jr.'s single light sensor left and right.

To implement a feedback strategy to monitor the position of the scanning light sensor, Tanky Jr. has as additional inputs two touch sensors mounted to the left and right of the light sensor. When the robot is first turned on, its wheels remain stationary while it performs a scanning procedure. The first part of the scanning procedure is to scan the light sensor to the right until the touch sensor dedicated to that side is activated. The program then updates a variable that serves as a record of the light sensor activity at that position. Next, the sensor is scanned in the opposite direction until the other touch sensor is activated. The reading of the light sensor in this position is then compared to the previous reading. If the difference in the light readings from the left and right positions are relatively negligible, the robot then moves straight ahead a short distance; otherwise, the robot will turn a bit in the direction of the greatest light reading before making its forward motion. The robot then stops and begins another run of the scanning procedure. The alternating repetition

[8] For a nice overview of the philosophical uses of robots as tools for both research and pedagogy with special focus on the LEGO® MINDSTORMS™ system, see J. Sullins (2002).

of these steps is quite effective in getting the robot to move toward a light stimulus, such as a spot of light shone on the floor from a flashlight.

Equally successful is a strategy that forgoes sensory feedback in favor of efference copies. In this latter condition, Tanky Jr.'s touch sensors are removed and the program is altered so that the commands involved in the scanning procedure do not specify that the scanning motion be ceased when the touch sensors are activated, but instead be ceased after a fraction of a second. The left and right light sensor variables are updated not as a response to touch sensor feedback but instead as a response to a record of what commands have been sent. Thus is this latter strategy describable as implementing a system that utilizes efference copies. The equivalence in performance of the efference copy and feedback solutions shows that the efference copy solution is no less representational than the feedback solution.

Is the Action-Oriented Solution Evolvable?
A Reply from Artificial Life

Tanky Jr. shows the feasibility of the solution, although the neural feasibility has not yet been addressed. Also unaddressed until now is the question of whether the efference copy solution is evolvable. In Mandik (2002, 2003), I discuss several artificial life experiments I have conducted to evolve various kinds of neural network controllers for artificial organisms solving simple yet representationally demanding perceptual tasks. Typical experiments involved the modeling of legged land creatures traversing a planar surface. Survival and other estimations of fitness depend on the capacities of the creatures to utilize sensor information to find food distributed through the environment. In Mandik (2003, pp. 118–122), I describe experiments designed to coax the evolution of action-oriented representations in these neural controllers. The artificial creature "Radar" utilized in these latter experiments had the general body structure depicted in Figure 8.3 and the general neural network topology depicted in Figure 8.4. Body structure and neural topology were specified by hand. An evolutionary algorithm was employed to evolve specifications of the neural weights. Radar's forward locomotion is effected by four limbs, and steering is effected by a single bending joint in the middle of its body. Food is detected utilizing a single sensor mounted on a scanning organ that moves left and right in a manner similar to the scanner used by Tanky Jr.

The top layer in figure 8.4 depicts the portion of Radar's nervous system serving as a central pattern generator that sends a sinusoidal signal to
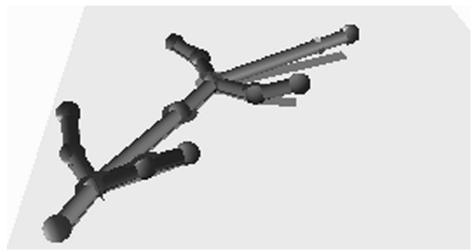
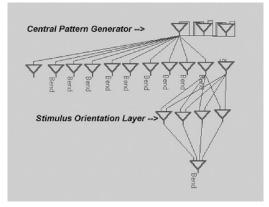FIGURE 8.3. The artificial life creature "Radar."

FIGURE 8.4. Radar's nervous system.

the muscles responsible for the forward walking motions as well as the left and right scanning of the sensor organ. Stimulus orientation is effected by a three-layer feed-forward network consisting of a two-neuron input layer, a four-neuron hidden layer, and a single-neuron output layer. The two inputs are the single food sensor and sensory feedback concerning the state of the scanning muscle. The four-unit hidden layer then feeds into the single orientation muscle. A second version of Radar's neural topology replaces the muscular feedback with an efference copy. Instead of receiving feedback from the scanning muscle, the hidden layer of the orientation network receives as input a copy of the command that the central pattern generator sends to the scanning muscle. Neural weights were evolved for three kinds of controller topologies: The first had orientation layer inputs from the sensor and the muscular feedback, the second had inputs from the sensor and an efference copy, and the third had only

sensor input without either muscular feedback or efference copies. On several occasions, populations with the first two topologies successfully evolved sets of neural weights that utilized both food sensor input and muscular input (either efferent copy or feedback) in order to maximize their life spans by finding food. However, I was somewhat disappointed to find that the efference copy and feedback conditions, while equally successful, did not consistently and significantly outperform the creatures that had only the single input from the food sensor feeding into the stimulus orientation layer.

To see what might be missing in the neural topologies to account for this result, it is instructive to compare Radar to Tanky Jr. When Tanky Jr. executes the scanning-procedure portion of the program, a crucial step involves using a single sensor to take two different readings – the left and right readings, respectively – of the local light levels. After the second of the two readings is taken, it is compared to a *memory record* of the first reading. This employment of a memory is, I suggest, the crucial difference between Tanky Jr. and Radar. The stimulus orientation network in Radar's nervous system is a three-layer *feed forward* network that lacks recurrent connections or any other means of instantiating a memory. In other words, it lacks the means of being sensitive to information spread out over time. But the task of comparing left and right readings gathered with a single scanning sensor is, crucially, a process that occurs over time. Therefore, future versions of Radar must incorporate some means (such as recurrent connections) of storing information about a previous sensor reading long enough for it to be compared to a current sensor reading.

While I have not yet experimented with versions of Radar that incorporate memory into the scanning procedure, in Mandik (2003, pp. 111–118), I discuss creatures that I have evolved to utilize memory in a similar task, namely, the comparison of a past and current stimulus. In these simulations, creatures with a single sensor did not scan it left and right but, however, did utilize it in a comparison between past and current stimuli by routing the sensor signal through two channels in the stimulus orientation network. One of the two channels passed its signal through more neurons, thus constituting a memory delay. The portion of the network that had to effect a comparison thus compares the current signal to a delayed signal. This can be part of an adaptive strategy for food-finding insofar as it, in combination with the tacit assumption that the creature is moving forward, allows the creature to draw something like the following inference: If the current value is higher than the

remembered value, then the creature must be heading toward the stimulus and should thus continue doing so, but if the current value is lower than the remembered value, then the creature must be heading away from the stimulus and must thus turn around. Such a use of memory has been shown to be used by *E. coli* bacteria to navigate up nutrient gradients (Koshland 1977, 1980).[9] This initial success with these artificial life simulations helps bolster the claim of the evolvability of the kinds of action-oriented representation solutions implemented in the robot Tanky Jr. Much remains open, however – in particular, the question to which I now turn: *Do human nervous systems utilize any action-oriented representations?*

Is the Action-Oriented Solution Instantiated in Human Nervous Systems? A Reply from Neuroscience

One especially promising line of evidence concerning whether efference copy–based action-oriented representations are employed in human nervous systems comes from research on visual stability during saccadic eye movements. The phenomenon to be explained here is how it is that we don't perceive the world to be jumping around, even though our eyes are constantly moving in the short jerky movements known as saccades. Hermann Helmholtz (1867) hypothesized that efference copies are used in the following manner. When the eye moves, there is a shift in the array of information transduced at the retina. When the eye movement is caused in the normal way – that is, by self-generated movements due to commands sent to ocular muscles – an efference copy is used to compute the amount to compensate for the anticipated shift in the retinal image. The amount of movement estimated on the basis of the character of the efference copy is thus used to offset the actual shift in retinal information, giving rise, ultimately, to a percept that contains no such shift.

This hypothesis implies that there should be a perception of a shift in cases in which the eye is moved in the absence of efference copies, as well as in cases in which efference copies are generated but no eye movement is produced. The first sort of case may be generated by eye movements produced by tapping or pushing on the eye. A quick way to verify this is to take your own (clean!) finger and gently push the side of your eye. Your eye is now moving with respect to the visual scene in a manner actually less extreme than in many saccadic motions. However, the instability of

---

[9] For further discussion of artificial life simulations involving neural representation of information about the past, see Mandik 2002, pp. 14–15, and Mandik 2003, pp. 111–118.

the visual scene – it jumps dramatically as you gently nudge your eye with your finger – far outstrips any visual instabilities that rapid saccades might occasion (Helmholtz 1867). The second sort of case arises when subjects have their ocular muscles paralyzed by a paralytic such as curare. When the subjects attempt to move their eyes, they perceive a shift in the visual scene even though no physical movement has actually occurred. (Mach 1885 and Stevens et al. 1976).

C. Colby (1999) hypothesizes that the lateral intraparietal area (LIP) constitutes the neural locus for the efference copy–based updating of the visual percept. LIP neural activity constitutes a retinocentric spatial representation. However, this activity reflects not just current retinal stimuli but also a memory record of previous stimulation. Additionally, the memory representation can be shifted in response to efference information and independently of current retinal stimulation. Colby (1999, pp. 114–116) reports experiments on monkeys in which LIP neural responses to a stimulus flashed for only 50 msec get remapped in response to a saccade. (The remapping is the shift of receptive fields from one set of neurons to another.) The duration of the stimulus was insufficiently short to account for the remapping; thus, the remapping must be due to the efference copy.

The previously discussed evidence concerning the role of efference copies in perceptual stability during saccades points to some crucial similarities between, on the one hand, complicated natural organisms such as humans and monkeys and, on the other hand, extremely simple artificial organisms such as Tanky Jr. and Radar. Both the natural and the artificial creatures actively scan their environments, and the content of the percept is, while underdetermined by sensory input, determined by the combined contribution of sensory input and efference copy information concerning motor output.

I turn now to consider a possible objection to my account. The account I'm offering here sees action-oriented representations as determining the character of many instances of perceptual experience. J. Prinz (2000) and A. Clark (2002) raise a concern about accounts such as this that postulate relatively tight connections between the determinants of action and the content and character of perceptual experience.[10] The worry stems from consideration of D. Milner and M. Goodale's (1995) hypothesis that visually guided action is localized primarily in the dorsal stream (cortical

---

[10] Such accounts include Grush 1998, Cotterill 1998, Hurley 1998, and O'Regan and Noë 2001.

areas leading from V1 to the posterior parietal area), whereas conscious perception is localized primarily in the ventral stream (cortical areas leading from V1 to inferotemporal cortex). The worry that Prinz and Clark raise is that action cannot be too closely coupled to perception, since the work of Milner and Goodale serves to show a dissociation between the processes that are most intimately involved in action and the processes that are most intimately involved with perceptual consciousness.

I have two responses to this worry. The first is that, unlike, say, R. Cotterill (1998) and O'Regan and Noë (2001), I am *not* saying that the sorts of contributions that action sometimes makes to perception will be either necessary or sufficient for a perceptual state to count as a conscious mental state. I am arguing merely that action-oriented processes sometimes contribute to the representational contents of perceptual consciousness. What contributes to the content of a conscious state need not be one and the same as what makes that state a conscious mental state. Indeed, there are plenty of accounts of consciousness that dissociate the conditions that make a state have a particular content and the conditions that make that state conscious. Two prominent examples are M. Tye's (1995) Poised Abstract Non-conceptual Intentional Content (PANIC) theory and D. Rosenthal's (1997) Higher-Order Thought (HOT) theory. Further, the theories of consciousness that Clark (2000a, 2000b) and Prinz (2000, 2001, this volume) advocate are consistent with this general sort of dissociation.

My second, and not unrelated, response to the stated worry is that there is evidence that activity in the dorsal stream does influence conscious perception. Such evidence includes the evidence previously described concerning parietal processing of efference copies for visual stability during saccades. Additionally, see V. Gallese et al. (1999) for a brief review of various imagery studies implicating parietal areas in conscious motor imagery. M. Jeannerod (1999) similarly questions whether dorsal stream activity should be regarded as irrelevant for conscious perception. He describes PET studies by I. Faillenot et al. (1997) that implicate parietal areas in both an action task involving grasping objects of various sizes and a perception task involving matching the objects with each other.

Conclusion

Perception oft involves processes whereby the perceiver is not a passive receptacle of sensory information but actively engages and explores the perceptible environment. Acknowledging the contributions

that action makes to perception involves a certain rethinking of perception. However, we are not thereby forced to abandon the view that perception is a representational process. Indeed, the impact of action on the mind is mediated through representations of action. In cases in which transducer input is insufficient to provide the requisite representations of action, efference copies of motor commands may be substituted, since they themselves are representations of action. Efference copies are examples of action-oriented representations, and insofar as they contribute to the makeup of perceptual contents, our perceptual states themselves become action-oriented representations.

Acknowledgments

References

Austin, J. (1964). *Sense and Sensibilia.* New York: Oxford University Press.

Bach-y-Rita, P. (1972). *Brain Mechanisms in Sensory Substitution.* New York and London: Academic Press.

Baum, D. (2002). *The Definitive Guide to LEGO MINDSTORMS, Second Edition.* Berkeley, CA: Apress.

Churchland, P. S., Ramachandran, V. S., and Sejnowski, T. J. (1994). "A critique of pure vision." In C. Koch and J. L. Davis (eds.), *Large-Scale Neuronal Theories of the Brain.* Cambridge, MA: MIT Press, pp. 23–60.

Clark, A. (1997). *Being There.* Cambridge, MA: MIT Press.

Clark, A. (2000a). "A case where access implies qualia?" *Analysis* 60: 1: 30–37.

Clark, A. (2000b). "Phenomenal immediacy and the doors of sensation." *Journal of Consciousness Studies* 7 (4): 21–24.

Clark, A. (2002). "Visual experience and motor action: Are the bonds too tight?" *Philosophical Review* 110: 495–520.

Colby, C. (1999). "Parietal cortex constructs action-oriented spatial representations." In N. Burgess, K. J. Jeffery, and J. O'Keefe (eds.), *The Hippocampal and Parietal Foundations of Spatial Cognition.* New York: Oxford University Press, pp. 104–126.

Cotterill, R. (1998). *Enchanted Looms: Conscious Networks in Brains and Computers.* Cambridge: Cambridge University Press.

Dretske, F. (1995). *Naturalizing the Mind.* Cambridge, MA: MIT Press.

Evans, G. (1985). "Molyneux's question." In Gareth Evans, *The Collected Papers of Gareth Evans.* London: Oxford University Press.

Faillenot, I., Toni, I., Decety, J., Gregoire, M. C., and Jeannerod, M. (1997). "Visual pathways for object-oriented action and object identification: Functional anatomy with PET." *Cerebral Cortex* 7: 77–85.

Fodor, J. (1987). *Psychosemantics.* Cambridge, MA: MIT Press.

Gallese, V., Craighero, Laila, Fadiga, Luciano, and Fogassi, Leonardo (1999). "Perception through action." *Psyche* 5 (21): http://psyche.cs.monash.edu.au/v5/psyche-5-21-gallese.html.

Gibson, J. (1966). *The Senses Considered as Perceptual Systems.* Boston: Houghton Mifflin.

Gibson, J. (1986). *The Ecological Approach to Visual Perception.* Hillsdale, NJ: Lawrence Erlbaum.

Gould, S. (1991). "Exaptation: A crucial tool for evolutionary psychology. *Journal of Social Issues* 47: 43–65.

Grice, H. (1961). "The causal theory of perception." *Proceedings of the Aristotelian Society,* sup. vol. 35: 121–152.

Grush, R. (1998). "Skill and spatial content." *Electronic Journal of Analytic Philosophy* 6: http://ejap.louisiana.edu/EJAP/1998/grusharticle98.html.

Grush, Rick (2001). "Self, world and space: On the meaning and mechanisms of egocentric and allocentric spatial representation." *Brain and Mind* 1 (1): 59–92.

Hanneton S., Gapenne, O., Genouel, C., Lenay, C., and Marque C. (1999). "Dynamics of shape recognition through a minimal visuo-tactile sensory substitution interface." *Third International Conference on Cognitive and Neural Systems,* pp. 26–29.

Helmholtz, H. (1867). *Handbuch der Physiologischen Optik.* In G. Karsten (ed.), *Allgemeine Encyklopädie der Physik,* vol. 9. Leipzig: Voss.

Hubel, D. H., and Wiesel, T. N. (1962). "Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex." *Journal of Physiology* 195: 215–243.

Hurley, S. (1998). *Consciousness in Action.* Cambridge, MA: Harvard University Press.

Hyman, J. (1992). "The causal theory of perception." *Philosophical Quarterly* 42 (168): 277–296.

Jarvilehto, T. (1998). "Efferent influences on receptors in knowledge formation." *Psycoloquy* 9 (41): http://www.cogsci.soton.ac.uk/cgi/psyc/newpsy?9.41.

Jeannerod, M. (1999). "A dichotomous visual brain?" *Psyche* 5 (25): http://psyche.cs.monash.edu.au/v5/psyche-5-25-jeannerod.html.

Keeley, B. (2002). "Making sense of the senses: Individuating modalities in humans and other animals." *Journal of Philosophy* 99: 5–28.

Koshland, D. (1977). "A response regulator model in a simple sensory system. *Science* 196: 1055–1063.

Koshland, D. (1980). "Bacterial chemotaxis in relation to neurobiology." In W. C. Cowan et al. (eds.), *Annual Review of Neurosciences 3.* Palo Alto: Annual Reviews, pp. 43–75.

Kosslyn, S., Ganis, G., and Thompson, W. (2001). "Neural foundations of imagery." *Nature Reviews Neuroscience* 2: 635–642.

Larsson J., Amunts, K., Gulyas, B., Malikovic, A., Zilles, K., and Roland, P. (1999). "Neuronal correlates of real and illusory contour perception: Functional anatomy with PET." *Eur J Neurosci.* 11 (11): 4024–4036.

Lenay, C., Cannu, S., and Villon, P. (1997). "Technology and perception: The contribution of sensory substitution systems." In *Second International Conference on Cognitive Technology.* Aizu, Japan, and Los Alamitos, CA: IEEE, pp. 44–53.

Livingstone, M., and Hubel, D. (1988). "Segregation of form, color, movement and depth: Anatomy, physiology and perception." *Science* 240 (4853): 740–749.

Mach, E. (1885). *Die Analyse der Empfindungen.* Jena: Fischer.

Mandik, P. (1999). "Qualia, space, and control." *Philosophical Psychology* 12 (1): 47–60.

Mandik, P. (2001). "Mental representation and the subjectivity of consciousness." *Philosophical Psychology* 14 (2): 179–202.

Mandik, P. (2002). "Synthetic neuroethology." In T. W. Bynum and J. H. Moor (eds.), *CyberPhilosophy: The Intersection of Philosophy and Computing.* New York: Blackwell, 2003, pp. 8–25.

Mandik, P. (2003). "Varieties of representation in evolved and embodied neural networks." *Biology and Philosophy* 18 (1): 95–130.

Millikan, R. (1996). "Pushmi-pullyu representations." In L. May, M. Friedman, and A. Clark (eds.), *Minds and Morals.* Cambridge, MA: MIT Press, pp. 145–161.

Milner, D., and Goodale, M. (1995). *The Visual Brain in Action.* Oxford: Oxford University Press.

Oakes, R. (1978). "How to rescue the traditional causal theory of perception." *Philosophy and Phenomenological Research* 38 (3): 370–383.

O'Craven, K., and Kanwisher, N. (2000). "Mental imagery of faces and places activates corresponding stimulus-specific brain regions." *Journal of Cognitive Neuroscience* 12: 1013–1023.

O'Regan, J., and Noë, A. (2001). "A sensorimotor account of vision and visual consciousness." *Behavioral and Brain Sciences* 24 (5): 939–1011.

Perrett, D., Mistlin, A., and Chitty, A. (1989). "Visual neurons responsive to faces." *Trends in Neurosciences* 10: 358–364.

Peterhans, E., and von der Heydt, R. (1991). "Subjective contours – bridging the gap between psychophysics and physiology." *Trends in Neurosciences* 14: 112–119.

Prinz, J. (2000). "The ins and outs of consciousness." *Brain and Mind* 1 (2): 245–256.

Prinz, J. (2001). "Functionalism, dualism, and the neural correlates of consciousness." In W. Bechtel, P. Mandik, J. Mundale, and R. Stufflebeam (eds.), *Philosophy and the Neurosciences: A Reader.* Oxford: Blackwell.

Rosenthal, D. (1997). "A theory of consciousness." In Ned Block, O. Flanagan, and G. Guzeldere (eds.), *The Nature of Consciousness.* Cambridge, MA: MIT Press.

Stevens, J., Emerson, R., Gerstein, G., Kallos, T., Neufield, G., Nichols, C., and Rosenquist, A. (1976). "Paralysis of the awake human: Visual perceptions." *Vision Research* 16: 93–98.

Sullins, J. (2002). "Building simple mechanical minds: Using LEGO$^{®}$ robots for research and teaching in philosophy." In T. W. Bynum and J. H. Moor (eds.), *CyberPhilosophy: The Intersection of Philosophy and Computing.* New York: Blackwell, 2003, pp. 104–116.

Tye, M. (1995). *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind.* Cambridge, MA: MIT Press.

von der Heydt, R., Peterhans, E., and Baumgartner, G. (1984). "Illusory contours and cortical neuron responses." *Science* 224: 1260–1262.