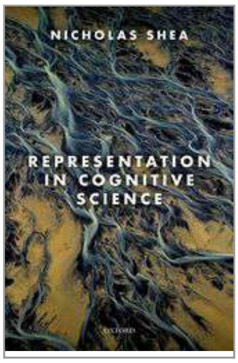

Representation in Cognitive Science
Nicholas Shea

Print publication date: 2018
Print ISBN-13: 9780198812883
Published to Oxford Scholarship Online: October 2018
DOI: 10.1093/oso/9780198812883.001.0001

Descriptive and Directive Representation

Nicholas Shea

DOI:10.1093/oso/9780198812883.003.0007

Abstract and Keywords

A distinction between descriptive and directive representations can readily be drawn within the varitel framework. Neither decoupling, nor the ability to keep track of goal satisfaction, are constitutive of having directive content, although both imply that content will be directive. The distinction drawn here has plausible consequences when applied to the case studies based on UE information and UE structural correspondence in previous chapters. There are several other kinds of sophistication which, while going along with a descriptive–directive difference, are not constitutive of it. Interestingly, the rat navigation case gives us a possible subpersonal example of a mode of representing that goes beyond the descriptive or the directive. Something like supposing may be involved when place cell activity is used offline to calculate shortest routes.

Keywords: descriptive, directive, indicative, imperative, direction of fit, mode of representing, pushmi-pullyu, goal, decoupling, supposition

7.1 Introduction

The representations we have discussed so far have in fact come in two distinct kinds: those with descriptive content and those with directive content. Speaking informally, descriptive representations are supposed to match the way things are in the world and are correct or true when they do. Directive representations are also associated with a worldly condition—a condition they are supposed to bring about. A directive representation is satisfied when its condition comes to obtain. This chapter engages with the long-standing philosophical debate about how that distinction should be drawn.

It has proven difficult to characterize the distinction precisely (Humberstone 1992) and I don't aim to do so in a theory-neutral way here. However, there is a clear distinction, in the right neighbourhood, which can be made within the varitel framework. Section 7.2, following some preliminary points of clarification (this section), gives my account of the distinction. Section 7.3 shows how the distinction works in our case studies. Section 7.4 compares my way of drawing the distinction to others in the literature. Section 7.5 briefly considers some further kinds of sophistication that arise in the case studies.

Many previous treatments have used the terms 'indicative' and 'imperative' for descriptive and directive content. However, those terms are used in linguistics and philosophy of language to label the grammatical mood of a sentence, which need not **(p.178)** align with its content (e.g. 'the door is open' is an indicative sentence that can have the directive content *shut the door*). Using 'descriptive' and 'directive' is designed to avoid confusion.

Beliefs and desires, respectively, are the paradigmatic examples of descriptive and directive representations. There are other types of propositional attitude at the personal level; for example, supposing, entertaining, and imagining. Although our case studies are simpler, it is not obvious that they are restricted to only having the descriptive or directive mode of representing. In §7.5b we will see that something like suppositional content may arise in one of our case studies, when representations are used in conditional reasoning.

I have characterized the descriptive–directive contrast as a difference in content. Others would say this is a difference in attitude between states that could have the same content. According to that usage, content is a worldly condition and the descriptive or directive is the attitude in which that content is represented. That picture derives from the model of beliefs and desires, where the same representational vehicle can be deployed in the attitude of belief (put in the 'belief box'), in the attitude of desire (put in the 'desire box'), in the attitude of intention, and so on. An account of the difference between belief and desire is given in terms of a representation with the same content figuring in two different functional roles (Fodor 1987a). These states share content in a restricted sense (e.g. that p), while differing in attitude.

I prefer to use ‘content’ for the full representational import of a state, so as to include a specification of its mode of representing. Used this way, a correctness condition like *the door is open* and a satisfaction condition like *open the door* are different contents. The functional role of a putative representation fixes this full specification, and it is part of the problem of content to say how that can be so. A theory of content that only delivered content in the more restricted sense would be incomplete. Furthermore, my terminology stops us simply assuming reusability. It’s not always the case that the same attitude-neutral vehicle can be reused in different functional roles so as to give it different modes of representing. That is a plausible hypothesis about beliefs and desires, making the content–attitude distinction useful there. However, it is not a general feature of the systems we have been considering. It does arise in certain cases (§§7.4 and 7.5 below). That can be captured by giving a full specification of the content, so as to include the mode of representing, while making clear that the same vehicles are reused in different functional roles, so that representations in different modes deploying the same vehicle involve the same worldly condition.

Another terminological choice is to talk about directive contents concerning a condition C. That encompasses outputs that are movements of the system and actions performed by the system. It is slightly awkward to call the movement of a body part a ‘condition’ produced by the organism, so it is worth emphasizing that I do intend the term to cover these cases. Limb movements can be individuated in terms of intrinsic properties of the system, but the limb moving that way counts as one type of condition C which the system brings about.

(p.179) 7.2 An Account of the Distinction

The accounts of content offered in Chapters 4 and 5 do not make a distinction between descriptive and directive contents. They home in on conditions whose obtaining or otherwise is important for explaining successful behaviour, without distinguishing between conditions which the system causes to obtain and conditions that obtain antecedently. Recall the central idea: an exploitable relation figures in an unmediated explanation of how a system performs task functions. In some cases that exploitable relation is a matter of receptivity: the system makes use of inputs in order to go into a state or states that stand in the exploitable relation, and the relation is exploited by relying on those states in further processing or by conditioning behaviour on them. In other cases the exploitable relation is a relation to outputs that the system produces: the role of the vehicle in achieving a task function is to bring about a certain result.

Exploitable relations that are part of an unmediated explanation of how a system achieves its task functions can be of either kind. Although the accounts of content do not differentiate between descriptive and directive modes of representing—nor need they do so in order for the obtaining or otherwise of the content-condition to explain success and failure of behaviour—we can supplement them so as to classify whether the exploitable relation plays a descriptive or directive role (or both, or neither).

A tempting first thought is that representations whose tokening is caused by inputs to the system are descriptive and those that cause outputs are directive. And that asymmetry does indeed exist in many cases. For example, the analogue magnitude system carries UE information about the numerosity of collections of objects, that correlation exists because of how the system is sensitive to objects in the world (§4.6a), and the representations do indeed have descriptive content (about the number of objects presented). Motor programs, by contrast, carry UE information about bodily movements and actions caused by the agent (§4.5), and they have directive content.

However, in varitel semantics content is based on explanatory considerations, not simply facts about causal sensitivity and causal effects. So, the descriptive–directive distinction should turn on whether producing condition C or reflecting condition C figures in the explanation (in an unmediated explanation of how—through the representations implementing an internal algorithm—S performs task functions F_j). Furthermore, there is no requirement in the theories above that the objects and properties represented should play any causal role at input. (Recall from §6.2c: the frog's retinal ganglion cell firing could carry correlational information about and represent the location of flies even if the cause of cell firing were patches of light that attract, and so precede, the presence of flies.)

Rather than causation by C, what is characteristic of the descriptive role is that the correlation of R with C, and hence C's obtaining better than chance, enters into the explanation of how downstream effects of R have or achieve their task functions. Contrast the directive case, where what is important is producing C. In the directive case, there would be no point tokening R if C obtained already. In the descriptive case, **(p.180)** the whole point of tokening R is that C does obtain already, or at least that it is likely to obtain at the point when behavioural outputs prompted by R occur.

A complication arises because there are cases of corollary discharge (§4.5) where a (directive) motor program acquires an additional descriptive content (having both modes of representing, it is a *pushmi-pullyu* representation). As well as having the job of causing a bodily movement, it also has the job of telling other subsystems that the movement is about to take place, so that they can produce appropriate outputs, such as balancing movements. To deal with this complication it is easier to define the directive mode of representing first.

Directive Content (based on UE information)

For internal component R carrying UE information about condition C in a system S with task function or functions F_j:

if R's producing C is part of an unmediated explanation of S's performance of task functions F_j,

then R has the directive content: *produce C*

Recall that we defined the explanandum, 'S's performance of task functions F_j', as including both the question of how outputs F_j have been stabilized (through evolution, learning, or contribution to an organism's persistence) and the question of how they have been robustly produced (§4.2a). Directive contents arise where exploitable correlational information becomes UE information because of R's role in causing a condition C to obtain. R's role in the mechanism that was stabilized and produces outputs robustly was to cause condition C to obtain.

We should recall a subtlety discussed in Chapter 4. Output-based UE information may concern the means by which a task function is brought about, as with a motor program that is used to bring about a task functional output F. But there are also directive representations that concern task functions directly, such as *get sugar*. So then the internal computations involve two directive representations: one that selects a task function to be achieved in the current context (*get sugar*) and another that programs the means to that end (*move the right hand to (x,y,z)*).[1] The correlation in the former case is simply a matter of production of a task functional outcome, rather than an explanation of how the system produces that outcome. We saw in §4.2a that the output correlation can nevertheless form part of an explanation of how the whole system comes to produce this outcome in a way that was stabilized and robust.

Having identified the UE information that is constitutive of directive content, we can now turn to descriptive content. The basic idea is that the explanatory role of a correlation in the descriptive case is to raise the probability that some condition C, relevant to how the system performs task functions, obtains. But an obvious complication **(p.181)** arises. The job of a directive is to bring about a certain output. In doing that it raises the probability that a certain condition will obtain (namely that the output is produced). The directive representation *get sugar* should not automatically be counted as also being a descriptive representation just because its satisfaction condition obtaining (the organism getting sugar) is part of an explanation of how the mechanism was stabilized. So, we need to exclude this kind of case in defining descriptive content. Descriptive content concerns a condition C whose obtaining when R is tokened figures in an unmediated explanation of robustness and stabilization, but not in virtue of R's having the causal role to produce condition C.

Descriptive Content (based on UE information)

> For internal component R carrying UE information about condition C in a system S with task function or functions F_j:
>
> if C's obtaining when R is tokened is part of an unmediated explanation, that does not go via R's producing C, of S's performance of task functions F_j,
>
> then R has the descriptive content: *C obtains*

Now we can return to the issue of corollary discharge. While we don't want our definition of descriptive content to automatically constitute every directive representation as also having descriptive content (as a pushmi-pullyu), we don't want to rule out that a directive representation can also acquire descriptive content, in virtue of a second functional role in plays in the system.

Consider a motor command which correlates with the production of a particular bodily movement, holding the right arm out horizontally. That is a role the motor command is supposed to play so it is a representation with directive content. Other motor systems react to this state by producing balancing movements that stabilize the body; for example, tensing muscles in the legs and pelvis (Bouisset and Zattara 1981). These movements are performed before any sensory feedback has come in to report that the right arm has gone up. They are based on the motor program, since it is a reliable sign that the arm will shortly be up, and hence of the need for compensatory adjustments. These compensatory adjustments are motor outputs, and a condition involved in explaining why they are stabilized is that the right arm is up (otherwise they would be a waste of effort and destabilize the body in the other direction). The motor program itself carries the correlational information that this condition is about to obtain. I would argue that *that* correlation figures in an unmediated explanation of how the organism achieves the task function of remaining upright. So, the motor command, in virtue of a second functional role that it plays in the system, also carries the descriptive content that the right arm is raised.

Does our definition of descriptive content allow that there is also descriptive content in this case? The organism performs the task function of staying upright. It does that when the arm is raised by producing balancing movements of the muscles of the legs and torso. It produces those balancing adjustments in appropriate circumstances by **(p.182)** relying on an internal vehicle R that correlates with the arm's being raised (condition C). So, R looks to have the descriptive content that the arm will shortly be raised. What about the caveat that the explanation 'does not go via R's producing C'? R does in fact produce C (i.e. cause condition C to obtain). But the way the obtaining of condition C combines with balancing movements to explain performance of the task function of remaining upright does not depend on R's producing C. It would work just as well if the correlation of R with C depended on R's detecting some other sign of

C and R had no role in producing C. The correlation does not fall under the exception, so qualifies as descriptive: the motor program does indeed have a second, descriptive content. So, this definition of descriptive content allows that representations can have both descriptive and directive content in appropriate cases, without going too far and entailing that all directives also carry descriptive content.

We can make a parallel descriptive–directive distinction within the account of UE structural correspondence (Chapter 5). In the case of structural representations, a representation is tokened in virtue of two or more vehicles standing in a certain relation within the system. So, a particular representation R_1 is realized when a relation V holds between two vehicles v_1 and v_2 (e.g. one place cell is activated after another). Under the UE structural correspondence which gives the content of the set of representations R_i which includes R_1, the two vehicles v_1 and v_2 stand for two entities in the world, say x_1 and x_2, and the tokened relation V stands for a relation between x_1 and x_2, call it H. R_1 then represents condition $H(x_1,x_2)$.

> *Directive Content (based on UE structural correspondence)*
>
> For representation R_1 from a set of representations R_i bearing the UE structural correspondence to relation H on a set of entities x_k under which R_1 corresponds to $H(x_1,x_2)$,
>
> if the fact that the R_i produce relation H on the entities x_i is part of an unmediated explanation of S’s performance of task functions F_j,
>
> then R_1 has the directive content: *produce $H(x_1,x_2)$*
>
> *Descriptive Content (based on UE structural correspondence)*
>
> For representation R_1 from a set of representations R_i bearing the UE structural correspondence to relation H on a set of entities x_k under which R_1 corresponds to $H(x_1,x_2)$,
>
> if the obtaining of relation H on the entities x_k when a representation R_i is tokened is part of an unmediated explanation, that does not go via R_i producing relation H on the entities x_i, of S’s performance of task functions F_j,
>
> then R_1 has the descriptive content: *$H(x_1,x_2)$ obtains*

Consider the spatial navigation system in the rat hippocampus. It exploits the correlation of a place cell with location to tell the rat where it is. Then it makes use of the co-activation structure to run through a series of routes offline until one joins up with **(p.183)** a rewarded location, selecting the shortest such route by selecting the shortest or fastest offline sequence. We have been supposing that this structure arises through learning because it has led rats to follow efficient routes to worthwhile locations, food sources say. Getting to previously encountered sources of food are then outcomes F_j that have been stabilized by learning. To explain how they were stabilized we point to the fact that, when the co-activation relation is tokened on place cells, the corresponding relation of spatial contiguity tends to exist on the corresponding locations; for example, that location y is near to location x. The co-activation relation therefore *descriptively* represents that *y is near to x*.

7.3 Application to Case Studies

(a) UE information

In this section I say briefly how my way of drawing the distinction applies to the case studies discussed in previous chapters. It agrees with Millikan's verdict about very simple throughput cases like the honeybee nectar dance. Those intermediates are pushmi-pullyu representations with both descriptive and directive content. Signalling evolved by natural selection and was stabilized because the dances have a tight correlation with the location of nectar at input and with the flight of foraging bees at output.

In the ALCOVE model (§4.3), the output representations are also pushmi-pullyu. The system exploits the fact that they correlate tightly with the category of object presented and with which box they cause the system to place the object in. I argued that the layer of exemplar nodes should not be seen as simply recapitulating with less fidelity the correlations used at the output (since that is less explanatory of how the system works). So, the explanatory correlation at the exemplar layer is with the identity of the exemplar, and not with any particular behaviour. They have descriptive content; similarly for the input nodes. Notice that the pushmi-pullyu representation at output has different (but related) conditions in the descriptive and directive contents (*the object is of type A; place the object in box A*). That is also true of the honeybee nectar dance.

There are also pushmi-pullyu representations where descriptive and directive content concern the same condition. Using motor programs to drive compensatory bodily movements gave us one example (§7.2 above). Corollary discharge will in many cases, for the same reason, have contents like this (§4.5). The predictive comparator mechanisms involved in motor control, discussed in §3.6a, are also somewhat like this in structure (Desmurget and Grafton 2000, Wolpert and Ghahramani 2000), except there the motor content is a pure directive and is transformed before a corresponding descriptive content is represented.

The schematic structure of a predictive comparator mechanism is illustrated in Figure 7.1 A motor command is used simultaneously to drive behaviour and to predict the sensory feedback that is likely to result. The discrepancy between predicted sensory feedback and desired state is used to adjust the motor command even before any **(p.184)** sensory feedback has had time to arrive (processing steps above and to the left of the diagonal dashed line in Figure 7.1). However, there the predictive descriptive representation (‘model state estimate’) is the result of a further processing step after the motor command, so the motor program itself will not have descriptive content. It is a pure directive. The fact that it correlates with outputs means that it can be relied on in order to generate another state with descriptive content about those outputs.

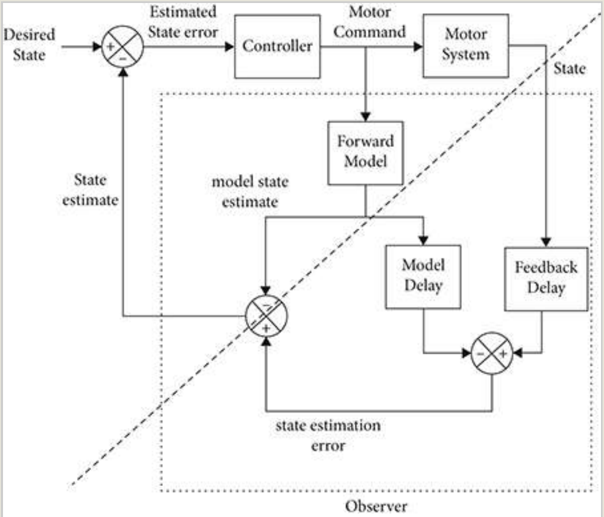

Figure 7.1 A predictive comparator model from Miall and Wolpert (1996, adapted). A copy of the motor command is fed into a mechanism that estimates the likely world state that will result (the ‘model state estimate’). That estimate is then compared with the target state (‘desired state’) so that the motor command can be adjusted even before sensory feedback is received (the processing steps above and to the left of the diagonal dashed line).

Recall the PFC colour/motion choice system described by Mante et al. (2013) (§4.6b). The computation involves two vehicles at input. One, the context representation, correlates with whether it is currently in a ‘colour context’ or a ‘motion context’ (i.e. whether colour or motion is going to be the basis of reward). The other vehicle, call it R_I, has two dimensions of variation, one corresponding to the colour of the stimulus and the other to the motion of the stimulus. R_I is then transformed into a one-dimensional vehicle that drives behaviour (left or right), which is formed by preserving the dimension of R_I that is relevant to choice behaviour on the current trial (colour or motion) and collapsing the other. R_I thus descriptively represents both the preponderant colour **(p. 185)** and the direction of motion of the current stimulus. The context representations are descriptive representations of whether colour or motion will be the basis of reward in the current context. R_I is transformed into a directive representation R_O which drives an action to the left or right. R_O also correlates with a disjunctive input condition (e.g. *the majority colour is red or the majority direction is left*), but since there are other components of the mechanism correlating more tightly with colour, motion, and choice context, the fact that this internal component correlates with this further disjunctive condition offers no additional explanatory purchase. It does not figure in an unmediated explanation of stabilization. Thus, even in this simple case, there are separate descriptive and directive representations, and no pushmi-pullyu representations.

Pure descriptive contents are found in many of the case studies. Analogue magnitude representations of numerosity are pure descriptives, as discussed in §7.2 above. Representations of individuals in the face recognition system are another clear case (§3.4c). No output behaviour at all need be involved in acquiring a representation of a new individual. The plaid motion detection system (§4.7) descriptively represents motion properties (in thick stripe V2) and chromatic properties (in thin stripe V2). It transforms these into a register R in MT that correlates with the direction of motion of surfaces. The system makes use of that correlation to condition behaviour that is appropriate to the motion direction, so R has descriptive content about the direction of plaid motion. In our stylized treatment we supposed that R is used directly to condition behaviour (a reaching movement in the corresponding direction). If the connection were that direct, R would also have directive content, however in the human visual system behaviour is likely to be highly context-dependent, so no such directive content would arise in MT.

In §4.8 we looked at a simple system that accumulates evidence and computes the probability that various available actions will be rewarded. We explained its operation in terms of exploited input correlations, taking into account probability distributions. Only its outputs have directive content, programming one or other action. This case shows that a vehicle of descriptive content can be activated with various strengths, with stronger activation representing that a worldly condition is more likely. That raises the possibility that directive representations could also come in varying strengths, an idea we will return to in §7.5a below.

(b) UE structural correspondence

I turn finally to our cases of UE structural correspondence. In the rat navigation case, the correspondence between co-activation structure on place cells and spatial structure on the corresponding places is used to calculate shortest routes. The rat doesn't put the locations into spatial relations. The behaviour it performs relies on the correspondence holding when it calculates a route. So, they have descriptive content.

It is harder to find simple cases of UE structural correspondence with directive content. We are familiar with high level cases like using a blueprint to build a building. A blueprint represents the way parts should be put into certain spatial relations. Chris Eliasmith's **(p.186)** spiking neuron model for solving the Tower of Hanoi problem uses representations of relations amongst discs and calculates over these relations in working out which disc to move when (Eliasmith 2013, pp. 191–8). These are arguably structural representations of spatial relations. The representation of the target end state uses this structure to represent an arrangement of discs. It is then a structural representation with directive content. There may be lower-level cases involving causal relations. If an organism plans a sequence of actions and represents them with some kind of directed causal graph, which it then consults in order to execute the actions in the right sequence, that causal graph would have directive content about the causal structure of the sequence of actions to be performed.

Dan Ryder's SINBAD model uncovers statistical structure in the world (Ryder 2004, Ryder forthcoming). Ryder claims that SINBAD constructs structural representations. A clever and attractive feature of the model is that it can be inverted so that representations can be used in directive mode so as to drive and guide action. The model discovers statistical regularities in the input, so that it ends up with individual cells that are tuned to underlying causes of statistical regularity (Ryder's 'sources of mutual information'); for example, to natural kinds. When it encounters a series of kinds to which it is tuned, a corresponding series of cells will be activated. Ryder calls the way cell activity mirrors the sources of mutual information encountered a 'dynamic isomorphism'.

The network can also learn which kinds are correlated with which, through forming lateral associative connections. So, if cell B is normally activated by sensory inputs when cell A is, the system will learn that association. Cell A will now activate cell B even in the absence of sensory inputs to cell B, and the converse (in Ryder's model learning produces bi-directional connections). That allows the network to 'fill in' with predictive activity in line with patterns it has encountered in the past. Ryder shows how this network can be taken offline and used in a directive mode of operation (Ryder forthcoming). One cell is specified as a target state. The rest of the network is allowed to adjust, in the absence of input, to reach values that would produce this target state. When the system is operating in this functional role, activity of each node has directive content. Directive content about a target sets off a chain of simple inferences until the system arrives at directive content that it can execute (which thereby acts as an intention or motor program).

As described so far, no structural correspondence between cell activity and conditions in the environment is doing any work. The system is exploiting a collection of vehicles carrying correlational information, both input correlations and output correlations. We saw an example of the way exploitable correlational information about a distal output like getting sugar can depend, for its existence, on an exploitable correlation with a more proximal output; for example, with pressing the left-hand button. Similarly, SINBAD has some exploitable output correlations that depend on others. Once it has undergone learning, the system has a vehicle which correlates with producing a distal outcome, getting hot water, and that correlation proceeds via another vehicle **(p.187)** which correlates with producing a more proximal outcome, turning the left-hand tap. Just as input information is transformed through a series of correlational vehicles (e.g. in the perceptual hierarchy) output information can also depend on intermediate transformations. The complete computation is making use of multiple correlational vehicles at output as well as at input, and the way they are computed over instantiates an algorithm for performing task functions. That is not yet a case of structural representation (§5.7a).

A more sophisticated model of the same type might contain structural representations. Suppose the chains of activity between cells correspond, not to co-instantiation of kinds, but to regularly encountered causal sequences. Suppose further, that when the network is used in directive mode, the steps of activity produced by the inverse model are acted on in sequence. T is set as a target, which in directive mode activates A1, then A2, and so on. Rather than connecting all of these to action simultaneously, the system deploys them in serial order, first doing the action driven by A2, then that driven by A1. (To work like this something more sophisticated than the simple bidirectional associations suggested above would have to have been learnt.) Using the network in this way would make use of the correspondence between the temporal order of activation (in the network) and the causal sequence of events (in the world). If this temporal order was important to the reasoning in some way (as described in §5.6b), then it would be a structural representation of the causal order of events.

A notable feature of this case is that there is something like a content–attitude distinction: the same vehicle can be redeployed in descriptive or directive mode. I noted above that this feature, characteristic of beliefs and desires, is not a necessary feature of representational systems in general. It is interesting to see that it can nevertheless arise in a case that is simpler than the belief-desire system.

In fact, the rat navigation system too may take place cells, whose descriptive content we have already captured, and redeploy the same vehicles in a different role which gives them directive content. Consider what happens once the rat has calculated a shortest route and needs to set off. It should follow a path corresponding to the sequence it has just calculated. It does that by putting its place cell system online again, so that its activity reflects the rat's current location. Now let's make a supposition, just to illustrate the point. Suppose that the place cell for the next step in the shortest route sequence becomes active, and the difference between that activity and the currently active place cell is calculated. If the rat moves so as to minimize this error signal, it will tend to move to the next location, where the process is then repeated. When active as a target, the place cell has a directive (correlational) content: *move to location y*. When activated by an input signal, a signal which is changed by movement, a place cell has descriptive (correlational) content: *I am at location x*. The same cell has directive or descriptive content about the same location, depending on its current functional role. (Note that the representational *relation*—co-activation—has descriptive (structural correspondence) content throughout.)

(p.188) 7.4 Comparison to Existing Accounts

In this section I compare my account with three broad approaches found in the literature: (i) canonical teleosemantic accounts, which like my account appeal to an asymmetry in the way conditions explain successful behaviour; (ii) accounts based on decoupling at input and output; and (iii) accounts based on the system being able to detect when a directive representation has been satisfied (i.e. when a goal has been reached). I also reply to an objection that has been made to teleosemantic approaches.

Canonical teleosemantic accounts of the descriptive–directive distinction (sometimes referred to as 'direction of fit') appeal to the existence of a producer and a consumer of a representation (§1.4), something that is not available in my framework. Nevertheless, my distinction is closely related to the canonical teleosemantic treatment. According to teleosemantics, a representation R has directive content C when it has the function of causing a consumer system to produce outcome C (Millikan 1984, ch. 6). (It follows that R itself has the function of producing C, by that means.) A representation R has descriptive content when it is produced by a producer mechanism that has the function of producing R when condition C obtains, which in turn means that C's obtaining figures in an explanation of how behaviour of the consumer, prompted by R, led systematically to survival and reproduction.[2]

My definition takes from teleosemantics the core idea that directive contents are outputs which a system generates and which explain why a pattern of behaviour was stabilized; and that descriptive contents are other conditions whose obtaining explains how those outputs are produced and lead to stabilization. (Stabilization for me is not restricted to selection; and I add explanation of robustness.) My approach covers cases with multiple interacting internal components (so does not require a producer and a consumer). It also has a correlation or correspondence requirement at input (Chapter 4, Chapter 5). The obtaining of an exploitable relation (correlation or correspondence), in virtue of which the representations had systematic connections with external conditions, figures in the explanation of task functions—not just the fact that those conditions happened to obtain on occasions when behaviour had results that led to its stabilization. However, when dividing approaches to mode of representing into broad camps, my approach falls into the same family as canonical teleosemantics.

Marc Artiga argues that Millikan's account of directive content entails that all representations have directive content, which is implausible (Artiga 2014a). (Artiga uses this conclusion to argue in favour of an account according to which only descriptive content is attributed to simple representations: p. 552.) For there to be descriptive content in the first place, there has to be some range of behaviours which representation R is supposed to prompt its consumer to produce. That range may be very wide: $O_1 v O_2 v \dots O_n$. Nevertheless, it follows, according to Artiga, that R has the directive content to produce $O_1 v O_2 v \dots O_n$.

(p.189) This kind of disjunctive condition is unlikely to arise as a directive content in my framework, since correlating with $O_1 v O_2 v \dots O_n$ is a poor candidate for exploitable correlational information or for explanation, on usual approaches to causal explanation. First off, highly disjunctive properties are generally poor candidates to figure in nomological generalizations, hence in exploitable correlational information. Furthermore, think about candidate algorithms for performing S's input–output function. No such algorithm is likely to call for a stage of processing where such a nondescript fact is represented. In any event, the fact that a very general umbrella condition like $O_1 v O_2 v \dots O_n$ obtains is unlikely to explain why a particular output F is robustly produced, nor to explain why it has been stabilized. My focus on how a collection of available internal states collectively explain task functions also means that there will generally be different contents for different vehicles (§6.2f), whereas Artiga's proposed content attribution results in very many representations having just the same (highly disjunctive) directive content. So, Artiga's objection does not apply to my account: a representation's having descriptive content does not entail that it has directive content.[3]

A second broad family of approaches identifies the descriptive–directive difference with an asymmetry in the coupling or decoupling that exists between a representation, on the one hand, and inputs and outputs, on the other. According to this approach, directive contents are still outputs that the system is supposed to produce, but there has to be a certain amount of decoupling between the representation and behaviour for it to count as a representation at all. So Price (2001) says directive content only arises when a representation has been selected to produce a particular outcome via a range of different behaviours in different circumstances (Price 2001, p. 141). It acts as a goal, on the basis of which the system selects, from a range of possibilities, a sequence of movements that is suitable for achieving that goal, performing a simple inference that also relies on descriptive representations about conditions in the environment. If the mediating behaviours are not themselves produced by a range of different means (if they are not robust outcome functions of the system, in our terminology), then they cannot be directively represented on Price's view. As well as lacking directive content, the motor programs driving these behaviours need not have descriptive content either. So, the important role of motor programs would be left out in the representational explanation of behaviour if we follow Price's definition.

Sterelny (2003) distinguishes between two kinds of decoupling (2003, pp. 30–40). Representations with ‘response breadth’ are not tightly functionally coupled to specific types of response (like Price’s condition on having directive content). Representations that show ‘robust tracking’ are those which function by making use of a variety of cues so as to correlate with a feature of the environment. Sterelny argues that response breadth is a defeasible way of distinguishing descriptive from directive **(p.190)** representations: descriptive representations tend to be strongly decoupled from producing any specific type of response.

By making the converse point about directive representations, we would have an asymmetry in decoupling which differentiates between descriptive and directive representations. Kevin Zollman makes just this distinction, working within the sender–receiver models developed by Skyrms (Zollman 2011). Zollman differs from Skyrms who embraces the idea that all signals have both descriptive and directive content, content for Skyrms just being a function of the correlational information carried about inputs and outputs (Skyrms 2010). Zollman (2011) argues that an asymmetry between the correlational information carried about inputs and that carried about outputs must exist if a representation is to have only one direction of fit (2011, p. 163). He describes a signalling game in which such asymmetries arise. Descriptive representations are more tightly tied to inputs—they carry more information about world states—and directive representations are more tightly tied to outputs—they carry more information about acts.[4]

My distinction tends to go along with this kind of decoupling. Where R has directive but not descriptive content, production of outputs figures in the explanation of task function performance, so R will tend to carry stronger correlational information about outputs and to be decoupled from any specific input. The converse applies to a representation that carries descriptive and not directive content. It will tend to be more decoupled on the output side. I noted above that it would be a mistake to make these causal facts the basis for the distinction. Furthermore, doing so fails to capture the descriptive contents that are present in the kind of case mentioned in §7.2, where a motor program for movement B is relied on in generating further kinds of behaviour which make sense in the light of the fact that B is produced (the motor program thereby acquiring secondary descriptive content).

My view differs more markedly from a different kind of decoupling proposal, where decoupling is not just a causal matter, but depends on the kind of psychological processes that are performed on a representation. This idea stretches back to the way Lewis draws the distinction (Lewis 1969), and gets its intuitive support from the way hearers interpret and react to linguistic commands. Characteristic of a command is that the hearer doesn’t have to deliberate about what to do (e.g. ‘attack by sea’). The hearer can comply just by following the command. Assertions also have consequences for action (e.g. if 'the British are coming', there is probably something you should do about it), but the hearer has to deliberate about what to do in the light of that information. Huttegger (2007) models the distinction in this way by building the possibility of deliberation into his model.

(p.191) Artiga suggests something similar when he says that decoupling is not just a causal matter, but that imperative representations are those that produce their behavioural results automatically (2014a, p. 558–9). That calls for a distinction between automatic and non-automatic use of a representation. The latter presumably involves deliberation, but as ordinarily understood, none of our case studies involve deliberation by the subject. There is processing involved in generating representations and in acting on them, but it is hard to see how a distinction between automatic and non-automatic processing can be applied. So, if 'not automatic' means something more than just 'decoupled', this is not a promising way to differentiate between descriptive and directive content in our cases.

A third general approach picks up on a different feature of a goal, which is that there is no point in pursuing it further once it has been brought about. On this view it is constitutive of having directive content about output C not only that the system is disposed to bring about C, but that it is sensitive to whether C obtains, so that behaviour prompted by the representation should cease once the system detects that C does obtain (Dickie 2015, p. 282).[5] Millikan makes this a criterion, not of having directive content at all, but of being a pure directive rather than a pushmi-pullyu representation (Millikan 2004).

That is a demanding constraint if it is taken to be a condition on having directive content at all, since it requires the system to have separate descriptive and directive representations about condition C, and to compare them. So, it would deny directive content in many of our simple cases, even when it is clear that it is the correlation with outputs that is explanatorily relevant, and whose coming to obtain (or not) explains the success or failure of behaviour. My approach is compatible with the intuition behind this view, however, because my distinction entails that representations which meet this demanding condition will have directive content. The toy model in Chapter 3 is an example (§3.6a and §4.1b). Where a comparator mechanism takes as input two vehicles and drives behaviour until they match, if the pattern of behaviour is such that one correlates strongly with the current state and the other with the output that tends to be produced, then the first will come out as descriptive and the second as directive (provided, as always, those correlations figure in an unmediated explanation of the system's performance of task functions). In short, there being a comparator mechanism is not a plausible requirement on the existence of directive content, but it is one way in which a system can come to have pure directives that are different from its pure descriptives. It is one way of making a system more sophisticated. The next section looks more closely at various kinds of cognitive sophistication that arise in directive systems.

(p.192) 7.5 Further Sophistication

(a) More complex directive systems

When we look at human beliefs and desires we see further layers of sophistication, which go beyond simply having separate descriptive and directive representations. This section briefly mentions four kinds.

In the human belief-desire system, orthodoxy holds that the same vehicle of content can be entertained with different attitudes. The vehicle of a belief can be redeployed to form a corresponding desire, and vice versa. That is not generally true in our case studies, but we did see two examples where it does arise: when place cells are used to guide action; and when the SINBAD network is used both to represent what is the case and to direct behaviour. We also saw, in the case of motor programs, that the correlational information carried by a directive representation could be made use of by other systems for its exploitable correlational information about outcomes so as to give it additional descriptive content.

The phenomenon of conditioned reinforcement is another example. In instrumental conditioning, obtaining a reward—that is detecting a certain kind of feedback—cements or promotes an internal configuration that encourages the disposition to produce the same kind of behaviour in same circumstances (Dretske 1988).[6] A stimulus that has never elicited a reward can itself become a reinforcing feedback if it has been repeatedly paired with a primary reinforcer. For example, if a rat has repeatedly observed a light paired with the delivery of food, then the light becomes a secondary reinforcer. The rat will then learn to act in ways that cause the light to go on, even if the light is not paired with food during that learning phase (Colwill and Rescorla 1988). The phenomenon of secondary reinforcement strongly suggests that a representation that descriptively represents that a condition C obtains can come to be a directive representation that functions to make the animal bring it about that C obtains.

Dickie’s account of directive content (Dickie 2015) and Millikan’s account of purely directive content (Millikan 2004) rely on a directive representation producing behaviour until the system detects that its satisfaction condition has come to obtain. It is usually assumed that this is achieved because the same vehicle is used for the descriptive and directive representation, making them readily comparable. On my account, keeping track of whether you have reached your goal is not required for having directive content, even for having purely directive content. Keeping track of satisfaction is an additional level of cognitive sophistication which, where it arises, does produce a descriptive–directive difference in the way that is usually assumed. The general-purpose redeployability arguably found in the belief-desire system is a further level of sophistication again.

One final level of sophistication is worth mentioning. Sometimes directive representations conflict: they concern conditions in the world that exclude one another, or the actions required to bring them about are different and cannot both be performed at **(p.193)** once. Many organisms have a system for sorting between their directive representations to prioritize which ones to act on. That is not the same thing as being able to engage in practical reasoning: to reason from a directive representation (bring about C) and a conditional belief (C is likely if B) to a new directive representation (bring about B). (Arguably secondary reinforcement is a simple instance of that pattern.) The ability to sort amongst and prioritize goals is a further characteristic feature of the human belief-desire system. Although not constitutive of a representation's having directive content, its operation is a set of functional roles that is likely to imply that the vehicles so deployed have directive content. A characteristic way of doing that is for desires to have different strengths.[7] Relative strength can vary over time, for example through the vehicles being more or less active.[8] The choice of action then depends both on the relative strength of the agent's desire for q and the agent's assessment of how likely it is that they can bring about q.

Taking stock, there are at least four levels of sophistication involving directive representations. (There are doubtless others too.) First, there is having separate representations with purely descriptive and purely directive content, something not found in the simplest pushmi-pullyu systems. Secondly, there is the ability to keep track of when a directive representation has been satisfied, hence to compare descriptive and directive representations concerning the same worldly condition. Thirdly, there is the ability to redeploy the vehicle of a descriptive representation in a mode that gives it directive content concerning the same condition, and the converse. Fourthly, there is the ability to calculate over directive representations in order to prioritize which ones to carry out. I have argued that none of levels two, three, or four are needed in order to have representations at level one, pure directives and pure descriptives.

(b) Another mode of representing

Propositional attitudes admit of other modes of representing in addition to the descriptive and the directive, for example supposing. Something like supposing may be at work in one of our cases. When the rat navigation system is run offline to calculate the shortest route, place cell activation does not correlate with or represent where the animal is at that moment. Instead, activation represents somewhere it might be. Co-activation still represents the spatial relations between places, so when one place cell causes the activation of another offline, that represents that one location is near another, or that being at one location the animal could move directly to the other. So, the relation between vehicles is still a descriptive representation of spatial structure. The system is using that representation to do a kind of conditional reasoning: if you were at x, then you could get to y, and then to z, and so on.

(p.194) Activation of an individual place cell in this process has neither descriptive nor directive content. One possible explanation for that is that it has unsaturated content, as we have seen (§6.3). Here I want to look at a second possible explanation, namely that it has content in a different mode of representing, a kind of suppositional or hypothetical content. Activation of a place cell offline concerns the same condition C as when it is activated online with descriptive content (Chapter 5) or directive content (§7.4 above), but it has a different overall content. It says something like *suppose you were at x*. That causes another place cell to fire, by virtue of the connection between vehicles. Instantiation of the co-activation relation represents, under the UE structural correspondence, that location y is near to location x. And the second place cell is also saying something hypothetical: *y would be nearby*. That is the conclusion of a little chain of reasoning: from *suppose you were at x*, and *y is near x*, to *y would be nearby*.

Now I have offered two possible explanations for why the activity of an individual place cell offline does not have a correctness condition: because it is unsaturated (§6.3) or because it has suppositional content (here). Which should we prefer?[9] It seems to me to count slightly against the unsaturated view that online place cell activity has a saturated content (*you are at x*). The suppositional view says that, when the vehicle is reused offline, the same saturated condition is in play, albeit in a different mode of representing (*suppose you were at x*). The unsaturated view calls for a switch from saturated content online to unsaturated content offline. On the other hand, the unsaturated proposal seems like a simpler account of the offline reasoning, since it just runs off straightforward descriptive representations like *location x is near location y*. I don't propose to resolve the issue. Either way, offline place cell activation introduces kind of sophistication which might, at first pass, be thought to be the preserve of propositional attitudes.

I won't speculate here on whether the suppositional content I have just described corresponds to the mode of representing of any propositional attitude state (e.g. supposing), or whether it can be captured properly by any linguistic term. I will rest with noting the interesting functional role of the representations involved in this offline reasoning.

7.6 Conclusion

This chapter draws the descriptive–directive distinction in a non-theory-neutral way: it shows how a distinction along these lines arises straightforwardly within the varitel framework. The account retains the virtues of the standard teleosemantic treatment of mode of representing, while being preferable to other existing theories, and also claiming some advantages of its own. Neither decoupling, nor the ability to keep track of goal satisfaction, are constitutive of having directive content. However, my account does imply that, where they exist, these features will give rise to a descriptive–directive difference as expected. The account can readily be applied to the case studies based on **(p. 195)** UE information (Chapter 4) and UE structural correspondence (Chapter 5), with plausible consequences. There are several other kinds of sophistication which, while going along with a descriptive–directive difference, are not fundamental to it. Interestingly, the rat navigation case gives us a possible subpersonal example of a mode of representing that goes beyond the descriptive or the directive. Something like supposing may be involved when place cell activity is used offline in calculating shortest routes. In short, the varitel framework supports a useful way of understanding the descriptive–directive distinction. **(p.196)**

Notes:

([1]) An output that is a means for producing a task functional outcome, e.g. moving the eyes thus-and-so, need not itself be a task functional output, e.g. if it does not meet the robustness condition.

([2]) In Millikan's terminology, C is a Normal condition in the most proximate Normal explanation of the behaviour of the consumer system prompted by R (Millikan 1984).

([3]) A similar reply is available to Millikan since directive contents derive from the 'most proximate Normal explanation' of behaviour, which will both count against disjunctive outputs figuring in the explanation, and count in favour of different representations in the same range having different contents.

([4]) This is not Sterelny's distinction, since 'robust tracking' implies strong correlation with (distal) inputs, but it is compatible with Sterelny's view if we take his 'robust tracking' condition to be a requirement on being a cognitive internal state at all, rather than a tool for distinguishing between directions of fit.

([5]) Smith (1987) is in the same spirit: it is constitutive of being a *desire* that p (having the directive mode of representing) that the state tends to endure in the face of the perception that not-p and dispose the subject in that state to bring it about that p (at p. 54).

([6]) See §4.2 and §8.2e.

([7]) An agent engaged in planning or means-end reasoning has to have some motivating state, some directive representation. If it has two or more, and they call for control of the same effectors, then there must be facts about how likely each is to bring about its outcome in the presence of the others. That is their relative strength.

([8]) We saw something similar with descriptive representations where the level of activation represented the probability that a particular world state obtains (§4.8).

([9]) At the cost of some complexity, it is even possible to combine the views, so that an unsaturated constituent is used suppositionally, rather as we might say ‘consider Nisha for a moment’.