

ANSGAR BECKERMANN

WHY TROPISTIC SYSTEMS ARE NOT
GENUINE INTENTIONAL SYSTEMS*

1.

If one uses a set of rather simple criteria for the ascription of intentional states one may be forced to reckon among the class of intentional systems not only men and higher animals, but also lower animals like starfish and jellyfish and perhaps even amoeba and paramecia. And, what seems worse to many, one may be forced to say that, according to these criteria, even cybernetic mechanisms like thermostats are intentional systems. Some philosophers have accepted this conclusion with a shrug, murmuring "Why not?". But most of us, I think, have a strong feeling to the contrary since to speak in such an inflationary way about intentional states would deprive this kind of talk of its whole point. So we are badly in need of better criteria which are less at variance with our intuitions. What really is the difference between systems which are genuine intentional systems and systems which are not?

Since I cannot hope to answer this question in general I shall restrict myself to a minor part of the problem: the development of a set of criteria which enable us to show that at least tropistic systems – e.g., amoeba and paramecia and thermostats – are not intentional systems. Intentional systems are, as is generally agreed, those systems which really have genuine intentional states (i.e., especially: genuine wants and beliefs) or – to put it in more cautious terms – those systems to which we can legitimately ascribe such intentional states. Hence, to claim that tropistic systems are not genuine intentional systems really is to claim that tropistic systems do not have genuine intentional states, that they do not have genuine wants and beliefs. I shall therefore try to show that this claim is indeed true, that there are good reasons for not ascribing genuine wants and beliefs to systems of this kind.

2.

I would like to begin my considerations with a simple example. Imagine a little mechanical bug – let us call it S_1 – which is con-

Erkenntnis 29 (1988) 125–142.

© 1988 by Kluwer Academic Publishers.

structed to move towards any bright enough light in its environment. This kind of behavior is, as we shall assume, brought about in the following way: on the top of the mechanical bug S_1 there is a circle of 24 photo-electric cells which especially respond to light coming from the direction to which they themselves are oriented (compare figure 1). For instance, if there is a bright light at an angle of 45 degrees in front of the bug then the corresponding photo-electric cell (cell #3) will have a considerable output while all other cells remain rather "silent". Given this situation S_1 will turn around (right or left, depending on whether the cell with the greatest output is on the right or on the left side of S_1) until the output of the front-cell (cell #0) is greatest and then move on straight forward towards the light. If the output of all cells lies under a certain threshold, S_1 will not move at all.

S_1 is what is commonly called a positively phototropic system. I

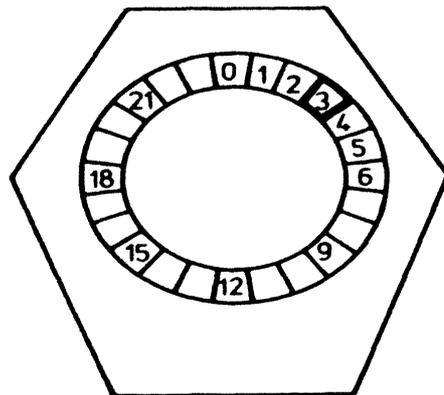
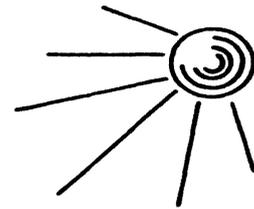


Fig. 1.

shall therefore try to elucidate by means of this simple mechanical bug what I think to be the main differences between tropistic and genuine intentional systems, i.e., what in my eyes are the reasons for not ascribing wants and beliefs to things like the system S_1 . Let us begin with the question why systems like S_1 are not the right kind of things to have genuine wants. As a first step to answering this question it is interesting to compare the system S_1 with objects which in a way show a quite similar behavior – though no one, I hope, is tempted to call these objects intentional systems. Think e.g., of a physical experiment which you will perhaps remember from your school days. A knitting-needle is magnetized, put through a cork and then the cork-plus-needle is placed in a little water basin so that only the north-pole of the needle stands out from the water. If now a bar-magnet is placed on the edge of the little basin the north-pole of the needle will move on a certain predictable course towards the south-pole of the bar-magnet.

One could ask whether there really is a difference in principle between the system S_1 and the magnetized knitting-needle since, after all, the behavior of these two objects is indeed rather similar. But in my opinion there is such a difference, and, as I see it, this difference lies in the fact that the knitting-needle is moved by *outward* forces operating on the needle while the system S_1 is moved from *within*. S_1 is, in the strict sense, an *automaton*. Systems like S_1 need a motor and a power supply to be able to move. This is already shown by the simple fact that they stop and stand still if their motor is broken or they have run out of fuel (or their battery is empty or whatever). In contrast to this it is characteristic of movements like the movement of the knitting-needle that they can only be prevented by the exertion of a sufficient counterforce. It is for this reason that we say that the needle is *attracted* by the bar-magnet while we wouldn't say – except metaphorically – that S_1 is attracted by the light, though, to be sure, even this system has an immanent tendency to approach the brightest light in its environment. Hence the difference between the knitting-needle and the system S_1 can be summed up like this: while the needle is pushed (or pulled) around by external forces the source of the motions of S_1 lies in this system itself. I think that by virtue of this property S_1 is a possible candidate for being an intentional system since we certainly would not ascribe wants and/or beliefs to any system which is only pushed around by outward forces. Hence, being an automaton is a necessary condition for being a possible candidate for

the ascription of genuine wants and beliefs. And, therefore, the following considerations are meant to apply only to automata.

That non-automata cannot have wants and beliefs, however, characterizes intentional systems only in a negative way. What can we say more positively about the ascription of wants and/or beliefs? Or, to begin with, what can we say more positively about the ascription of wants or want-like states?

Having a certain want is – or at least implies – being disposed to behave in a certain way. Wants involve behavioral dispositions. This much is, I think, philosophical common sense. But if so, what then is the difference between wants and want-like states on the one side and other behavioral dispositions on the other? What is it that makes wants intentional states which are to be characterized by their content? The answer is that the behavior which constitutes the realization of a non-intentional behavioral disposition – e.g., the brittleness of a windowpane or the ignitability of a match – are best characterized in purely physical terms. A plane *breaks* if it is hit by a stone, a match *ignites* if it is struck. On the other hand the behavior which constitutes the realization of an intentional behavioral disposition cannot be characterized this way. If someone wants to open the window then he is, after all, disposed to behave in a certain way. But this behavior cannot be characterized in physical terms. For example, he is not disposed to turn left, move his legs in a physically describable way, raise his arm, etc. Rather he is disposed to do *something which will* – at least, as he believes – *have the effect that the window is open*. The behavior which constitutes the realization of an intentional behavioral disposition thus is best characterized by *the state of affairs which it will probably bring about* or, to put in another way, by the *goal-state G* which it is meant to realize. Hence, having the want that *G* implies having the disposition to do something which is apt to bring about *G* until finally *G* is achieved. So presumably we can start from the fact that any automaton *A* which has the genuine want that *G* satisfies the following two conditions:

- (1) In many situations *A* exhibits a behavior that under the given circumstances is apt to bring about *G*
- (2) *A* ceases to exhibit this behavior if *G* is achieved.

These two conditions express an important point about wants, namely, that wants imply goal-directed behavior, i.e., behavior that (normally)

brings about a certain goal and that comes to an end if the goal is achieved. But condition (1), as formulated above, may lead to difficulties since it does not exclude the case that *A* only under very specific conditions exhibits a behavior appropriate to bring about *G* and, at least in my opinion, we should not ascribe wants or want-like states to a system which would not under varying circumstances do what is under the respective circumstances apt to bring about *G*. Nor should we ascribe wants or want-like states to a system which exhibited the same kind of behavior under all circumstances even if this behavior under all these circumstances would have the same result. Think e.g., of a time-fuse bomb. No one would say that such a bomb wants to destroy its environment and itself at the preset time. So we had better replace condition (1) by the new condition

- (1') In a variety of different situations *A* exhibits a behavior which is apt to bring about *G* and the behavior is not the same in all situations.

Now, as to conditions (1') and (2) it might seem that even the system S_1 exhibits some kind of goal-directed behavior. But with regard to this system there remains the crucial question: What exactly is the goal-state which the behavior of S_1 tends to realize? The description given above makes clear that S_1 has a tendency to move towards the brightest light in its environment, but it mentions no state which will bring the behavior of this system to an end. And so S_1 – like many other tropistic systems – in fact does not have a definite goal. It moves in a certain direction, but there seems to be no definite state which it is really tending to achieve. The situation would, however, be different if S_1 were only to turn around until the output of cell #0 is greatest without then moving straight forward. For in that case the goal-state of this slightly modified system – let us call it S'_1 – could be identified as “the front of S'_1 is facing the brightest part of its environment”. Hence, the system S'_1 satisfies both conditions (1') and (2) and thus exhibits a kind of goal-directed behavior which is at least a necessary condition for the ascription of genuine wants. But obviously this is not yet enough for us to be able to ascribe a genuine want to this system. Hence, the question now is, what more is needed? Or, in other words: what are the reasons for not ascribing genuine wants to tropistic systems like the system S'_1 even though they satisfy conditions (1') and (2)?

I see at least two points that are relevant here. First, though tropistic systems are automata – self-moving systems – they are in an important sense *passive* and not active systems. If we take a closer look at the behavior of S_1 or S'_1 we notice that these systems turn around or stand still *only if and as long as there is a corresponding causal impact from the environment on them*. If a bright enough light in the environment causes one of the photoelectric cells of S'_1 (except cell #0) to emit a greater output than all other cells, then S'_1 turns either right or left depending on which cell emits the greatest output. But this movement stops at the same moment at which the causal impact of the light on S'_1 is inhibited – either by dimming the light, or by turning it off, or by screening, or by whatever means. As I see it, this is a characteristic of tropistic systems in general. If the object they tend to approach or to move away from is prevented from having a causal influence on the system, all behavior stops at once. This is the reason why these systems show a behavior that is in many respects indiscernible e.g., from the behavior of a magnetized knitting-needle under the causal influence of a bar-magnet, i.e., why these systems *seem* to be attracted or repelled by the objects they tend to approach or to move away from.

Imagine in contrast a very simple machine which I would like to call a random seeker. This little system also exhibits a kind of goal-directed behavior since it is designed to stop if it encounters a certain state of affairs. But, as we shall further assume, the behavior of this system is entirely random. It moves straight on for a certain while, then changes its direction by a random angle, moves straight forward for some time, changes its direction again, etc. This little system, it seems to me, is a better candidate for the ascription of a genuine want than any tropistic system for the simple reason that *what* the system does is not entirely dependent on the varying states of its environment. It is not only an automaton, a self-moving system, but in a way also a *self-directed* system.

But let us come to the second point which is, in my view, the decisive reason for our reluctance to ascribe genuine wants to tropistic systems. As we noted before, tropistic systems exhibit in many different situations a behavior which in each of these situations is apt to bring about a certain state of affairs. But in fact they do not exhibit this behavior *because* it is apt to bring about this state of affairs. This comes out clearly if we imagine some possible situations in which e.g.,

the behavior of the system S'_1 would be absolutely inapt to achieve the goal-state of this system. Let us assume for instance that someone picks up S'_1 and suspends it in such a way that the wheels of the system no longer touch the floor. Then, if a bright light is put at an angle of 45 degrees in front of S'_1 (which causes cell #3 to have the greatest output) S'_1 will begin to move its wheels in a certain direction. But, evidently, this will have no effect, the position of S'_1 will not change a bit. Therefore, if the light is not removed or screened, the wheels of S'_1 will continue to move until finally S'_1 runs out of energy. Nothing will have been achieved as far as the goal-state of S'_1 is concerned. And nothing could have been achieved since in the mentioned situation none of the possible movements of S'_1 would have been able to bring S'_1 nearer to its goal-state.

Take another case. Let us assume that the system S'_1 is put on a disc which rotates counter-clockwise at the same speed at which S'_1 normally performs its turns. If we now put a light in the environment of S'_1 in such a way that at the moment it is put there it is located on the right side of S'_1 , S'_1 will begin to turn right, again with the effect that its relative position to its environment including the light remains unchanged. But this situation is different from the situation mentioned before because in this situation S'_1 could do something to achieve its goal: first do nothing until the light is just in front of it and then start turning right. It's easy to see that this doing things just in reversed order would have the desired result. But S'_1 does nothing like this. "Stupid" as it is it clings to doing what it always does even if this will not bring it nearer to its goal. This is speaking metaphorically. In fact this insensitivity to the special aspects of the situation is due to the already mentioned simple design of S'_1 . Every time a certain causal impact on the system occurs it will respond with a certain kind of movement – no matter whether this movement is apt to bring about one state or another. If there is an unscreened bright light on one side of or behind S'_1 , S'_1 will move its wheels in one direction or the other at a preset speed and, if the bright light is just in front of it, S'_1 will do nothing. That this way of moving normally leads to a certain result is, as I see it, an entirely *contingent* fact. Change some features in what can be called a normal situation and the same behavior will lead to completely different results. This makes clear that, even though the behavior of tropistic systems like the system S'_1 is as a matter of fact apt to bring about a certain goal-state in many different situations, it is

nevertheless false to say that systems of this kind exhibit such behavior *because* it is apt to bring about this state of affairs.

These considerations, in my opinion, suggest the conclusion that systems which have genuine wants differ from merely tropistic systems in the following way. Both kinds of systems exhibit in a variety of different situations a behavior which is apt to bring about a certain state of affairs. But with regard to tropistic systems this is no more than a contingent fact. Given a certain causal impact from its environment a tropistic system behaves in a certain way no matter whether this behavior will bring about a certain state of affairs or not. With regard to systems with genuine wants, however, it is not a contingent fact that they show a suitable behavior since these systems behave as they do *because* their behavior is apt to bring about a certain state of affairs. Perhaps one could even say that these systems *choose* their behavior in virtue of its being apt to bring about a certain goal-state. Hence, if an automaton *A* is a system with genuine wants it seems to satisfy not only conditions (1') and (2) but also the additional condition

- (3) The suitable behavior of *A* is due to an internal mechanism which *chooses* this behavior in virtue of its being apt to bring about *G*.

It might seem strange that in the formulation of this condition the notion of "choice" is used though this notion, too, seems to have at most a metaphorical application in this context. But there are good reasons for using this notion here. For, if it is not a contingent fact with regard to systems with genuine wants that their behavior in varying circumstances is apt to bring about a certain goal-state, then this behavior *must be caused by an internal mechanism which ensures that this is indeed the case*. This mechanism, therefore, must effect that of all possible actions of a system that action is executed which under the given circumstances will bring about the desired goal-state – granted that there is such an action at all. Such a mechanism, I think, can well be called a choice-mechanism. Hence, there is nothing mysterious in using the notion of choice in this context.

To bring this out a little bit more I would like to sketch a possible mechanism of this kind. Imagine a system *S* which is able to exhibit three different kinds of movement *a*, *b*, and *c*. Let us assume that this system has an internal model of its environment (I shall say something

about what internal models presumably are later on) and let us further assume that it is able to produce for any given model M and for any of its possible movements a , b , and c an effect-model M_x , which would be a model of the changed environment of S , which in turn would be the result of its carrying out x . Now, a simple choice mechanism could work as follows: In a given situation, which is represented in S by the model M , the mechanism first checks whether M is the model of a situation in which the goal-state G of S is realized. If so it causes S to do nothing, if not it produces the effect-model M_a . It checks whether M_a is the model of a situation in which G is realized. If so it causes a to be carried out, if not it produces the effect-model M_b . It checks whether M_b is the model of a situation in which G is realized. If so it causes b to be carried out, if not it produces the effect-model M_c . Finally it checks whether M_c is the model of a situation in which G is realized. If so it causes c to be carried out, if not it stops and thereby causes S to do nothing. It is easy to see that this mechanism does what it is supposed to do. If there is a behavior which will lead to the desired goal-state G it will bring it about that S exhibits this behavior. If not it will bring it about that S does nothing. And even that is what one would expect a reasonable system to do which has the want that G .

Obviously, the mechanism just sketched is nothing else than the realization of a rudimentary form of a well known method in the field of mechanical problem-solving: the traversing of a search-tree in order to find a path to the solution of a problem. And this being so, it would not be difficult to enlarge this mechanism to enable it to check not only single actions but whole series of actions to find out whether they would lead to a given goal-state. But I do not want to labor this point further. For in the present context my aim is only to make clear that the having of genuine wants implies the capacity to make plans or at least the capacity to evaluate possible actions with regard to whether in the given circumstances they will bring about a certain (desired) state of affairs. And to make clear that tropistic systems are not the right kind of things to have genuine wants since they completely lack these capacities.

I think that the reason for drawing the dividing line just here can also be brought out by the following consideration. If we have a system A with an internal choice-mechanism that ensures that in any given situation A exhibits one of those actions that lead to a certain

goal-state G – if there are any – then the behavior of this system can be explained by the law-like sentence

- (4) In any situation A exhibits a behavior that brings about G , provided that there is such a behavior.

And this law-like sentence ascribes an intentional (want-like) state to A since it says in effect that A has a disposition which is characterized by reference to a certain state of affairs. Hence, this state has a *content*, namely, the goal-state G .

On the other hand the behavior e.g., of system S'_1 cannot be explained by the law-like sentence

- (5) In any situation S'_1 exhibits a behavior that brings about that the front of S'_1 is facing the brightest part of its environment – provided that there is such a behavior.

but rather by the three lawlike sentences

- (6)a. If the brightest part of its environment is at the right side of S'_1 , S'_1 moves its wheels in one direction.
 b. If the brightest part of its environment is at the left side of S'_1 , S'_1 moves its wheels in the other direction.
 c. If the brightest part of its environment is just in front of S'_1 , S'_1 does nothing,

since it is by these three statements that we are able to explain what S'_1 does in the two above mentioned queer situations – and in all other situations as well – and not by the statement (5). And this, I think, is the reason why, in the end, we are not entitled to ascribe to S'_1 the intentional (want-like) disposition that corresponds to the lawlike sentence (5).¹

3.

After these remarks on why tropistic systems do not have genuine wants, let us now turn to the second part of my thesis, the claim that tropistic systems do not have genuine beliefs either. One of the key concepts which have a special significance with regard to this claim has already been mentioned – the concept of an “internal model” or an “internal representation”. It is not the only important concept in this context. But I shall restrict my considerations to the claim that

tropistic systems have no internal representations of their environment.

If again we take system S_1 as an example of a tropistic system it might seem – at first sight – that this claim is plainly false and that on the contrary there are good reasons for attributing such internal representations to the system, since it is equipped with a subsystem (the circle of photo-electric cells on the top of S_1) whose states might well be interpreted in this way. For, within the limits imposed by the number and the sensitivity of the cells, each state of this circle stands in an obvious one-to-one relation to a certain state of the environment. In the same way that the above mentioned state in which cell #3 has a considerable output while all other cells remain silent corresponds to the environmental state “bright light at an angle of 45 degrees in front of S ”, the internal state “cell #6 high output, all other cells low output” corresponds to the environmental state “bright light exactly on the right-hand side of S_1 ”, and the internal state “cell #12 high output, all other cells low output” corresponds to the environmental state “bright light exactly behind S_1 ”.

Moreover, the states of the circle of photo-electric cells have a feature which any system of representing states should have. *They change in rather perfect congruence with corresponding changes in the environment.* If a bright light moves steadily from the position shown in figure 1 to a position behind system S_1 the states of the circle change correspondingly, e.g., from the state “cell #3 high output, all other cells low output” to the state “cell #4 high output, all other cells low output” until finally state “cell #12 high output, all other cells low output” is reached. So there is a kind of isomorphism between the states of the circle and certain states of the environment. And this isomorphism also holds if system S_1 itself moves, since then the states of the circle also change in an appropriate way. This is certainly necessary. For what the states of the circle stand for – if they stand for anything at all – is something that can only be expressed in system-relative coordinates. Hence, summing up we can say that there exists a set of internal states of system S_1 such that there is an isomorphism between the elements of this set and certain states in the environment of S_1 . That is to say, for any state e of these environmental states there exists a corresponding internal state r with the property that S_1 is in the state r if and only if the environment is in state e . Why not say that the internal state r is a representation of the environmental state e ?²

In my opinion the answer to this question is simply that we do not need to explain the behavior of system S_1 by reference to any internal representations whatever. And this, in turn, is simply due to the fact that the states of S_1 which could possibly count as representing certain states of the environment *occur only if and as long as they are caused by these environmental states*.³ If one wants to explain a certain piece of behavior of S_1 by reference to an allegedly representing state r of S_1 , one can explain this piece of behavior just as well – or perhaps even better – by reference to the state which r is said to represent. For example, if we want to explain a certain behavior of S_1 by the fact that S_1 is in the internal state “cell #6 high output, all other cells low output” we could as well or even better say that S_1 behaved the way it did because there was a bright light on the right-hand side of S_1 .

So there is one negative lesson to be learned here about internal representations. No internal state r of a system S will count as representing an environmental state e if r occurs only if and as long as it is caused by e . *Internal representations must have a certain degree of causal independence from the states they represent*. More positively, this can be illustrated by a slightly modified version of S_1 , let us call it S_2 . Imagine the following situation. There is a bright light in the neighbourhood of S_1 which is not standing still, but moving slowly in a certain direction. In this situation system S_1 will, according to its construction, follow the light as if it were trying to reach it. But imagine further that now for a short while the light disappears behind a screen or a wall or some other kind of non-transparent object. If we assume that it is rather dark in the neighbourhood of S_1 while the light is behind the screen, S_1 clearly will stop and do nothing until the light appears again. And then it will turn to the new direction and move on.

Now let us suppose that the modified version of S_1 , S_2 , has a rather different construction according to which things happen as follows. If the target of S_2 – e.g., a light – disappears for a while behind a screen, then S_2 by means of some internal mechanism extrapolates the previous course of the target and by this means calculates a probable position of the target. Then S_2 moves as if there were a bright light at this calculated position. Obviously, S_2 will move the whole time as S_1 would move if there were no screen. And this means that S_2 – given that its calculations are correct – will follow its target even if it is not able to “see” it; better: even if the target has no causal impact on system S_2 .

For this reason the behavior of S_2 cannot be explained by the target alone. It is true: if the calculations of S_2 are correct, the system of internal states, which corresponds to the various positions of the light, is the whole time in perfect accordance with the real position of the light. But this is not due to the fact that the system of these states is the whole time appropriately changing under the causal influence of the moving light. Rather the updating is – at least sometimes – accomplished by an internal mechanism of S_2 itself. Accordingly, it is not possible too explain the behavior of S_2 solely by reference to the real position of the light. For while the light is behind the screen it has no causal effect on S_2 whatsoever. At least during this period of time S_2 is therefore following its own compass, i.e., at least during this time its behavior is effected solely by its own states one of which seems really to be a representation of the actual position of the light behind the screen.

I think there are two general points that must be noted here. First, it is not possible to explain the behavior of S_2 entirely by reference to certain features of its environment. This corresponds to the fact that the behavior of S_2 is non-uniform in the sense that the system behaves differently in situations of the same kind, i.e., situations which are qualitatively the same for the system itself, which have the same causal impact on the system. If e.g., two lights with different trajectories disappear for a while we will have darkness in both cases and that means the two situations are qualitatively the same for S_2 ; but S_2 will behave differently in these two situations according to its different calculations of the paths of the lights. Therefore, if we want to explain the behavior of S_2 at all, we must take into account the internal states of the system. This, I think, really is a very general point. We must ascribe different internal states to a system S if its behavior is non-uniform (but not random) in the sense that the system behaves differently in situations which have the same causal impact on it. Second, the internal states which are in fact responsible for the behavior of S_2 belong to a system of states which – partly by internal, partly by external causes – change in such a way that these states of S_2 are the whole time in exact correspondence to certain states of its environment. In my opinion, these two features make system S_2 a very plausible candidate for the ascription of internal representations.

To put it in more general terms: seen from outside, we are entitled to ascribe internal representations to a system if the system behaves in

certain situations *as if* its behavior were caused by a certain state of its environment while in fact the behavior is caused by an internal state of the system. Seen from within, we are entitled to ascribe internal representations to a system if its behavior is caused by a set of internal states which (at least normally) change in perfect accordance with corresponding changes in the system's environment and if the changes of these states of the system are not simply caused by the corresponding states of the environment themselves but are – at least in part – due to an *internal updating mechanism*. These two criteria, I think, are by and large two sides of the same coin.

These considerations bear a certain similarity to the considerations presented by J. A. Fodor in his recent paper 'Why Paramecia Don't Have Mental Representations', in which he tries to develop a criterion for having mental representations and thereby also to give a necessary condition for being an intentional system. Since, according to Fodor, a system without mental representations a fortiori cannot be a genuine intentional system, his main suggestion in this paper is that the fundamental difference between systems with and systems without mental representations lies in the fact that systems with mental representations are able to respond selectively to non-nomic properties of objects in their environment whereas systems without mental representations lack this ability. At least:

... *any* system that can respond selectively to non-nomic properties is, intuitively speaking, a plausible candidate for the ascription of mental representations; and any system that can't, isn't. (Fodor 1986, p. 11)

Obviously, this suggestion presupposes that we can make sense of the idea that there are non-nomic properties, i.e., properties which do not enter in any lawful relations. But if we grant the existence of such properties here for the sake of argument, the rationale of Fodor's proposal seems to be this. Even some tropistic systems are able to respond selectively to some specific features of their environment, say e.g., to the fact that an object *O* in their neighbourhood has a certain property *P*. Let us assume that *S* is such a system. How then has the selective behavior of *S* been brought about? In general the answer to this question will be: *O*'s having *P* causes a certain state in *S* to occur and this state in turn causes the selective behavior of *S* or, in more simple cases, *O*'s having *P* directly causes the behavior of *S*. Now, if *P* is a non-nomic property then *O*'s having *P* per definitionem does

not cause anything. How then can systems react selectively to non-nomic properties at all? According to Fodor there is only one possibility: the behavior of the system has to be effected by *S's representing O as P*. In these cases it is not *O's having P* that causes – directly or indirectly – the behavior of *S*, but a state of *S* which is a representation of *O's having P*. Fodor's argument, therefore, can be summed up in the following way. Every system which is able to react selectively to non-nomic stimuli must possess mental representations. For non-nomic stimuli themselves cannot be causes of behavior, so there must be intermediate states, representations of the stimuli, which take over this role.

Let us take a closer look at this argument. If a system *S* responds selectively to a certain non-nomic property *P*, then this system behaves differently in situations with and situations without an object with the property *P* (short: *P*-object) in its environment. That is to say, it shows a certain behavior *B* if and only if there is a *P*-object in its neighbourhood, or in more general terms: if and only if a *P*-object is standing in a certain relation to it. But, since *P* is a non-nomic property *S's* doing *B* cannot be caused by the fact that there is a *P*-object in the neighbourhood of *S*. And since obviously *S's* doing *B* cannot be caused by any other feature of its environment either (otherwise it could not count as a selective response to the property *P*), it must be caused by some internal state of *S*. Let us call this state *Z*. Now, since *S* does *B* every time it is in state *Z* and since *S* does *B* if and only if there is a *P*-object in its environment, there is after all an internal state *Z* of *S* such that *S* is in *Z* if and only if there is a *P*-object in the environment of *S*. (If doing *B* also can be caused by another internal state *Z'* of *S* we can choose the disjunctive state "*Z* or *Z'*" instead of *Z* alone.) But, since *P* is non-nomic not even *S's* being in state *Z* can be caused by there being a *P*-object present. Hence, according to Fodor, we are forced to say that a certain system *S* has internal (mental) representations of its environment if, first, *S* has an internal state *Z* which corresponds perfectly to there being a *P*-object in its environment, if, second, being in state *Z* is the cause of *S's* doing *B*, and if, finally, *S's* being in state *Z* itself is not caused by there being a *P*-object in the environment.

So, there is a rather similar structure between Fodor's paradigm case for the ascription of internal representations and the conclusions drawn from the example of system *S'₁*. But there remains a difference

in that Fodor requires the property *P* which is represented by state *Z* to be a non-nomic property. In my eyes, however, that is asking too much. For, as I see it, the kind of argument that Fodor puts forward applies already to cases in which a system behaves *as if* this behavior were caused by a certain state of the environment while *in fact* it is not caused this way even if the state of the environment does not consist in the presence of an object having a non-nomic property. That is to say, the basic criterion is not that a behavior which constitutes a selective response to a specific feature of the environment *cannot* be caused by this feature but that it is *in fact* not caused by it.

4.

It is time to give a short résumé. My claim was that there are good reasons for not ascribing genuine wants and beliefs to tropistic systems and hence for not reckoning these systems among the genuine intentional ones. These reasons can be summarized in the following way.

We should not ascribe genuine beliefs to tropistic systems because they have no internal representations of their environment. And they do not have such internal representations because the only feasible internal states of these systems do not have the necessary causal independence of the environment. For they occur only if and as long as they are caused by the environmental states they presumably stand for. It is this fact which also has the consequence that in explaining the behavior of tropistic systems we are not forced to take internal representations into account. For everything that can be explained by allegedly representing states can as well – or even better – be explained by the corresponding environmental states.

The reasons for not ascribing wants to tropistic systems are in a way very similar. For on the one hand it is true that tropistic systems often show a kind of goal-directed behavior. But on the other hand everything that a tropistic system does is caused by certain states of its environment. As a consequence of this fact a tropistic system can easily be ‘tricked’ by its environment into exhibiting a certain behavior even in situations in which this behavior has no chance to bring about the “desired” goal-state. Hence, tropistic systems have no genuine wants because they do not choose their behavior in virtue of its being apt to bring about a certain goal-state. Or to put it in more general terms: Because their behavior is not caused by an internal mechanism

which ensures that the behavior is exhibited because under the given circumstances it is apt to lead to the “desired” result.

These formulations, however, force me to add a last note. I have deliberately tried to avoid the problems of misrepresentation and miscalculation⁴ even though they have lurked around every corner. For it is certainly not true that an internal state *r* can represent an external state *e* only if *r* occurs if and only if *e* occurs. And it is certainly not true that a choice-mechanism which evaluates certain actions with respect to their being apt to bring about a certain result can be called a respective choice-mechanism only if it never makes mistakes. But a thorough discussion of these problems would have led us too far away from my original claim, the claim that tropistic systems have no genuine wants and beliefs.

NOTES

* Obviously, there is a certain resemblance between the title of this paper and the title of Fodor’s recent paper ‘Why Paramecia Don’t Have Mental Representations’. This is no accident, since in this paper I address myself to almost the same questions as Fodor did in his. I shall come back to Fodor’s answers at the end of section 3. An earlier draft of this paper was presented at a conference on ‘Aspects of Consciousness and Awareness’ at the ZiF in Bielefeld. I am very much indebted to W. Ewald and K. Hillebrandt for correcting my English.

¹ Now, this certainly is a little bit of cheating. For even if all systems *without* choice mechanisms are bound to make mistakes, systems *with* choice mechanisms also will make mistakes now and then. For absolutely fool-proof choice mechanisms, which never make any mistakes of any kind, will be at least very rare (in fact they are, as I think, impossible). So, the observable behavior of a system may not be decisive. And this leads to the conclusion that the real difference lies in the way the behavior of a system may not be decisive. And this leads to the conclusion that the real difference lies in the way the behavior of a system is produced – by means of an internal choice mechanism or by some other device (like a simple feed-back mechanism or whatever). In this respect my account of wanting is not purely functionalistic, but has some implications as to the internal structure of a system.

² Compare e.g., F. Dretske (1981). In chapter 7 of this book Dretske defines “semantic content” in the following way:

- (1) Structure *S* has the fact that *t* is *F* as its *semantic content* iff *S* carries the information that *t* is *F* in digital form. (p. 177)

And before that, in chapter 6, he writes:

- (2) Structure *S* carries the information that *t* is *F* in digital form iff *S* carries no additional information about *t*, no information that is not already nested in *t*’s being *F*. (p. 137).

From these two definitions one can derive, as it seems to me:

- (3) Structure *S* has the fact that *t* is *F* as its *semantic content* iff it is true that *S* occurs if and only if *t* is *F*.

At least, this is possible if we interpret (2) in the following way:

- (2') Structure *S* carries the information that *t* is *F* in *digital form* iff the following two conditions are satisfied:
 (a) if *S* is the case then *t* is *F*
 (b) there is no $F' \neq F$ with: if *S* is the case then *t* is F' , and if *t* is F' then *t* is *F*.

It must be noted here, however, that Dretske – in order to account for the possibility of false beliefs – does *not* identify beliefs with states having a certain semantic content.

³ This already comes out if we ask again what exactly the environmental state is which e.g., the internal state “cell #6 high output, all other cells low output” stands for. In the penultimate paragraph I said that this is the state “bright light on the righthand side of S_1 ”. But strictly speaking this is not true. For if we put a screen or a wall or some other non-transparent object between the light and the system S_1 , S_1 will no longer be in the mentioned state. And therefore it is simply false to say that S_1 is in this state if and only if there is a bright light on its righthand side. What one could say at most therefore seems to be that S_1 is in the state “cell #6 high output, all other cells low output” if and only if there is an *unscreened* bright light on its righthand side. And the “unscreened” here is best understood figuratively as “not blocking the causal impact of the light on the system S_1 ”.

⁴ Except for the few remarks in note 1 above.

REFERENCES

- Dretske, F. I.: 1981, *Knowledge and the Flow of Information*, Oxford.
 Fodor, J. A.: 1986, ‘Why Paramecia Don’t Have Mental Representations’, *Midwest Studies in Philosophy* 10, 3–23.

Manuscript received 5 August 1987

Philosophisches Seminar
 Georg-August-Universität
 Nikolausberger Weg 9c
 3400 Göttingen
 Germany