Some Foundational Issues Concerning Anticipatory Systems

Ron Chrisley
School of Computer Science, University of Birmingham, United Kingdom
Fax: +441214144281; rlc@cs.bham.ac.uk; http://www.cs.bham.ac.uk/˜rlc/

Abstract Some foundational conceptual issues concerning anticipatory systems are identified and discussed: 1) The doubly temporal nature of anticipation is noted: anticipations are directed toward one time, and exist at another; 2) Anticipatory systems can be open: they can perturb and be perturbed by states external to the system; 3) Anticipation may be facilitated by a system modeling the relation between its own output, its environment, and its future input; 4) Anticipations must be a part of the system whose anticipations they are. Each of these points are made more precise by considering what changes they require to be made to the basic equation characterising anticipatory systems. In addition, some philosophical questions concerning the content of anticipatory representations are considered.

Keywords: computing anticipatory systems, weak anticipatory systems, modeling, temporal representation, representational content.

1 Introduction

This paper discusses some fundamental issues surrounding the notion of anticipatory systems. The goal is a characterisation of anticipatory systems that is general enough to include not only simple cases of anticipation, but also some sophisticated forms of modeling. Of course, there are many anticipatory systems that lack this sophistication, but it will be beneficial to have a general theory with which we can understand both the sophisticated and less sophisticated systems as special cases of the same general form.

A clear definition of anticipatory systems is given in [Rosen, 1985]:

> An anticipatory system is a system containing a predictive model of itself and/or of its environment, which allows it to change state at an instant in accord with the model's predictions pertaining to a latter instant.

This is frequently made more formal by saying that an anticipatory system is a system X whose dynamical evolution is governed by:

$$X(t+1) = F(X(t), X^*(t+1)), \tag{1}$$

where $X^*(t+1)$ is X's anticipation, or prediction, of what its state will be at time $t+1$.

Because of a prior interest in cognitive science and artificial intelligence, which require explanation in causal terms, the discussion in this paper is restricted to anticipatory systems for which eqn. 1 and its descendant forms capture a *causal* relation. It will not be enough for a system merely to be *describable* in terms of such equations; the equation must model causal dependencies in the systems characterised. Thus, the systems under discussion are what have been called "weak" anticipatory systems (e.g., [Dubois, 2000]).

One can go further and inquire as to the extent to which such causal anticipatory systems are *computational*. Are there interesting sub-classes of causal anticipatory systems which compute their models of self and world? Which compute their next state based on their models of self and world? Must all causal anticipatory systems compute in this sense, or can there be non-computational yet causal anticipatory systems? These questions will have to be addressed at a later date. For now, no assumption is made as to the computational or non-computational nature of these systems; only causality is assumed.

2 The Doubly Temporal Nature of Anticipatory Systems

The intuitive notion of anticipation is of a relation involving not one but two points in time: not only the time of the state being anticipated, but the time at which the anticipation is occurring. In general, systems can anticipate more than one state at a time, and can, at different times, anticipate what things will be like at a given time. I can anticipate what December 25th will be like on both December 23rd and December 24th; and on December 25th 2002 I can anticipate what it will be like on December 25th 2003 and December 25th 2004. Therefore, the term $X^*(t+1)$ is only well defined for a particular subclass of constrained anticipatory systems, systems for which the function $X^*(t)$ is constant over the temporal extent of that anticipation. For anticipatory systems described by eqn. 1, the constancy of this function is trivially ensured by virtue of the fact that a state is only ever anticipated once, at the time immediately prior to the time of the anticipated state.

Thus, a natural generalisation of the formal notion of an anticipatory system would be to parameterise the anticipation function X^* with the time of anticipation (second parameter) in addition to the time anticipated (first parameter). The special case of eqn. 1 is then expressed as:

$$X(t+1) = F(X(t), X^*(t+1, t)). \tag{2}$$

But we now have a way of expressing a wider variety of anticipatory systems. For example, a "reminiscent" system which takes all of its past anticipations concerning the next time step into account in determining its next state would be described by:

$$X(t+1) = F(X(t), X^*(t+1, t), X^*(t+1, t-1), X^*(t+1, t-2)...). \tag{3}$$

However, having an unrestricted number of arguments may be problematic. Furthermore, there is the intuition that taking into account all of one's previous anticipations is typically done by modifying one's current anticipation. That is, when one takes past anticipations into account, one does so by altering one's current provisional anticipation $X^{**}(t)$ in the light of one's (memories of) previous anticipations $X^*(t - \tau)$. So perhaps a better way of modeling this kind of dependence would be to stipulate the recursive definition:

$$X^*(\tau, t) = G_{X^*}(X^{**}(\tau, t), X^*(\tau, t-1)), \tau > t. \quad (4)$$

With this definition we can capture both "reminiscent" and the more usual "forgetful" anticipatory systems, providing suitable alterations are made to G_{X^*}. For the case of the usual, "forgetful" system, one could have:

$$G_{X^*}(x, y) = x, \quad (5)$$

while a G_{X^*} which truly depends on both of its arguments would provide a form of reminiscence.

(One could allow for a more general system by adding a temporal parameter to G to allow it to vary with time. This same point goes for all the functions internal to X that are introduced below (e.g., I^*, F^*, E^*, O^*, etc.). To avoid further complexity, these possibilities will not be discussed further in this paper, and will not be expressed in the notation.)

3 The Environment and Openness

The formulations of the notion of an anticipatory system just considered still fall short of the general case. Rosen's intuitive notion includes anticipation of states external to the system, while the equations considered so far only model anticipation of the states of the system itself. Admittedly, [Rosen, 1978a] shows how a formulation in terms of self-anticipation can be augmented with a simple mapping in order to cover the case of the anticipation of another system's states, but an explicit inclusion of both cases into the formalism would facilitate the analysis of many systems.

The focus on the case of self-anticipation goes hand-in-hand with a presupposition on the part of the equations considered so far: that X is a closed system; X's evolution does not depend on anything outside X. Yet Rosen's intuitive notion, if not his formal analysis, takes the general case of an anticipatory system to be an open system, receiving "inputs" or "data" from the world. Although Rosen stresses that an anticipatory system acts on the basis of its model, not on the data directly, he nevertheless assumes, even requires, that the model be formed from the input data:

> We stress again that the common feature of all these examples is the transduction of present data into future data (i.e. into predictions) through the agency of a model of the world. The essential point to bear in mind is that it is the *prediction*, rather than the initial data, which is the actual stimulus, and it is the model relating the two which underlies the adaptive character of the behavior so generated [Rosen, 1978a, page 156, original emphasis].

(This passage will also be relevant when discussing the assumptions of causality in anticipatory systems, especially computing anticipatory systems.)

Once these two possibilities, of other-modeling and openness, are explicitly acknowledged, a question arises as to how to incorporate them into the formalism of anticipatory systems. Let us first consider the generalisation of X from a closed to an open system.

3.1 Openness I: Input

To accommodate openness, one need only add an input function $I(t)$ which can provide a point of coupling between X and any other system, open or closed:

$$X(t+1) = F(X(t), X^*(t+1,t), I(t)), \tag{6}$$

Thus, I will in general be a function from the real numbers to a vector of variables which capture the aspects of the environment which impinge on X.

(Note concerning notation: Like many of the functions introduced in this paper, I is dependent on the system X under consideration, but rather than clutter the notation with a ubiquitous X subscript, the dependency has been elided.)

But just as X's future state depends not only on its current state but also on its anticipation of its future state, so also will its future state depend not only on the actual input it receives but also on its anticipation I^* of what that input will be. For an open system adequately to predict its future state, it must model not only itself but also its inputs, since its future state will depend on those inputs. And like X^*, I^* will be doubly temporally indexed, to reflect the time of the anticipation as well as the time of the anticipated input.

However, a question arises concerning how best to reflect this possibility in the formalism. One option is to add the expectation of input explicitly into eqn. 6, but this would not properly reflect the role I^* is playing in the system. The function of I^* is to assist in the computation of the anticipation X^{**}. This computation requires a model, F^*, of how the state of X is affected by its input I. So a better formulation would not involve a change to the basic state equation for X, but to the equation defining the current contribution X^{**} to the anticipation X^* of X:

$$X^{**}(\tau, t) = F^*(X^*(\tau-1,t), I^*(\tau-1,t)), \tau > t. \tag{7}$$

Note that this does not duplicate what was captured in eqn. 4. To illustrate, eqn. 4 captures how your past anticipations of what will happen on December 25th affect your current anticipation of what will happen on that date. Eqn. 7, on the other hand, captures the effect that your current anticipation of what will happen on December 24th has on your current anticipation of what will happen on December 25th.

Furthermore, we can, in a manner similar to what we did for $X^*(\tau, t)$, make $I^*(\tau, t)$ dependent on previous estimates $I^*(\tau, t-n)$:

$$I^*(\tau, t) = G_{I^*}(I^{**}(\tau, t), I^*(\tau, t-1)), \tau > t. \tag{8}$$

where:

- $I^{**}(\tau, t)$ is X's provisional anticipation of $I(\tau)$ at time t;
- G_{I^*} is a function similar to G_{X^*}, above, which models the effect, if any, that previous anticipations of $I(\tau)$ have on the current anticipation $I^*(\tau, t)$ of $I(\tau)$.

At this point, a general methodological pattern emerges from two reasonable principles:

- **Principle 1:** If the best model we can have of a system x's basic dynamics is given by an equation e, then the best model that x can have will have the structure of e.
- **Principle 2:** The modeling relation distributes over functional composition; that is, the following postulate seems plausible:A model x^* of a system $x = f(y, z)$ should take the form $x^* = f^*(y^*, z^*)$.

If we were to take these principles at face value, eqn. 6 would imply the following definition of X^*:

$$X^{**}(\tau, t) = F^*(X^*(\tau - 1, t), (X^*)^*(\tau - 1, t), I^*(\tau - 1, t)), \tau > t. \tag{9}$$

(One might raise the issue here that X needs to model the variables τ and t as well, but this will not be discussed in this paper.)

Of particular interest is the second argument to F^*, X's model of its model of itself, $(X^*)^*$. Obviously, this is the beginning of a regress that can be iterated indefinitely, providing arbitrary levels of reflection. Until further analysis of this aspect of anticipatory systems can be carried out, such possibilities will be ignored: we will restrict ourselves to systems describable in terms of eqn. 7.

3.2 Openness II: Output

A major insight of the cybernetic and general systems approaches is that systems and their environments mutually determine their states in a dynamic coupling. That is, X typically will have some effect on the world, or "output", which typically will alter the state of the world. This change in world state may in turn alter the inputs to X, and so on. This can be captured by the following equations:

$$O(t) = H(X(t)) \tag{10}$$

$$I(t + 1) = J(E(t)) \tag{11}$$

$$E(t + 1) = K(E(t), O(t)) \tag{12}$$

where:

- $O(t)$ is the output of X at time t;

- H is a function which determines what the output of X is given X's state;
- $E(t)$ represents the state of the environment, excluding X, but including anything which may potentially causally impinge on X within one unit of time from t;
- J captures how the state of the environment determines the input to X at the next time step;
- K captures how the current state of the environment, along with X via its output O, determine the next state of the environment.

Some explanations of the chosen form of these equations can be given.

It might be thought that J should depend on X; that is, one might believe that the environment alone does not determine the input to X. Consider the case of, say, a paramecium receiving inputs via stimulation of its cilia. A stimulation on the left side of the organism is distinct from a stimulation on the right side. Yet the environmental conditions which cause these distinct inputs may be identical (stimulation impinging on the organism from the north, say); the only difference may be the orientation of the paramecium within that environmental space (the left side facing north vs the right side facing north). In this example, the question concerning J comes down to: is the orientation of the paramecium part of the state of the environment in which it is located, or part of the paramecium itself? Properly speaking, it is neither; it is a relation between the two. But such relations are difficult to capture in the function-based framework being considered here, so I have arbitrarily decided to model as part of E the relations between X and E which affect the input to X. This has the advantage of simplifying eqn. 11 by avoiding a reference to X.

One might wonder if it would be better to eliminate some of the variables, by not having input and output functions I and O, and instead having X and E co-defined, such as:

$$X(t+1) = F(X(t), X^*(t), E(t)); \tag{13}$$

$$E(t+1) = O_{alt}(X(t)), \tag{14}$$

where O_{alt} captures the effect that X has on the environment. While technically adequate, eqns. 13 and 14 fail to reveal and exploit structure that we know to exist: the regions of interface captured by I and O in eqns. 10 and 11. Highlighting these interfaces is not an arbitrary act; the very notion of a system is that of a relatively dense region of dynamical activity which is marked off from other clusters of activity by a perimeter of relatively less activity. Input and output are the migrations of casual effect across this perimeter. Merely noting that system depends on environment which depends on system ignores the reduction in complexity that can be achieved by acknowledging this perimeter.

Given that X's future inputs may depend on its own outputs, it follows that in order for X properly to anticipate its future inputs and thus its own future states, it must anticipate its own outputs:

$$O^{**}(\tau, t) = H^*(X^*(\tau, t)), \tau > t; \tag{15}$$

$$O^*(\tau, t) = G_{O^*}(O^{**}(\tau, t), O^*(\tau, t-1)), \tau > t, \tag{16}$$

where:

- O^* and O^{**} capture X's anticipations of its output in the same way as X^* and X^{**} do for anticipations of the state of X, and as I^* and I^{**} do for anticipations of the input to X, respectively;
- H^* is X's model of H: of how the output of X is determined by the state of X;
- G_{O^*} plays the same combining role for O^* and O^{**} as G_{X^*} does for X^* and X^{**} and G_{I^*} does for I^* and I^{**}.

3.3 Modeling Self by Modeling Environment

By Principles 1 and 2, X should take these anticipations of its output into account, along with an anticipation of the state of the environment, when forming an anticipation of its input. This can be modeled by creating the anticipatory versions of eqns. 11 and 12, along with an equation which does for X's model of the environment, E^*, what eqns. 7 and 8 do for X^* and I^*, respectively:

$$I^{**}(\tau, t) = J^*(E^*(\tau, t)), \tau > t; \tag{17}$$

$$E^{**}(\tau, t) = K^*(E^*(\tau, t-1), O^*(\tau, t-1)), \tau > t; \tag{18}$$

$$E^*(\tau, t) = G_{E^*}(E^{**}(\tau, t), E^*(\tau, t-1)), \tau > t. \tag{19}$$

Since we are trying to produce a model of X that allows for sophisticated anticipatory capabilities, every elaboration on the dynamics of X prompts a corresponding elaboration of the dynamics of X^*.

The analysis of the environment given by eqn. 12 will in most cases be inadequate. Most environments are too complex to be understood in terms of a single function E; the simplification that results from, e.g., dividing the environment up into objects, locations, events, properties, etc. is necessary. But if *we* require a structured model of the environment, then it is very likely that any sophisticated anticipator, X, will require one also. That is, if changes are made to E, then by principles 1 and 2, corresponding changes will be made to E^*.

However, perhaps we have over-applied Principles 1 and 2. Some applications of the Principles still seem correct: A typical open system X can't hope to anticipate its own states unless it models its inputs. But even if we are interested in the most sophisticated anticipators, why should X's model of the world have to mirror our own? Or, putting the question into epistemological terms, why should we be bound to believe that X's model of the world must mirror our own? Perhaps X needs to have structure in its model of E if it is to have any hope of success, but must we assume that its model be isomorphic to our own? Can we not make sense of a system that models the world in terms of objects, locations events and properties which we do not? If so, what becomes of Principles 1 and 2: how can they be altered so that they require mirroring between our model and X^* when, and only when, such mirroring *is* actually a necessity?

Here we quickly get into deep philosophical waters; a full treatment cannot be given here. But one point can be made: there are some (e.g. [Davidson, 1974]) who have argued that this notion of a conceptual scheme different from ours is incoherent. In suggesting that I can specify an X that carves up the world differently than I do, a contradiction appears. If I really can specify such an X, then I must possess the concepts used in that specification, and so X's scheme is not distinct from my own. Thus, there is a strong mirroring constraint between our model and X^*.

The preceding argument seems a bit hasty, however. It assumes that the only way to specify a model is by using the same concepts as that model. That is, if I am to specify an X that anticipates environmental states with a model E^* that employs the concept *things within reach*, then I will have to use the concept *things within reach* in doing so. We can call specifications which incur this requirement *conceptual* specifications. But if there are *non-conceptual* specifications, ones which do not require the theorist specifying a model to possess the concepts that the specified model employs, then the possibility of specifying an X which models the world in a manner radically different from our own reappears. (The distinction between conceptual and non-conceptual specifications made here is similar to, but crucially different from, the distinctions made in the literature on non-conceptual content; e.g. in [Cussins, 1990].) One way of making a non-conceptual specification of X^* would be to describe the causal mechanism that realises X^*; but the benefit of such a specification can only be retained if one gives up the ability fully and correctly to analyse the causal mechanism on a conceptual level.

4 A System's Anticipations are Part of that System

In section 1 it was stated that an emphasis on anticipatory systems which compute, and compute using, their anticipations requires an analysis which reflects the causal structure of the system. In section 2, we saw that this point had the result of of dividing the aspects of a system which realise an anticipation (of $X(t+1)$, say) into two components: the current contribution to the anticipation, and the part of the current state which carries the effect, if any, of previous anticipations of $X(t+1)$.

A similar point can be made concerning the evolution of X itself. The first thing to notice is that if we want the structure of our models to illuminate the causal structure of the systems we are modeling, then we need to modify all of our equations which define X (e.g., eqn. 6). Specifically, the anticipations that X has must be, strictly speaking, part of system X itself. Consider the first sentence of the abstract of [Rosen, 1978a]:

> An anticipatory system is one which contains a *subsystem* which can serve as a predictive model of the world [p 155, emphasis added].

(Note that this definition mentions modeling *the world*, rather the system itself, which supports the approach in section 3.)

Since the anticipations of X, be they of X itself or some external system, are subsystems of X, it is redundant to have future values of X depend both on the total state of X *and* its anticipations. So rather than using something based on eqn. 6, one might prefer:

$$X(t+1) = F(\bar{X}(t), X^*(t+1,t), I(t)), \tag{20}$$

where $\bar{X}(t)$ is the state of the non-anticipatory parts of X at time t.

However, while eqn. 20 might delineate an interesting and manageable sub-class of anticipatory systems, it may fail to capture the general case. Specifically, it seems possible for there to be a system which *is* redundant in the way just described. It might be that the physical aspects of a system which realise an anticipation might have an effect on the behaviour of the system by virtue of their raw physical properties, in addition to their effect as (things which realise) an anticipation. Thus, the most general account requires retention of eqn. 6.

In a similar manner, we can acknowledge a distinction between the aspects of X which carry information about previous states of X, and those which do not. Since these "memory traces" may be incorrect or incomplete, we can denote them with a superscript notation similar to that used to denote anticipations, which are also potentially inaccurate or incomplete: $X^-(t-n)$. In order to collect together all memory traces into one variable (to avoid having an indefinite number of parameter places), we can stipulate:

$$X(t+1) = F(\bar{X}(t), X^*(t+1), X^-(t-1)), \tag{21}$$

where $X^-(t-n)$ is defined recursively by:

$$X^-(t-1) = J(X^{--}(t-1), X^-(t-2)), \tag{22}$$

where $X^{--}(t-1)$ represents the contribution at time t-1 to the total memory state X^-.

5 The Content of Anticipation

What is it that makes one state an anticipation of another? This difficult question is a special case of another difficult question: what makes one state a representation of another? It is a substantive matter, since we are interested in anticipations which can play a causal role in the functioning of a system. That is, the notion of anticipation in play is not entirely observer-relative; the intrinsic properties of the system place some constraint on what anticipations can be ascribed to it. A stone cannot become an anticipatory system in the sense employed in this paper just by someone deciding to see it as such (compare [Chrisley, 1994]).

There are four parts to the content of an anticipation that must be accounted for. What is it about a state that determines, or plays a part in determining:

1. that it is an anticipation at all (as opposed to, say, a memory)?
2. that it is an anticipation of the system/state/object it is an anticipation of (of, say, the state of a light, rather than of a clock)?
3. that it is an anticipation of the time that it is an anticipation of (as opposed to some other time)?
4. that it is an anticipation with the predicative content it has (an anticipation that, e.g, the light will be green rather than an anticipation that it will be red)?

Question one might be answered functionally. That is, a state is an anticipation if it is used as such: if the state guides behaviour in the way characteristic of anticipations, rather than, say, memories. If a state makes me put out my hand in a configuration appropriate for catching a ball, then, ceteris paribus, it is understandable as an anticipation that a ball will pass by me soon.

But anticipations can be directed toward particular objects: I can anticipate that *that particular ball* will move past me in two seconds. Answering question two seems to require more than an appeal to function; any conventional notion of function will not be able, on its own, to yield the directedness or intentionality of anticipations toward particular objects. A standard tactic in trying to explain particularity is to appeal to causal relations between the representing state and some object or state in order to single it out as the particular being represented. Such a tactic seems impossible here, since the state being represented is in the future. Many, even of those who are interested in anticipatory systems, are unwilling to talk of backwards causation, for familiar reasons ([Rosen, 1978a, p 157] cites [Windeknecht, 1967] to support this caution, but one can find more philosophical objections to backwards causation, such as the bilking argument in [Black, 1956]). But if we assume that states are composed of objects with various properties and in various relations to each other, then we can use the temporal extension of objects to provide us with causal relations that are not toward future events. My anticipation is that *that particular ball* will pass by me in two seconds because that ball in the future is the same ball which has already had causal impact on my retina and caused me to form the anticipation in the first place.

Nevertheless, there are cases where we have an anticipation involving an object in the future, with which we have had no causal interaction at the time of the anticipation. Thus, the solution just outlined is not available in such cases; what should we do? One suggestion is to abandon particularity for such anticipations. One can only have anticipations toward particular objects with which one has had some causal interaction (however remote or indirect). If no such interaction exists, then the form of one's anticipation is not that of being directed toward a particular object, but rather toward any object that happens to meet some descriptive condition. When I anticipate that there will be a car in front of me when I drive to work tomorrow, my anticipation is not about any particular car, but rather whichever car happens to occupy that place.

The temporal aspects of anticipation raised in question three also seem to be decomposable into a functional and a particular part. That an anticipation is directed toward some time, and specifically, some time in the future, falls out of the very notion of an anticipation. So it would seem to be a functional matter, involving the way that a state is used, that determines that it is directed toward some future time. It may even be a functional matter that the state is anticipating a time two minutes in the future from the time of the anticipation. But what makes the anticipation one of 12:02 on December 25th rather than 23:59 on December 31st is the fact that the state occurred at 12:00 on December 25th instead of 23:57 on December 31st. So the time of the state being anticipated will depend not only on the functional properties of the anticipating state, but the particular, absolute, non-functional temporal properties as well.

The fourth question is perhaps the most difficult. The philosophical community has only a few candidate theories of representational content, which attempt to explain the source of representation in terms of causal relations, evolution, asymmetric dependence, or conceptual role. There is no space here to go into detail; but a few quick points concerning the content of anticipatory states can be made by comparing causal and evolutionary accounts.

Causal theories of representation (e.g., [Dretske, 1981]) usually take the representation to be about whatever caused the representation to become active or come into existence. This is not a satisfactory account for the content of anticipations, since by their very nature they are about things in the future, which could not have been the cause of the anticipation, and may not even occur at all.

Evolutionary theories of representation (e.g., [Millikan, 1984]) require a representation to be something which has been copied or reproduced over time. Such theories then take the content of a representation to be whatever condition in the environment the ancestors of the representation covaried with, that also explains why the representation "survived" or was reproduced. To give a simplistic and implausible yet illustrative example, a particular state of an organism's nervous system will have the content "predator nearby" if that state has been copied from states which happened to be active when a predator was near the organism's ancestors, and that coincidence is part of the explanation why the state has been reproduced (presumably because it caused behaviour which allowed the organism to avoid the predator and reproduce, thus passing on the "predator nearby" state to its offspring.)

It might be thought that a backward-looking theory such as this would suffer from the problem just identified for causal approaches: anticipations are of things in the future, so how could an account of representational content which appeals to evolutionary history, which is even more remote, apply to anticipations? But evolutionary accounts are more subtle than that. For example, it could be part of the explanation for why the ancestors of a state survived and were reproduced that they occurred ten seconds before a predator crossed the plain to attack. That is, it could be a covariance with an event in the future which explains a representation's evolutionary value. In such a case, the content of the representation would be an anticipatory one.

However, this raises a general worry, not with just an evolutionary account, but with any account of anticipatory content. In the example just given, on what grounds do we say that that the content is “in thirty seconds, predator will be here” rather than “predator is now thirty seconds away”? That is, can’t any purported anticipation be re-interpreted as a representation that is non-anticipatory, instead indicating what conditions are like now? This may be a problem without substance; it may be that nothing of import depends on which of the two interpretations we give. But if it is substantive, then some philosophical work will have to be done in order to clarify this fundamental aspect of anticipation.

Any account of representational content has to be such that it is possible for representations to misrepresent, to be in error. For example, a causal account that identifies the content of a representation with whatever caused it fails to meet this criterion, since it is impossible on such an account for a representation to be about something other than what caused it. Evolutionary accounts fare better, in that it is possible for r to have the function of representing “predatory nearby”, as a result of r’s evolutionary history, and yet for r to misrepresent in virtue of the current facts (i.e., r is active, yet no predator is nearby). But the ability of evolutionary accounts to explain the possibility of misrepresentation has been questioned.

In any case, it seems that any theory of representation that can account for anticipation will most likely solve the misrepresentation problem anyway. Anticipations are directed toward non-actual states of affairs, those that are in the future and may or may not become actual. If one can explain how this fundamental problem of intentionality, dating back to Brentano and even the ancient Greeks, can be solved, then one will be able to use that solution to show how a representation can be directed toward a non-actual state of affairs in the present, which is just what we mean by a false or misrepresenting representation.

6 Conclusion

The equations given above allow us, in a precise manner, to broaden our notion of an anticipatory system to include some sophisticated forms of anticipation, including anticipation of times that are not a fixed distance into the future; and anticipation of a system’s output, environment and input. Furthermore, we can now begin to explore what constraints are placed on the dynamics of a system once one recognises that the anticipations a system has must be realised somehow in the causal structure of that system. Finally, some conceptual questions concerning the content of anticipatory representations which must be answered have been identified.

Acknowledgements My thanks to Aaron Sloman for inspirational discussions during the writing of this paper. This work was supported by a Research Fellowship funded by a grant from the Leverhulme Trust.

References

Black, M. (1956). Why cannot an effect precede its cause? *Analysis*, 16:49–58.

Chrisley, R. (1994). Why everything doesn't realize every computation. *Minds and Machines*, 4(4):403–420.

Cussins, A. (1990). The connectionist construction of concepts. In Boden, M., editor, *The Philosophy of Artificial Intelligence*, pages 368–440. Oxford University Press, Oxford.

Davidson, D. (1974). On the very idea of a conceptual scheme. *Proceedings and Addresses of the American Philosophical Association*, 47.

Dretske, F. (1981). *Knowledge and the Flow of Information*. Blackwell, Oxford.

Dubois, D. (2000). Review of incursive, hyperincursive and anticipatory systems: Foundation of anticipation in electromagnetism. In Dubois, D., editor, *Computing Anticipatory Systems: CASYS '99 - Third International Conference*, volume 517 of *AIP Conference Proceedings*, pages 3–30, Melville, NY. American Institute of Physics.

Millikan, R. (1984). *Language, Thought, and Other Biological Categories*. MIT Press, Cambridge.

Rosen, R. (1978a). On anticipatory systems: I. When can a system contain a predictive model of another? *Journal of Social & Biological Structures*, 1:155–162.

Rosen, R. (1978b). On anticipatory systems: II. The nature of the modelling relation between systems. *Journal of Social & Biological Structures*, 1:163–180.

Rosen, R. (1979). Anticipatory systems in retrospect & prospect. *General Systems: Yearbook of the Society for General Systems Research*, 24:11–23.

Rosen, R. (1985). *Anticipatory Systems*. Pergamon Press, New York.

Windeknecht, T. G. (1967). *Math. Syst. Theory*, 1:279.