Embodied Cognition as Interactive Cognition

R.A.Young

Philosophy Department
University of Dundee
r.a.young@dundee.ac.uk

Abstract

What abstract conditions concerning embodiment are necessary, together with formal conditions, for a specification of sufficient conditions for cognition? A material environment does more than realise one formal specification. There is no one base specification on which all the formal specifications that it realises supervene. Thus a relationship with a material environment may defy comprehensive formal analysis. Cognitive capacities are capacities to have relationships with material environments as such. Thus they are capacities which are not analysable in terms of any one formal model of the cogniser and its environment. An account of cognisers as autonomous is developed. The questions of whether a cogniser and its environment can be computational is explored.

I. Introduction

A question for this workshop is:

> Q1: What computational and representational structures are necessary and/or sufficient for enabling embodied cognition?

There is scope for argument as to what constitutes a computational or representational structure. It has been standard to regard computation as syntactic in a sense meaning formally specifiable and it has been standard amongst many philosophers to regard representational ***structure*** as formal. A formal specification is a finite specification of a set of elements S_e and a finite specification of a set of orderings of those elements S_o, where any set of elements that can be uniquely correlated with the elements of S_e, and which has the same orderings, constitutes a realisation of the formal system. Thus a finite specification of a language with a finite alphabet counts as formal so long as the coding of the alphabet is regarded as arbitrary. A formal system has a finite set of elements, such as an alphabet and a set of orderings of those elements, which is such that the difference between a valid ordering and an invalid ordering is finitely specifiable.

This is compatible with the orderings themselves not being finite orderings. Thus, if we define an analogue system as involving real numbers, some of which are not Turing-computable, we can still think of these real numbers as each uniquely correlating with an infinite ordering of characters from a finite alphabet and with a finitely specifiable grammar. Therefore we can think of analogue systems and chaotic systems as formally specifiable. These systems are capable of being instantiated in different material realisations, the set of material elements out of which they are constituted can be replaced, and, so long as the relationships between the replacement elements place them in the same orderings as the original elements, we have an instantiation of the same formal system. The respects in which there must be the same orderings, if a replacement system is to instantiate the same formal system as its original, are finitely specifiable in the grammar of the system. If computational and even representational ***structures*** are formal, and ***it is also held that to be embodied more is required than simply being the realisation of a formal structure***, then one component of Q1 is redundant viz. the sub-question:

> Q2: Are any computational and representational structures sufficient for enabling embodied cognition?

The argument is as follows: a set of sufficient conditions for embodied cognition must include conditions pertaining to embodiment, therefore no conditions pertaining purely to formal structures can be sufficient, hence no computational and representational structures can be sufficient for cognition. One can dismiss the argument by holding that computational and/or representational structures are not simply formally specifiable; or one can dismiss it by treating any realisation of a formal structure as an embodiment; but however one responds to the argument the following are crucial questions, in the debate about Q2:

> Q3 Is it possible to specify conditions of embodiment non-formally?

> Q4 If this is possible, then how is it to be done?

To the non-formalist it may seem that, if a non-formal specification of physical embodiment is sufficiently detailed, then it will itself specify sufficient conditions for cognition because computational and representational structures will be supervenient upon physical conditions. Thus, for sufficiency, formal computational and representational structures might appear irrelevant. It may be that this is Searle's position. If so, his dismissal of formal conditions is too swift, because a science of cognition must aspire to generalise and generalisation may only be possible by introducing formal specifications of computation and representation. Nevertheless we must at least recognise that material

conditions are pertinent to generalisation. Thus, if we are to generalise about sufficient conditions, then, even if we hold formal conditions to be crucial for generalisation, we must also address the following question:

> Q5: What abstract conditions concerning embodiment are necessary, together with formal conditions, for a specification of sufficient conditions for cognition?

Indeed it may only be by addressing this question that we can hope to make progress on the question of which computational and representational structures are necessary for cognition. The formal requirements for cognition may only be intelligible when the material conditions have been made intelligible. Thus even to give a formalist answer to the following sub-question of the original question Q1 viz.:

> Q6: Are any computational and/or representational structures necessary for cognition and, if so, what are they?

we may need first to address question Q5.

II. Three Kinds of Machine Specification

Question Q5 is puzzling: how are material conditions to be specified, moreover how are they to be specified ***abstractly*** (thus allowing generalisation) but not ***formally***? Consider three types of machine specification:

> A: a formal specification of a system with output determined by input,
>
> B: a material specification of a way of embodying a type A specification,
>
> C: a specification of a capacity that a machine is to possess that is given by reference to a ***particular*** environment.

Specifications of type A are formal specifications of abstract machines, including digital computers such as Turing Machines and digitised neural networks, but also including analogue machines, dynamical systems and chaotic systems. Specifications of type B specify input and output in terms of physically measurable states - instead of abstractly in terms of functional roles. They identify ways of embodying type A specifications. Yet they do not afford us any way of understanding how certain material conditions rather than others can be necessary for cognition. It appears arbitrary to say that one material realisation of a certain formal system (e.g. a brain as a realisation of a neural network) embodies cognition whereas another realisation (e.g. a silicon-based electronic realisation) does not. At least it is arbitrary unless we can explain either how the material condition that supports cognition realises some further formal specification or how it is that a material condition contributes to supporting cognition.

Specifications of type C ***refer to a particular environment***. Non-cognitive specifications are often of this kind. Thus NASA might specify that a surface vehicle is to be capable of manoeuvring in some region of the surface of Mars. At first sight such specifications may appear to have nothing to do with cognition. Surely cognitive capacities must be general and therefore they must be capacities to relate to any instance of a type of environment and not simply to a particular environment? The answer to this question may be affirmative, but it does not follow that it is possible to give specifications of cognitive capacities except by reference, implicit or explicit, to a particular environment.

Type B specifications may be construed as akin to Type C specifications insofar as a material is specified by reference to samples in some particular environment (see the discussion of water in the next section). Even so Type B is different from Type C, because Type C specifications are concerned with a relational capacity to interact with an environment. Thus Type C specifications may be understood as specifying a capacity to interact with a material environment, where the kind of material environment in question is specified by reference to a particular instance, or instances, of it. The material in which a formally specified entity is realised is an arbitrary matter, therefore if cognition could be specified formally, then variations in the material embodiment of an entity would be irrelevant to its cognitive status. However if cognition is properly specified by a Type C specification, then it may necessarily the case that, if one is to have a capacity to relate in certain ways to a material environment, then one must have a certain kind of material embodiment.

III. Type C specifications of Cognition

Consider, as an example of a type C specification, Turing's specification of thinker-equivalence in terms of an imitation game played using a ***human*** comparator, whom the machine is to imitate, and a ***human*** interlocutor who is to try to discriminate between the comparator and his mechanical imitator (Turing 1950). This specification appears general, but implicitly refers to a particular environment, because the term "human" has its extension fixed by reference to members of our family of organisms on this planet. Our understanding of the term "human" even when we say, "There might have been humans, even if all of the humans who have actually existed had never existed.", is dependent upon reference to our particular world of existing humans.

Type C specifications may, unlike Turing's, involve reference to an environment which does not consist of intentional agents. Consider the specification "has the capacity to discriminate water from other substances and so apply the word 'water' to water". With Putnam, we can think of even a word such as "water" as having its intension over possible worlds determined by how its reference is fixed to particular masses of material in the actual environment of the language user, rather than being entirely determined by the language-user's theory of those masses (Putnam, 1975). If the masses of material consist of H_2O, then water is necessarily H_2O. If the material consists of XYZ, then water is necessarily XYZ. Even if water is H_2O, and understood to be such, that is not the end of the under-determination of the meaning of "water" by theory, not if the properties of the constituents of Hydrogen and Oxygen atoms are not themselves fully expressed in the theory. Consider a cogniser that itself has the ability, as opposed to the latent capacity, to refer to water. This, just as much as the ability of a robot to manoeuvre around the surface of Mars, is an ability which is not formally specifiable.

Thus on this view cognisers are agents situated in a material world. The material world is such that it is neither entirely explicated in the language and theory of those cognisers themselves nor even in the language and theory of any further cognisers who may attempt to interpret them or specify their capacities. ***Cognition itself is to be explicated in terms of relations to the material world, relations which a***

cogniser can never fully articulate in a formal specification of cogniser and world. On this view not only is the material world not comprehensively specifiable as an instantiation of a formal system, but also it is not possible fully to analyse all the relevant aspects of the cogniser's relationship with its world, by identifying a formal system of relations between the cogniser and its world. Even in the case where a cogniser has a formal system as object of cognition, as when a logician works on a formal system, cognitive relationships are more than merely formal, because the cogniser must be able to recognise material realisations of that formal system, and distinguish between material elements and the abstract syntactic role that they realise.

IV. Autonomous Intentional Agents and the External World

Turing's attempt to characterise thinker-equivalence was cited as an example of a type C specification of a cogniser's capacities. This seems insufficiently abstract, because it is too anthropomorphic. Surely there could be alien cognisers, whose alien cognitive capacities would not allow them to mimic humans well enough to pass the Turing Test. There are however other specifications of cognisers in the literature which, like Turing's, involve reference to an environment of cognisers and yet which are not so anthropomorphic.

It may be argued that Dennett's account of cognisers in terms of the intentional stance is one such account (Dennett 1990). Arguably Wittgenstein's account of rule-following provides us with another such account, or perhaps a family of such accounts, when we consider the elaborations upon it that have been developed by Wittgenstein's followers. On Dennett's account a True Believer is an entity which is best predicted from the intentional stance and the best interpretation of an individual entity from the intentional stance is the interpretation which is best for predicting that individual's behaviour. However, like Turing, Dennett presupposes the existence of beings who can engage in the intentional stance and in other stances. If his specification of what it is to be a True Believer is to be correct and ***non-trivializable*** then it must be impossible to give a comprehensive formal account of the relationship between a True Believer, its environment and the being in its environment who takes the intentional stance towards it. Dennett's account is more abstract than Turing's, because it does not require that those who take the intentional stance must themselves be human beings. Nevertheless Dennett's account is only intelligible because we can refer to intentional beings in our particular environment. His specification of "True Believer" is implicitly a type C specification.

In applying it we can begin by presupposing that some set of humans in our particular environment is a set of cognisers. We then treat non-humans whom those cognisers can best predict from the intentional stance as further cognisers. Then by extension we can accept yet further beings who are best predicted by those non-humans as cognisers and so on. This may lead to the recognition of cognisers whose behaviour is not intelligible to the original humans. We can even conceive of possible worlds in which there are no humans, but which are inhabited by such alien cognisers.

But is there a principled argument to the conclusion that it is not possible to give a comprehensive formal account of the relationship between a True Believer, its environment and beings in its environment who take the intentional stance towards it? After all an interpretation from the intentional stance can be presented in the form of an hypothesis that the true believer has certain premises, certain desires, and a capacity to change premises in accordance with changes in the environment. This hypothesis seems to be susceptible to formalization. Thus we could envisage a formal model of interacting agents with a "True Believer", an environment and a set of interpreters, whose "interpretations" themselves were formally expressed. Could an arbitrary implementation of such a formal model, digital or analogue, itself realise a world of genuine cognisers? If not why not?

Answers to these questions can be developed from an idea, which I have already canvassed in this paper. Cognitive relations are to be understood as relations which are established in a material world and are not merely formal. Cognitive relations are relations to the material world, not just relations to some formal aspect of it. No formal specification of that world comprehensively captures all of its properties. We can think of cognitive representation as representation which is not analysable in terms of a comprehensive formal specification of world and cogniser, because the representational system is sufficiently flexible to allow it to vary as it interacts with different aspects of its world, and the world is not fully specifiable as a realisation of any one formal system.

According to this account a cogniser is an autonomous intentional agent where such an agent is defined as one that is capable of categorising its environment in ways which differ from any other agents who interact with it (see Young 1994 for a fuller definition). An interpreter of such an agent will need to bear in mind that it may have its own categories and its own semantics. If tokens of a type of agent are each to count as autonomous intentional agents then they will have to be capable of diverging from each other in their categorisations and if this divergence is not to be purely arbitrary, then it will reflect variations in their relation to their environment, variations which have enabled each of them to learn different categories. If one of these agents recognises others as autonomous intentional agents, then it will ipso facto recognise its environment as one for which it may learn novel categorisations beyond any formal analyses of its world which it has yet developed.

V. Cognition and Computation

It may seem at first sight as if this account of cognition is antagonistic to a computational account of mental processes. How can a computationally finite cogniser be capable of indefinite flexibility in its categorisations of its world? Consider a robot that can learn to vary its categories by varying its system of internal representation. If the robot's representational capacity is understood to be limited such that the number of distinct representational states can be no more than the number of its distinct syntactic or internal functional states, then it cannot show indefinite flexibility in how it categorises. The fundamental answer here is that the same internal syntactic state may represent different objects in different environments. But this is an insufficient answer. Even if we take a very inflexible robotic system, say an industrial robot used for welding, without the ability to learn new categories, then it might be said to be capable of representing (and welding) an infinite variety of different objects. Not in the sense that it could distinguish between each of these different objects, but simply in the sense that it could interact with an infinite variety of them. Its procedures for representing (and welding) might be completely inflexible. Thus the robot would not show an indefinite flexibility in representation: it would have no learning capacity at all.

A finite cogniser, whose internal workings can be explained computationally may nevertheless be capable of indefinite variety in categorisation. One way in which this can occur is if the cogniser uses external storage, as we do when we use writing systems, electronic information systems and so forth. Just as the finite state machine embedded in a Universal Turing Machine can constitute, together with its tape, a system which is capable of indefinitely large computations, so a cogniser, with a finite brain, together with external systems of storage, can be capable of indefinitely various ways of categorising its environment. If we put the matter like this, then it may appear as if only a very sophisticated entity, with the capacity to learn to write and use external methods of calculation, can fulfil the conditions herein specified for an autonomous intentional agent. Thus less sophisticated organisms than humans would appear to be excluded from the ranks of cognisers. However this is a misunderstanding. Less sophisticated organisms may have their own ways of storing information in an environment, for example by chemical marking. Leaving that aside, if they have a capacity for co-operation with other cognisers, then their methods of categorisation may be constituted out of co-operative processes, and not just out of their individual internal processes. In the human case a philosopher like Putnam may, as he remarks, not be able to tell the difference between elms and beeches, but he may still recognise a distinction between them, because other humans, more expert with trees, can make a distinction. On the other hand Putnam has his own skills and abilities to contribute to the human system of categorisation. This is to put the point at a very intellectual level, but correspondingly a truffle hunter may rely on a dog to make distinctions of smell that he could not make himself, and the dog may rely on the truffle hunter for other distinctions.

The conception of categorisation being embedded in a society of cognisers may help to explain its flexibility and variety, but it may appear to be in conflict with the very conception of autonomy that was introduced above. In the previous section the following claim was made:

> a cogniser is an autonomous intentional agent where such an agent is defined as one that is capable of categorising its environment in ways which differ from any other agents who interact with it.

Fortunately this conflict is only apparent. A being may be capable of combining its own internal processes with those of surrounding cognisers in a novel way even though it relies on other cognisers to apply its categories. For example a scientist, who may be very dependent upon the expertise of other scientists, may nonetheless introduce a novel change in a theory or in the practice of applying it.

An account of a cogniser's representational states may be based in part on Dretske's paper, `Misrepresentation', in which he is concerned to understand, in causal terms, how an inner state may be a natural sign of an external property and not simply a natural sign of an intervening state of the sensory organs (Dretske 1990). He takes the example of an associative learner and argues that, as its environment varies, different sensory cues may be associated with the presence of a property (e.g. the property of toxicity). Thus he argues that the internal symbol for the property of toxicity need not correlate with any one sensory state, nor even with a disjunction of such states. The symbol for toxicity will only correlate with what is external not with what is internal. Note that in a sufficiently sophisticated learning system the response to toxicity may also vary and so may the internal symbol used for toxicity. In Dretske's example representation depends upon the causal regularities on which the system will stabilise through learning in each of a potential infinity of different environments. Because of this, because of the points made above about externally based ways of categorising, and because any given internal symbol may come to correlate with different external states, a computationally finite cogniser may be capable of indefinitely varying representation.

In his paper Dretske attempts to deal with misrepresentation and does this by invoking the concept of a state having the function of representing an object. Since Dretske sets out to give an account of misrepresentation, the term "function" here cannot simply be understood in terms of causal role or correlation. If "function" in this sense in terms of evolutionary history or even individual history of learning, then the view is problematic, because a Doppelgänger of a cogniser lacking the background history may, in the right environment, be a cogniser. To avoid this problem Dretske's point about the contingency of representation on the environment may be incorporated in Dennett's account of the intentional stance. Instead of identifying the function that a state has on the evolutionary, or learning, history of the organism, we can identify the function of the state in terms of the role it has in the best interpretation of the organism according to the intentional stance. This is the most predictive interpretation according to Dennett. Thus an account of misrepresentation that can allow for Doppelgängers may be developed.

VI. Externalism, Semantics and Algorithms

A basic intuition of externalist semantics is that description of internal procedures is inadequate for specifying the truth conditions and conditions of referential success and failure of states with content. A corollary is that there can be no internal algorithms that guarantee correctness of representation. A given content presupposes the background of a certain environment. The success of internal "algorithms" presupposes that background and thus the conditions of referential success are presupposed, rather than being themselves computed. A basic question that arises for any externalist is: "How can beings have a content when they do not have any internal procedure that guarantees referential success or truth?" The answer is that procedures may enable reference and truth even though they do not guarantee it and that conditions of reference and truth are determined by facts about how those procedures enable interaction with a cogniser's environment. In the XYZ world my twin interacts with XYZ in just the same way as I interact with H_2O. The presence of the different substances in the two worlds is what makes the difference to reference and ultimately to truth conditions. This answer gives rise to a problem: if my twin exactly parallels me in all functional respects, and has not discerned any property of XYZ that is different from H_2O, then why should we say that my twin has different content from me? Might it not be that the referential extension of the "water" symbol includes both H_2O and XYZ? The problem is that if this is admitted then externalism is undermined: the extension of the water symbol will include anything with which a twin of mine could interact in the way in which I have so far interacted with H_2O

The externalist can respond to this challenge in different ways. One form of response is to say that my twin and I have different interactions with the world, and therefore different contents, even though the difference between XYZ and H_2O makes no functional difference between us whatsoever and never could make any such difference. On this view it is just a brute fact that my twin interacts with XYZ and I interact with H_2O - a brute fact that is postulated in the metaphysics of the semanticist, but which makes no

difference that myself or my twin could discover. But there is another kind of externalism which does allow my twin and myself to learn about its basis. On this second view the difference in content between my twin and myself does depend upon the capacity of my twin and myself to learn about properties of H_2O which are not properties of XYZ and/or properties of XYZ which are not properties of H_2O. On this latter view the conditions of reference and truth that I now have are not solely determined by my current internal procedures for applying the "water" symbol, but they are also determined by my capacity to learn other more sophisticated procedures in the future. I use the "water" symbol as a term for a substance about which I am learning. It is a term which I treat as having the same conditions of reference and truth even though I change my procedures for applying it. Thus a child can begin with procedures which enable the term "water" to be applied to anything which is perceived as a colourless sloppy fluid with a neutral taste. But the child accepts that these procedures may be defective and is willing to learn better procedures for applying the term from the adults around it. It will find itself applying the term to steam and to snow and ice. Adults themselves are willing to learn better procedures as they interact with water in the environment, thus they can come to distinguish pure water from solutions of other substances in water.

Thus both children and adults are capable of learning to change their current procedures for applying the term "water". As they learn new procedures they not only correct their misdescriptions of the world around them but they also learn to correct their misperceptions. The child comes to see ice, steam and clouds as water and to see colourless fluids as such and not as water. The child learns that cues which were once used to distinguish water from non-water have only a transitory and ephemeral role to play in the learning process about water. Note that this means that the child will, in a certain sense' have different concepts of water as it learns. The sense is one in which the possession conditions of concepts are defined in procedural terms, and thus on this view concepts can change while referential content remains the same. Throughout the learning process it is assumed that the referential extension and truth conditions of the term remain the same. The extension and truth conditions are ones that are discriminable using the learning procedure in the long run. They are discriminable by the learning procedure given the particular environment and on the basis of initial procedures for applying the term. Myself and my twin have different content for the water symbol to the extent that there are discriminable properties of the structure and effects of H_2O which are not discriminable properties of XYZ. This difference in content does not arise only when these properties are recognised. We differ in content as soon as we begin to apply the term "water" because we are applying our learning procedure in different environments, and thus what we are learning about differs.

Some externalists, e.g. Peacocke, propose that we should think in terms of externalist content-involving algorithms (Peacocke, 1994). Consider Peacocke's account of algorithms in the light of the second version of externalism. If we think of initial procedures and the learning procedure as algorithms, then do we want to say that my twin in the XYZ world has different initial algorithms and a different learning algorithm from myself merely because it has XYZ in its world whereas I have H_2O in mine? If we did say this then it would undermine the conception of reference and truth conditions being determined by what is ultimately learnable in an environment given initial procedures for applying a term and the learning procedure itself. On Peacocke's conception we would not be able to explain the difference in content between my twin and myself in terms of different outputs in the long run from the same learning procedure given the same initial internal procedures, instead my twin and myself would have algorithms that were specific to our environments from the very beginning.

On the syntactic or non-semantic view of algorithms we can recognise that my twin and myself have the same algorithms, but nevertheless recognise that these algorithmic processes realise different content-involving states. Difference in referential content occurs where there is difference in external relations that is manifestable as a difference in syntactic outcome in the long term. It is tempting to think that if a content-involving algorithm were a general learning algorithm then an explanation in terms of it could have sufficient generality to explain the difference in content between the two worlds. The temptation is to think that there could be a general content-involving algorithm which explained the difference in content as a computed outcome arising from different non-semantic inputs to the algorithm in different worlds. But according to the externalist view, which Peacocke accepts, my twin and I differ in content from when we first apply the "water" symbol, whilst the input that we have (he from XYZ and I from H_2O) has exactly the same internal non-semantic properties. Thus a content-involving algorithm that specified how difference in non-semantic input generated difference in content could not explain how my twin and I have different contents from the beginning. On the syntactic view of algorithms we can recognise that the same algorithm is being applied by my twin and myself even though my twin and I have different content. We can do this because we keep content and algorithms distinct. Thus an externalist has reason to keep content and algorithms distinct.

The psychologist also has a reason for keeping algorithms and externalist content distinct. If the psychologist does this then it is possible for the psychologist to construct a science that generalises over situations where the externalist content varies, but inner psychological processes remain the same. Consider myself and my twin before we make any divergent discoveries about H_2O and XYZ. We may have different contents when we use the "water" symbol, but there are powerful reasons for saying that our psychology is the same. Ex hypothesi our thought processes are formally identical, they use exactly the same internal resources of short and long term memory, and they produce the same bodily behaviour. At first sight the way for the psychologist to generalise over these situations is to avoid content-involving descriptions and to concentrate on syntactic, that is formal, explanations of our thought processes.

However this is too simple a position for the psychologist to take. If the externalist is correct about content, then the psychologist must aspire to a psychological explanation of how it is possible for us to sustain content of that kind. It has been argued in the above discussion of externalist semantics that content can be sustained by us, because of our learning capacities. According to the externalist our procedures for applying a symbol will be inadequate when we begin to apply that symbol, but when we introduce a symbol, we presume that our learning procedures, as they are applied in our environment, will be capable of developing our initial procedures to make them progressively more adequate. Throughout such a process of learning a symbol can have as its extension a set of objects about which we are learning, a set which is ill-determined by our initial procedures taken by themselves, but which is identifiable as the set for which our learning algorithm or algorithms can develop more adequate procedures. These further procedures generated by learning may or may not coincide in each one of us. It is possible, according to externalism, for there to be intellectual division of labour and therefore for there to be no one individual who embodies all the procedures which now constitute our

communal basis for deciding truth claims about water. Moreover it is possible that it is beyond the intellectual capacity of any one of us to embody all the relevant procedures . Thus we need not think of the outcome of learning as a psychological outcome, in the sense of an individual outcome, so much as a sociological, that is communal outcome. Nevertheless the psychologist has a legitimate objective in seeking to understand how such outcomes can be sustained by psychological processes embedded in social and environmental interaction.

If the psychologist is to aspire to explain how it is possible for us to sustain content, then the psychologist needs an account of learning algorithms, but, on the syntactic account of what an algorithm is, these algorithms will be functions from syntactic patterns to syntactic patterns. This is indeed how neural network algorithms are specified as functions specifying the activation of neurones for given inputs and functions for adjusting the weighting of inputs. There is also the configuration of a network to consider, but that too can be understood as an implementation of a set of functions. In themselves syntactically specified learning algorithms will not wholly explain how it is that we are capable of learning to discriminate sets of objects in an environment. In order to understand that, we need to understand how neural network algorithms (or other learning algorithms) can, given the appropriate embodiment and environment, enable an organism to develop capacities to discriminate objects, a task which cannot be altogether severed from understanding how it is possible for algorithms to control movement of the organism in an environment. Understanding how symbols can more or less reliably be applied in an environment to meet conditions of reference and truth involves more than simply an understanding of syntactic algorithms, it involves an understanding of sensory systems and systems of locomotion, of how the objects in the environment impinge on sensory systems, and of how objects serve to help or hinder locomotion. Syntactic algorithms and a theory of computational learning will have their part to play in all this. For example the theory of computational learning can be used to analyse the amount of memory that is required to solve the problems of syntactic pattern recognition that are necessary in order to respond with appropriate discriminations to sensory input. Understanding the amount of memory that needs to be used in order to solve a problem is at the heart of the computational aspect of psychological theorizing. In order to have a measure of the amount of memory being used the psychologist needs to be able to think at a syntactic level. At the level of content we do not really have such a measure. The conception of the minimum number of bits that it takes to code some information is essentially a syntactic notion.

Thus the psychologist has reason to think in terms of syntactically specified algorithms rather than in terms of Peacocke's "content-involving" algorithms. These reasons are (1) that syntactically specified content allows the psychologist to develop a general science by identifying the same psychological processes in cases where content differs (as in the case of myself with H_2O content and my twin with XYZ content) and (2) that a quantitative measure of memory needs to be syntactic and the analysis of memory utilisation in syntactic terms lies at the heart of our reasons for regarding psychological processes as computational.

VII. Embodiment and Computational Environments

Consider an attempt to implement a cogniser as a process in a computer where the designer of this process seeks to interpret the "cognising" process as having an environment which itself consists of processes within the computer. Is such a project inconsistent with the view expressed in the present paper? The view is that a cogniser is an autonomous intentional agent that is capable of diverging in its categorisation of its environment from any other cogniser, including even its own designer and the designer of its environment. The environment itself is supposed to be a material environment which ipso facto cannot be comprehensively modelled using any one formal model. If the designer's project is to realise the relations between a cogniser and its environment in processes which can be comprehensively formally specified, then it is indeed inconsistent with the view herein expressed.

The computational model with which the designer has worked is just one way of interpreting the material processes which, subject to error and breakdown, realise that model. Thus the agent is realised in certain material processes and so is its environment. Are we to envisage the agent embodied in this material process as having the capacity to develop a way of categorising its environment according to a different model from that of its designer? According to the view herein expressed it is only if the agent has the capacity to develop autonomous categories for the material realisation of its computational environment, and/or the larger environment that extends beyond its computational environment, that it counts as an autonomous intentional agent or as a cogniser.

VIII. Conclusion

In this paper I have attempted to argue that capacity to interact with the material environment is what is fundamentally important in embodiment of cognition. On the view put forward the internal workings of an embodied cogniser may realise a formal specification, but representation is an interactive capacity and it depends upon the material capacity that a cogniser has to interact with its material environment Thus my answer to question Q5 above, namely the question, "What abstract conditions concerning embodiment are necessary, together with formal conditions, for a specification of sufficient conditions for cognition?", is that a cogniser must constitute an autonomous intentional agent with respect to a material environment which is specified by reference to a particular environment rather than as a realisation of a formal system.

References

Dennett, D. 1990 `True Believers', pp.150-166 in Lycan,ed., *Mind and Cognition*, Blackwell,

Dretske,F, 1990, `Misrepresentation', pp. 129-143 in Lycan,ed., *Mind and Cognition*, Blackwell,

Peacocke,C.,1994, `Content, Computation and Externalism', 303-335, *Mind and Language*, Vol. 9, No. 3

Putnam, H.,1975, `The Meaning of Meaning', *Philosophical Papers: Mind, Language and Reality*, Cambridge, University Press,

Turing, 1950, `Computing Machinery and Intelligence', *Mind*, LIX, no 236,

Young, R.A., 1994, `The Mentality of Robots', pp. 199-227, *Proceedings of the Aristotelian Society*, Supplementary Volume LXVIII