ON THE COGNITIVE ARCHITECTURE OF INSECTS AND OTHER INFORMATION-PROCESSING SYSTEMS*

Francisco Calvo Garzón
Departamento de Filosofía. Campus de Espinardo. Universidad de Murcia
fjcalvo@um.es

Abstract

According to Carruthers (2004) ants and bees have minds. This claim is to be understood realistically. We do not interpret the overt behaviour of ants and bees by ascribing to them beliefs and desires in an instrumental manner. They rather possess minds in the relevant cognitive sense. In this paper, I propose to pave the way for a *reductio* against such a polemic view. In particular, I shall argue that if ants and bees have minds, by the same token, plants do have minds too. In my view, the problem has to do with Carruthers' (2002) underlying technical concept of *cognitive architecture*; a concept which, as I shall argue, can be called into question both on empirical and conceptual grounds.

KEY WORDS: Cognitive architecture; Connectionism; Plant neurobiology

Resumen

Según Carruthers (2004), las hormigas y las abejas tienen mente. Esta afirmación debe entenderse de modo realista. No es que interpretemos la conducta abierta de hormigas y abejas en términos de creencias y deseos atribuidos instrumentalmente. Se trata más bien de que tienen tales estados mentales en el sentido cognitivo relevante. En este trabajo, me propongo llevar esta polémica concepción a una reductio. En particular, argumentaré que si las hormigas y las abejas tienen mente, por la misma razón las plantas también la tendrían. A mi modo de ver, el problema tiene que ver con el concepto técnico de *arquitectura cognitiva* de Carruthers (2002); un concepto que puede ser cuestionado sobre bases empíricas y conceptuales, según argüiré.

PALABRAS CLAVE: Arquitectura cognitiva; Conexionismo; Neurobiología vegetal

---

* I would like to thank Peter Carruthers and Toni Gomila for their helpful comments and suggestions. A previous version of this paper was presented at the *XVII Interuniversity Workshop on Philosophy and Cognitive Science*, held in Palma de Mallorca (Spain) in May of 2007. I wish to thank the audience at this meeting as well. Preparation of this manuscript was supported by DGICYT Projects BFF2003-129/BFF2003-08335-C03-02/HUM2006-11603-C02-01 (Spanish Ministry of Science and Education and Feder Funds).

1. Introduction

According to Carruthers (2004), the conditions for an information-processing system to qualify as minded allow in for a wider range of systems than most people would be happy to admit at first sight. This claim is not put forward with an eye towards classic experiments in cognitive ethology or comparative psychology in relation to, say, great apes' manifest competencies. Rather, much "lower-level" animals such as ants and bees, however radical this position may sound, are meant to have minds in the relevant cognitive sense.

The link between alleged insect cognition and cognitive architecture is pretty straightforward.[1] Carruthers (2004) asks what is required for an information-processing system in order to have a mind. The answer is the possession of a "certain core cognitive architecture" (p. 207). Put bluntly, by a "certain core cognitive architecture" Carruthers means a belief/desire architecture. He claims,

> Having a mind means being a subject of perceptual states, where those states are used to inform a set of belief states which guide behavior, and where the belief states in turn interact with a set of desire states in ways that depend upon their contents, to select from amongst an array of action schemata so as to determine the form of the behavior. (p. 207)

For the foes of Carruthers' polemic thesis, one option might be to tackle it head-on by appealing, for instance, to the privileged status of language-using animals. This may be the end of the story, were we to interpret the possession of a belief/desire architecture exclusively in propositional attitude terms, as far as the alleged cognitive capacities of non-human animals go. Only humans, the argument would run, can realize (proposition-based) cognitive architectures.

Carruthers (2004), however, has in mind a more embracing picture. As he points out, the error lies in the "assumption that thought contents *must* be specifiable by means of that-clauses" (p. 206). The critical condition

---

[1] A word of caution is needed before we proceed. We may distinguish between an extended *mindedness* thesis, as non-human animals are ascribed a mind in Carruthers' philosophical sense (see below), and a thesis of extended *cognition*, insofar as particular cognitive abilities such as, say, attention, perception, learning or memory are the empirical focus of interest in the animal cognition literature (Bekoff et al., 2002). For present purposes I shall speak interchangeably about the possession of a mind and the possession of cognitive abilities, bearing in mind that I mean to address the former thesis.

is not that the format of inner states can be realistically equated with propositional states, but rather that the states in question, whatever form they have, are intrinsically contentful. Non-human animals have states about external objects, and insofar as those states stand in for them, a belief/desire architecture can be said to apply to them: "[We] don't know how much the ape knows about termites, nor how exactly she conceptualizes them, but we do know that she believes of the termites in that mound that they are there, and we know that she wants to eat them" (p. 207).

In this way, the issue of the intrinsically contentful character of the system's states in question boils down to a question about the complexity of the internal economy of the system itself.[2] Mediating states between perceptual input and motor output are arranged so as to exclude a reactive (non-cognitive) interpretation of the behavioural repertoire of the system. What really matters is that it is "unlikely that possession of perceptual states alone, where those states are used to guide a suite of innate behavioral programs or fixed action schemata, could be sufficient for a creature to count as possessing a mind." (p. 207). In this way, structural mediating complexity, rather than the positing of propositional attitude states, is what can be exploited as a demarcation between reactive and cognitive overt behaviour.

For present purposes, I am pretty happy to bite the bullet for the sake of argumentation. In principle I am sympathetic to any principle of demarcation that is not cast in anthropocentric terms. Plausibly, evolutionarily speaking, minds emerged as a result of the structural complexity that obtains as an increasing number of subsystems play a partial causal role in the integration of information. And although adaptive responses to contingencies can certainly grow in sophistication, I see no reason (beyond multiple realizability arguments, which I shall ignore in this treatment, although see Bickle, 2003, and Churchland, 2005) why accounting for complexity necessarily requires human-based conceptualizations. That is, I am happy to endorse a stance without anthropocentric preconceptions, and consider any sort of information-processing system, natural or artificial, in order to assess whether it possesses the kind of belief/desire architecture that Carruthers demands.

In what follows, I propose to pave the way for a *reductio* against Carruthers' polemic thesis. In particular, I shall argue that if ants and

---

[2] It is far from clear what conditions an information-processing system must satisfy before it can qualify as representational (for two recent approaches, see Ramsey, 2007, and Calvo Garzón, in print), but we may grant Carruthers' setting for arguments' sake (although see sections 4 and 5 below).

bees have minds, by the same token, plants do have minds too. In my view, the problem has to do with Carruthers' (2002) underlying technical concept of *cognitive architecture*; a concept which, as I shall argue, can be called into question both on empirical and conceptual grounds.

2. Towards a principle for mindedness

Once we adopt Carruthers' neutral stance we need to spell out in some detail the core features of the belief/desire architecture in order to assess the alleged mentality of information-processing candidates. In this section, I shall first introduce Carruther's proposed principle for mindedness, in order to consider subsequently, as an alternative candidate to flesh up the belief/desire architecture, a connectionist-inspired principle for mindedness.

2.1 Carruthers' principle for mindedness

Carruthers exploits a "condition on mindedness" according to which:

> **Carruthers' Principle for Mindedness (CPM)** *To be minded means … having distinct belief states and desire states that are discrete, structured, and causally efficacious in virtue of their structural properties*. (p. 208)

CPM is clearly based upon the sort of features that a Fodorian (good-old fashion AI - GOFAI) computational picture of the mind provides. The underlying idea is that thinking can be seen as logic-like inferential processing. In this picture, context-independence is a key feature. Mental representations are formed out of context-independent constituents in such a way that constituents appear in different thoughts as syntactically identical tokens with the same content. I shall refer to this kind of context-independence, as classical constituency (Calvo Garzón, 2000). Carruthers is, as a matter of fact, taking for granted a concept of classical constituency insofar as the belief/desire architecture demands a Fodorian view of structural compositionality of the sort championed by Fodor and Pylyhsyn (1988).

The question I want to raise, put bluntly, is: Why should classical constituency be *the* condition on mindedness? Connectionism and dynamicism are probably the most well-known alternatives in the philosophical community, but there are a number of other hypotheses in the market as far as the architecture of cognition is concerned that range from hybrid architectures to full-fledged embodied and embedded models (for an overview see Gomila and Calvo, forthcoming). However, for purposes of illustration, and assuming that the general reader will be more familiar with connectionist theory, in what follows I shall consider connectionism as *the* alternative, bearing in mind that parallel arguments to the ones I shall be offering can easily be recast, for example, from a dynamicist perspective (Port and van Gelder, 1995).

2.2 A distributed and superposed principle for mindedness

According to connectionist theory, cognitive activity does not consist of formal operations performed on internal representations according to syntactic rules. Instead, the cognitive function is approximated by adjusting the synaptic weights that configure the skeleton of the neural network under the light of the statistical (sub)regularities that the network is recurrently fed with. The result is the failure to endorse the 'computational theory of the mind', as classically understood.[3] Neural networks represent states internally as fully-distributed set of values, such that individual hidden units defy semantic interpretation. In addition, since learning boils down to the adjustment of a single set of weights in order to produce the appropriate input/output correlation for the training set, knowledge is said to be stored superpositionally in the weight matrix as a whole. In this way, individual units and connection weights embody subtler information than the one being represented and processed in terms of classical constituents.

Connectionist networks, insofar as they employ fully-distributed representations and superpositional storage techniques, are incompatible with CPM above. In particular, they are incompatible with the idea that logically independent belief/desire states have distinct causal roles.[4] Since information is stored in a fully-distributed superpositional fashion, it makes no sense to talk of the distinct causal efficacious role played by the representation of a particular state. Notice that each hidden unit and each weight encode information about every single informational state that has been presented at the input level. The content of each different state is determined by the superposition of all the available representational

---

[3] For a philosophy-oriented introduction to connectionism, see Bechtel and Abrahamsen (1991).

[4] For a defence of the elimination of the posits of folk psychology that exploits the connectionist features of distribution and superposition, see Ramsey, Stich and Garon (1990). For a recent development of this line of argument, see Ramsey (2007).

resources; resources that are a function of the whole range of input/output patterns that the network has been trained on. Put bluntly, there is no discrete causal efficacy in connectionist networks and there is no canonical representation to be singled out which is common to all instantiations. Instead, there are many different representations encoding for each different context. Constituency can be kept context-dependent by encoding the precise location of each individual pattern of activation. Connectionism, in this way, differs in its way of representing constituency. Whereas in the classical symbolic approach, constituency is context-independent, connectionist constituency is said to be context-dependent (Calvo Garzón, 2000).

With this toolkit, I can propose now the following alternative to CPM:

> **Distributed and Superposed Principle for Mindedness (DSPM)** *To be minded means having distinct belief states and desire states that are fully-distributed, and causally efficacious in virtue of their superposed statistical properties*.

I am aware that this rendering of the dispute between classicists and connectionists has been contested from both the classicist and the connectionist fronts (e.g., Fodor, 1998; Clark, 1995). Moreover, there's nowadays a heated debate as to the sort of entity which can be the vehicle of content in a neural network (see Shea, 2007). Nevertheless, we can ignore this set of issues for present purposes. All I need in order to pave the way for the *reductio* is to bear in mind the cleavage between rule-governed manipulation of symbols of the sort assumed by CPM, on the one hand, and fully-superposed (that is, fully distributed and superposed) manipulation of subsymbols of the sort put forward by the friend of connectionism, on the other. In the following section I shall offer empirical and conceptual considerations in order to show that statistically-driven subsymbolic computation is on a par with the GOFAI framework. The first question then is: Can the Distributed and Superposed Principle for Mindedness (DSPM) be used in order to assess Carruthers' thesis?

3. The architecture of cognition: Empirical and conceptual challenges

Carruthers is of course well aware of the threat that the application of DSPM, instead of CPM, would represent. In his view, however, connectionism is not a "viable [candidate] as an overall model of cognition"

(2002, p. 714).[5] But what are the reasons for rejecting connectionist theory? Carruthers contends that connectionist models have had a very limited impact and that their success reduces to partial explanations in specific areas of research such as pattern recognition (Rumelhart, McClelland et al., 1986). By contrast, with respect to classical cognitive science, he claims that "many of us believe that this form of psychology represents easily our best hope (perhaps our *only* hope) for understanding how mental processes can be realized in physical ones" (p. 705).

Unfortunately, being a very good candidate does not mean being the one and only candidate. It seems to me that the burden of proof is on the side of the defender of CPM to spell out in detail the reasons why non-classical models have a different status. In an attempt to clarify his main thesis in relation to the cognitive functions of language, Carruthers (2002) contends that:

> i) there are principled reasons for thinking that [connectionist] models cannot explain the kinds of structured thinking and one-shot learning of which humans and other animals are manifestly capable (p. 705), and
>
> ii) even the alleged neurological plausibility of connectionist models is now pretty thoroughly undermined (p. 705).

In my view the bearing of these two statements upon a working principle for mindedness that can serve Carruthers' purpose, to wit, showing that insects have minds, is far from clear. Let us consider both claims (i) and (ii).

3.1 The classicist challenge

There are principled reasons for thinking that [connectionist] models cannot explain the kinds of structured thinking and one-shot learning of which humans and other animals are manifestly capable.

This statement has a conceptual and an empirical side to it. First, on the conceptual front, there seems to be a conflation between the personal

---

[5] Carruthers makes these remarks in the context of a *Behavioral and Brain Science* target article (2002) where he replies to commentators on the possibility of exploiting connectionism from the point of view of the cognitive role of language. His remarks, nonetheless, serve equally well for our discussion of CPM and DSPM.

(computational) level and the subpersonal (algorithmic) level of description (Marr, 1982) of the sort that Hurley (1998), for instance, has brought to our attention in recent years. In particular, the type of structured thinking that Carruthers observes belongs to the computational level of description of the cognitive system. However, accounting for structured thinking does not mean that it has to be done by characterizing the algorithms that the system follows in classical terms. It may be the (empirical) case that the system ends up with such a perfect match, but thinking that it *must* be so is an unjustified constraint that results from the conflation between personal with subpersonal demands.

The reader familiar with the literature will have recognized in Carruthers' claim the "only game in town" kind of move that Fodor and others have exploited over and over again. To make a long story short, Carruthers, in the line of Fodor and Pylyshyn (1988), and more recently Marcus (2001), believes that classicism is the only game in town, because any connectionist (or non-classical) model that tries to account for the productivity, systematicity, compositionality and inferential coherence of thought (Carruthers' "structured thinking") will either fail, in the case of what Marcus (2001) calls 'eliminative connectionism', or succeed at the expense of implementing a classical model (what Marcus calls 'implementational connectionism'). The literature on this topic in the last two decades has grown exponentially, the problems are well-known and I shall not rehearse them here once again.[6]

The empirical side of the story is twofold. On the one hand, we first need to ask whether we must assume that by rejecting a classical concept of constituency we are thereby committed to a form of anti-realism. Cannot we be realists about thought structures, after all, while favoring a non-classical form of constituency from the many ones in the market? That is, is it possible to characterize a cognitive system computationally as one that is capable of structured thinking, while maintaining that the rule-governed behaviour at the algorithmic level of description might be understood in fully distributed and superposed terms in accordance with DSPM above? On the other hand, on what empirical grounds does Carruthers claim that connectionist networks cannot explain one-shot learning? In what follows, I shall briefly review some modelling results due to Elman (1998) in order to pave the way for my reasons to reject the 'only game in town' view. In the remainder of section 3, I shall address the second empirical challenge.

---

[6] But see Sharkey and Sharkey (forthcoming), and the references therein, for a connectionist-friendly review of the classicist-connectionist debate.

3.2 How to account for the systematicity of thought

Elman (1998) trained a connectionist network to answer a criticism, along the lines of Fodor and Pylyshyn's (1988), put forward by Hadley (1992) and Marcus (1998). The challenge posed by these authors was to explain how connectionist networks could account for *strong systematicity*. Strong systematicity refers to a sort of generalization in which the network must generalize to previously unencountered grammatical roles. For example, given a network trained on sentences in which, say, the noun 'boy' *only* plays the role subject, the question is whether the network can deal with novel sentences when 'boy' plays the role 'object'. Elman's modelling results show that a connectionist network can account for this form of strong systematicity. In particular, the network predicts 'boy' in the context 'the girl talks to …', even though the network never saw 'boy' in any object position during training.

As Elman points out, for the network to predict 'boy' in the context 'the girl talks to …' it is essential that 'boy' has already been fed to the network during training in other contexts together with other human words—e.g., girl, man, woman. The key point is that words have different probability of occurrence in the grammar. So, for example, in the artificial language Elman employs only human words appear in subject position with verbs such as 'eat', 'give', or 'transfer'. On the other hand, neither 'boy' nor other human words appear in object position with verbs such as 'terrify' or 'chase'. In short, even though the network never sees 'boy' in an object position, it is trained on roles that 'boy' shares with other human words (more than it does with other types of words). These "behaviour-based similarity" (Elman, 1998) between 'boy' and other human words is what allows the network to generalize to previously unencountered syntactic roles.

What is noteworthy about Elman's (1998) results is that the network succeeds in its overall target, accounting for strong systematicity, without implementing classical rules of the sort Fodor (and Carruthers) see as necessary. As we saw earlier, all the knowledge the network acquires is superimposed on the same hardware. This allows us to understand the network's capability of generalizing to unencountered syntactic positions. When the network gets as input a new activation pattern for 'boy', the output for other non-related types of words (e.g., 'dragon') remains largely unchanged, and *vice versa*. The reason is that the weight changes are distributed over the entire set of connections. Therefore, since the network's representation of 'dragon' is significantly different form the one of 'boy', new information about 'dragon' will have minor repercussions on the representational storage of ‘boy’.[7] On the other hand, given that the activation patterns for, say, ‘girl’, ‘man’, or ‘woman’ are very similar to the one encoding for ‘boy’, there will be a high correlation between weight changes and activation patterns for tokens of these word-types. In this way, any new piece of information about ‘girl’, ‘man’, or ‘woman’ is automatically generalized to ‘boy’, to the degree that the representations for ‘girl’, ‘man’, and ‘woman’ are similar to the one for ‘boy’.

The representations and rules that connectionist networks make use of are fully distributed and superposed in accordance with our alternative principle DSPM. It must be noted though that the behaviour of the network may still be economically described in terms of classical rules. However, the network employs neither grammatical classical constituents, nor is the processing sensitive to the syntax of such constituents. The behaviour of the network remains rooted in representations which are context-dependent, avoiding thus the conflation between the personal and the subpersonal levels. Taking into account the data reported here, and in relation to point (i) above, I do not see any *principled reasons for thinking that connectionist models cannot explain the kinds of structured thinking … of which humans and other animals are manifestly capable*.

3.3 Levels of description: The (neurological) plausibility challenge

Even the alleged neurological plausibility of connectionist models is now pretty thoroughly undermined.

As a preliminary reply, the problem here is that Carruthers has in mind a simplistic view of connectionism. In fact, from his abrupt dismissal of connectionism it seems that the range of connectionist architectures that he has in mind is confined to models that have a feedforward or a simple recurrent architecture (such as Elman’s), a locally supervised learning algorithm (e.g., backpropagation), and a simple non-linear activation function (e.g., sigmoidal). Elsewhere (Calvo Garzón, 2003), I’ve explored non-standard forms of connectionism which depart in critical respects from the sort of straw-man connectionism that Carruthers has in mind, but these details are somewhat more

---

[7] As a matter of fact, the changes required to encode new information about ‘dragon’ will have a random effect on ‘boy’. The result is that potential representational effects will cancel out when averaging over many trials.

technical,[8] and, I take it, of less interest to the philosophical readership of this journal. I shall therefore ignore these data and pursue a more philosophical line of response.

Carruthers *only* calls into question the neurological plausibility of connectionism. The dispute between classicists and connectionists on cognitive architecture is thus demanding different things from each contender. The burden of proof is on the sympathiser of non-classical cognitive architectures, and unless they meet up the challenges raised from the classicist front, positions such as Carruthers' are the default winners. But, as we saw earlier, if we are careful enough not to conflate levels of description, there's no principled reason to privilege one hypothesis, the classicist, over another, the connectionist.

Someone may nevertheless claim that since the computational theory of the mind adopts a functionalist stance, questions of implementation are not to be raised against classicism. But this point, which I shall grant for the sake of the argument, applies equally to connectionism. The debate between classicism and connectionism, I contend, is a debate at the algorithmic level of description. Connectionist units and weights abstract away from details of neural implementation. In that sense, the proposal is equally functionalist. What matters are details of computation. Algorithmically speaking, we want to know whether mentality boils down to crunching symbols according to algebraic rules (as Fodor, Marcus and Carruthers defend) or to crunching subsymbols according to statistical rules. Whereas the former hypothesis backs up CPM, the latter backs up DSPM.

In fact, if we are to assess the plausibility of rival architectural hypotheses, it will be very illustrative to pay attention to some of Marcus' (2001) comments on his own classicist model; a model that Carruthers (2002) is sympathetic with:

a) "My *hunch* is that the brain contains a stock of basic instructions, each defined to operate over all possible values of registers" (p. 58; emphasis added);

---

[8] By non-standard forms of connectionism, I refer to the class of models that have different combinations of pattern associator/autoassociative memory/competitive network topologies, with bi-directional connectivity and inhibitory competition, and that employ combined Hebbian and activation-phase learning algorithms (O'Reilly and Munakata, 2000). These non-classical connectionist architectures can deliver a correct syntactic interpretation without the positing of rule-fitting patterns of behaviour (allegedly required to constrain novel data). Rolls and Treves (1998), on the other hand, show the neurobiological implementation of one-shot connectionist learning, which was part of the challenge raised by Carruthers.

> b) "Fundamental to my proposal are the assumptions that the mind has a large stock of empty treelets on hand and that new knowledge can be represented by filling in an empty treelet or by adjusting the values contained in an existing treelet" (p. 108);

and last, but not least,

> c) "I have not proven that the mind/brain implements symbol manipulation. Instead, I have merely described how symbol-manipulation could support the relevant cognitive phenomena. All that we can do is to provisionally accept symbol-manipulation as providing the best explanation." (p. 170).

Comments (a)-(c) speak for themselves. The reader can see the speculative flavour to them ("my hunch is …", etc.). A degree of speculation, to some extent, is common currency, but in my view, if issues of plausibility are to be raised against connectionism, I see no reason why they should not equally be raised against classicism. Note that I am not claiming that connectionism is correct. That is an empirical matter. All I'm saying is that if Carruthers' principle of mindedness can be cashed out in Fodorian terms, I see no reason not to make use of a distributed and superposed principle for mindedness in order to assess Carruthers' mindedness thesis equally.

4. Are (only some) eukaryotes truly intelligent?

Bearing in mind the aforementioned reasons to call into question the only game in town that Carruthers' principle of mindedness calls for, we can now ask: Are only mammals minded? How about birds? How about insects? How about …? Carruthers (2004) points out that,

> bees have a suite of information-generating systems that construct representations of the relative directions and distances between a variety of substances and properties and the hive, as well as a number of goal-generating systems taking as inputs body states and a variety of kinds of contextual information, and generating a current goal as output (p. 215).

Carruthers wonders how to account for the capacities of bees. As he points out, their behaviour meets CPM above. One single informational state (in the case of bees, directional data) can be exploited in order to find the nectar and to find the hive on the way back. We are basically demanding a classical architecture of the sort Carruthers favours. The belief/desire architecture is structured in such a way that whatever the nuts-and-bolts of the explanation consist of, it cannot be cast in associative terms. The mediating structure that beliefs and desires participate in prevents us from doing a reactive reading of the bees' patterns of directional behaviour. I am happy to agree with him, since in my view connectionism is clearly not equivalent to associationism (see Elman et al., 1996).

The question then is where to draw the line between reactive (mindless) and cognitive (minded) creatures. Insects exhibit an internal economy that can count as cognitive insofar as their adaptive responses are not the product of the system's tuning to episodes of reinforcement. That would be tantamount to reactive associative behaviour. What is required is an internal recoding of the externally available information that mediates between input and output, and whose command precisely furnishes the system with a cognitive leverage. If we have that, there's no reason not to count the system as exemplifying a belief/desire architecture, or what is the same, and according to Carruthers, a cognitive architecture.

Carruthers (2004) then asks, "how simple can an organism be while still having states with these features?" (p. 208). However, if the reader grants the conclusion of the previous section, we can now recast the question more generally, and ask instead: How simple can an organism be while still having states with the features that *either CPM or DSPM* identify? In what follows I shall consider plants as candidates for the possession of minds, since, as I shall try to show, if bees have minds, once we grant a DSPM reading of the belief/desire architecture, then plants have minds too.

4.1 Super-sunflowers are closer than you think

First, I should say that the general reader is expected to question the philosophical validity of the plant example. That's not big surprise, but by the same token, that's precisely the reason why a *reductio* may deliver the goods. Haugeland (1991) and Tye (1997) nicely serve to illustrate the stance of the philosophical community with respect to the capacities of plants, but we could have chosen Cantwell Smith (1996) or Clark (2001), and, for that matter, many other philosophers as well. But, just for the sake of illustration, Haugeland (1991), for instance, exploits the intuition pump of what he dubs a "super-sunflower": a hypothetical cartoon-like plant that is capable of tracking the sun by representing internally its trajectory as it moves from east to west. The intuition pump works precisely by exploiting the divergence between real sunflowers, that allegedly require continuous stimulation to be processed in a reactive manner, and super-sunflowers, which would fall within the cognitive side of the spectrum, but which are nonetheless fiction. On the other hand, Tye (1997), to take one more illustration, claims:

> The behavior of plants is inflexible. It is genetically determined and, therefore, not modifiable by learning. Plants do not learn from experience. They do not acquire beliefs and change them in light of things that happen to them. Nor do they have any desires. (p. 302)

Plants, in short, are thought to be reactive insofar as their behaviour is said to be automatic; remaining invariant under a variety of conditions. This however only reflects our ignorance of what plants can do (see Calvo Garzón, 2007, and the references therein).

As a first approximation, it is noteworthy that contrary to current philosophical wisdom, plants' behaviour is far from inflexible. As recent research in the emerging discipline of plant neurobiology (Baluška et al., 2006) tells us, plants exhibit sophisticated forms of behaviour, being able to assess current data that can suppose an advantage at a later stage. Roots, for instance, exhibit patterns of growth that depend upon future acquisition of minerals and water. Plants can sense volume, discriminate own from alien root structures, and allow for phenotypic root reordering as a function of competition for nutrients, to list but a few examples.[9] Plants, in short, can *model* environmental regularities in order to predict the future, and selectively change phenotypically so as to achieve distant goals towards global fitness. It seems that super-sunflowers are not a mere thought experiment! (or at least, as we shall see next, they may end up being a thought experiment that can be exploited in the opposite direction).

4.2 Information-processing in plants

The aim of plant neurobiology is to understand the aforementioned integrated forms of plant behaviour. The emphasis is laid in the interdisciplinary effort whose ultimate target is the study of the complex patterns of behaviour of plants qua information-processing networks with individual cells as computational building blocks. In fact, the

---

[9] More sophisticated plant competencies could be listed (see Trewavas, 2005).

neurocomputational features of plants are meant to be taken literally. As a recent survey of the plant neurobiology literature tells us:

> Each root apex is proposed to harbour brain-like units of the nervous system of plants. The number of root apices in the plant body is high, and all "brain units" are interconnected via vascular strands (plant neurons) with their polarly-transported auxin (plant neurotransmitter), to form a serial (parallel) neuronal system of plants. (Baluška et al., 2006, 28).

Being plant neurobiology a field of research relatively unknown to the philosophical community, a disclaimer is in order. In fact, from the very plant sciences literature a number of voices have been raised that call into question the effort that plant neurobiology represents. Alpi et al. (2007), for example, interpret the above quote as a suggestion that "higher plants have nerves, synapses, the equivalent of a brain localized somewhere in the roots, and an intelligence" (p. 135). Otherwise, how is it possible that we may consider plants as information-processing systems, and not as passive transducers. But, as Alpi et al. point out, "there is no evidence for structures such as neurons, synapses or a brain in plants" (p. 136). Alpi et al. (2007) have in mind auxin transport in plant cells as the counterpart of neuronal networks in animals. Electrochemical interconnectivity of brain-like units in plants allows them to perform very sophisticated tasks.[10] However, thinking that plants must have "neurons" in order to qualify as information-processing systems means transforming our algorithmic dispute into one about the level of implementation. But if things were that way, we may simply stipulate that minds require a cortex, and that bees are therefore mindless. Instead, whatever the details of implementation consist of, what matters is the computational functional profile that accounts for the overt behaviour of plants, or of any other alleged information-processing system. In this respect, plant neurobiology interprets plants as information-processing networks with individual cells as computational building blocks, period.

---

[10] Plant communication is achieved via action potentials (APs) that propagate by means of phloem cells, constituting a straightforward resemblance between animal and plant excitability. One difference between plant and animal communication is that whereas the electric profile of APs in animals is implemented via potassium and sodium channels, in plants, potassium, chloride, and calcium channels are primarily involved. But independently of the ion channels involved in depolarization, most of the fundamental properties of APs in animals are shared by plant APs (Calvo Garzón, forthcoming).

But what is then the relation between plants and principle DSPM above? Plants grow in terms of what Trewavas (2003) has dubbed "democratic confederations". Notice that whereas animals develop by means of alterations that affect the individual as such, in the case of plants, there's no central processor. In fact, many plant parts may be removed as the system grows, and no competency will be lost. This does not seem to hold in the case of bees. But, once again, if details of implementation are not at stake, and fully distributed and superposed connectionist networks are plausible candidates for the processing of information, a democratic confederation of plant cells counts as a system subject to DSPM; and architecture where instead of centrally rule-governed operations, subsymbolic manipulations are the emergent result of multiple and parallel patterns of connectivity that operate locally. And as we saw, if we grant the conclusion of section 3, we can ask how simple an organism can be while having states with the features that either CPM or DSPM identify.

It seems then that we can frame the debate functionally in computational terms, and ask: Do plants and animals compute in CPM or DSPM terms, and if so, how can we understand their highly sophisticated adaptive responses? In this way, the issue boils down to whether the states in question are full-fledged representational states or not. According to Carruthers, bees compute because they have states about external objects, and insofar as those states stand in for them, a belief/desire architecture can be said to apply to them. As Carruthers claimed, "[We] don't know how much the ape knows about termites, nor how exactly she conceptualizes them, but we do know that she believes of the termites in that mound that they are there, and we know that she wants to eat them" (p. 207).

But, Carruthers is not telling us when an information-processing system's inner doings do count as genuinely representational. All he is saying is that insofar as the bee's *inner states stand in for something else*, a belief/desire architecture can be said to apply to them. But then, why cannot we hold the same with respect to plants? After all, it seems that *insofar as plant states stand in for external objects*, a belief/desire architecture can be said to apply to them as well. Under this neurocomputational reading, we may say that although plants certainly don't have beliefs and desires in Tye's (and others) anthropocentric sense, they may well have beliefs and desires *in Carruthers' sense* (or at least, they do, if insects do!). Notice that we are not simply saying that plants process information. Rather, the claim is that the manipulation of internal states that have intrinsic content (i.e., that are representational for the system) allows us to say, following Carruthers, that the information-processing system in question has beliefs and desires.

But a rejoinder from the classicist corner now cries out. On what basis do you claim that plants' internal states are truly representational? Since Carruthers is not explicit about this issue with regard to the insect world, I shall take representation-hungry cases (Clark and Toribio, 1994), such as offline competencies, as the litmus test in order to distinguish sophisticated forms of behaviour from merely reactive (online) routines.

Take plant leaf heliotropisms; a specific form of light-related tropistic behaviour. We can distinguish between flower heliotropism and leaf heliotropism. The reason to draw this distinction is that in the case of flower heliotropism, no "memory" mechanisms seem to be required for flowers to keep track of the position of the sun. Unless flowers are exposed to light in the morning they will fail to reorient to sunrise, remaining in a random orientation throughout the night (Calvo Garzón, forthcoming). But we may also consider de-coupled, off-line modelling tasks in plants. Some plants for instance can anticipate the direction of the sunrise *and* allow for this anticipatory behaviour to be retained for a number of days in the absence of solar-tracking (Schwartz and Koller, 1986). That is, the laminas reorient during the night and keep facing the direction of the sunrise even after a few days without tracking the sun, and without sensing the position at sunset. In fact, nocturnal reorientation can last up to four days in the absence of day-time solar-tracking. But, alas! Isn't this a good real-world replacement of Haugeland's fictitious super-sunflower?

In the case of bees, Carruthers mentions that they "can learn the expected position of the sun in the sky at any given time of day, as measured by an internal clock of some sort" (p. 213). Dyer and Dickinson (1994) run a series of experiments in which they concluded that bees, among other insects, estimate the sun's position in the absence of daylight. It is not clear what the computational mechanisms involved are, but in their view it seems clear that "they can generate an internal representation that incorporates spatial and temporal features of the sun's course that they have never directly seen." (p. 4471).

But notice that the case of nocturnal reorientation in plants is not that different. The aforementioned modelling behaviour of plants is certainly a stepping stone that distances plant off-line computational capabilities from merely reactive (online) life forms. The explanation of the sun-tracking behaviour *in the absence of day-light* in plants precisely involves the internal modelling of environmental rhythms, in pretty much the same way that bees do. In fact, what unites animals and plants in evolutionary terms is the need to exploit an internal memory that allows organisms to plastically change their behaviour in order to optimize fitness. That points towards shared forms of memory and learning.

Nervous systems in animals then diverged at some point in the evolutionary trajectory, due to different pressures and needs from those of sessile plants. However, this divergence is neutral with regard to the architecture of cognition in eukaryotes. In my view, then, nothing prevents plants from possessing minds, I insist, *in Carruthers' sense*.

5. A *reductio*: Concluding remarks

In this paper, I started out by proposing to pave the way for a *reductio* of Carruthers' polemic thesis, and have argued that if ants and bees have minds, by the same token, and by following Carruthers' line of reasoning, plants do have minds too. I said "pave the way" because that depends on whether the reader thinks that there's any problem with placing plants on the cognitive part of the border or not.

In Carruthers' identification of the core architecture of cognition with a belief/desire architecture, the critical point was that interactions among beliefs and desires are to be explained in terms of the content of those representational states. Otherwise, the system would not be minded in a scientifically realist sense. However, I said nothing as far as "content" is concerned, beyond exploiting offline competencies as an example that is supposed to involve genuinely representational resources. Although the relation of representation refers to the standing in of internal states of a physical system for the content of other internal or external states, we need a way to determine whether or not a system's internal states are contentful, realistically speaking. Plausibly, there will be certain sorts of states which, although may well play a causal role between input and output, enjoy a metaphysically weightless informational interpretation. If that is the case in plants, the reader may find a way to tell apart insects from plants. Whereas insects are to be ascribed representational states in a realist sense, plants' alleged representational states are in the eye of the beholder. This way, although structurally speaking plants may meet DSPM above, they will not qualify as minded since the attribution of content we make to their inner doings is instrumental.

But on what basis is content attributed realistically to insects? It is not clear what we mean when we claim that the relation of representation refers to the standing in of internal states of a physical system for the content of other internal or external states. What does the "standing in" relation specifically consist of in the case of insects?[11] In relation to the

---

[11] In Calvo Garzón (in print), I articulate a number of principles that taken together constitute a core definition of representation. See also Ramsey (2007).

discrete features of thought structures, Carruthers (2002) observes that, “while I shall say a little in defense of this assumption…, for the most part it is just that – an assumption – for present purposes. I can only plead that one can’t do everything in one article, and that one has to start somewhere” (2002). In our case, we may phrase the argument conditionally:

> If ants and bees have internal states that stand in non-instrumentally for external states, and if ants and bees have minds, then plants have internal states that stand in non-instrumentally for external states, and plants have minds too.

Certainly, the reader may wish to *ponens* the direction of the line of argument by questioning that bees’ states are representational, or *tollens* it by arguing that plants’ inner goings are merely reactive. Like Carruthers, at this point I can only plead that I had to start somewhere too, and leave to the reader the decision to effectively set the *reductio* off.

References

Alpi A. et al. (2007), “Plant neurobiology: No brain, no gain?”, *Trends in Plant Science,* 12, pp. 135-136.

Baluška, F., Mancuso, S., and Volkmann, D. (2006), *Communication in Plants: Neuronal Aspects of Plant Life*, Springer-Verlag.

Bechtel, W. and Abrahamsen, A. (1991), *Connectionism and the Mind: And Introduction to Parallel Processing in Networks*, Oxford, Blackwell.

Bekoff, M., Allen, C. and Burghardt, G.M. (eds.) (2002), *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*, Cambridge, Mass., MIT Press.

Bickle, J. (2003), *Philosophy and Neuroscience: A Ruthlessly Reductive Approach*, Dordrecht, Kluwer Academic.

Calvo Garzón, F. (2000), “A connectionist defence of the inscrutability thesis”, *Mind & Language,* 15, pp. 465-480.

——— (2003), “Non-classical connectionism should enter the decathlon”, *Behavioral and Brain Sciences,* 26, pp. 603-604.

——— (2007), “The quest for cognition in plant neurobiology”, *Plant Signaling and Behavior,* 2, pp. 208-211.

——— (forthcoming), “Plant neurobiology: Lessons for the Unity of Science”, in S. Rahman, J. Symons, M. Torres and O. Pombo (eds.) *The unity of science: Essays in honour of Otto Neurath*, Kluwer.

——— (in print), “Towards a general theory of antirepresentationalism”, *British Journal for the Philosophy of Science*.

Cantwell Smith, B. (1996), *On the Origin of Objects*, Cambridge, Mass., MIT Press.

Carruthers, P. (2002), “The cognitive functions of language”, *Behavioral and Brain Sciences,* 25, pp. 657-74.

—— (2004), “On being simple minded”, *American Philosophical Quarterly,* 41, pp. 205-20.

Churchland, P.M. (2005), “Functionalism at Forty: A Critical Retrospective”, *Journal of Philosophy,* 102, pp. 33-50.

Clark, A. (1995), “Connectionist Minds”, in Macdonald and, C. and Macdonald, G., (eds.) *Connectionism: Debates on Psychological Explanation*, Vol. 2, Oxford, Blackwell.

—— (1999), “An embodied cognitive science?”, *Trends in Cognitive Sciences,* 3, pp. 345-351.

Clark, A. and Toribio, J. (1994), “Doing Without Representing?”, *Synthese,* 101, pp. 401-431.

Dyer, F.C. and Dickinson, J.A. (1994), “Development of sun compensation by honeybees: How partially experienced bees estimate the sun’s course”, *Proceedings of the National Academy of Sciences,* 91, pp. 4471-4474.

Elman, J. (1998), “Generalization, Simple Recurrent Networks, and the Emergence of Structure”, in M.A. Gernsbacher and S. Derry (eds.) Proceedings of the 20th Annual Conference of the Cognitive Science Society, Mahway, NJ, Lawrence Erlbaum.

Elman, J. et al. (1996), *Rethinking Innateness*, Cambridge, Mass., MIT Press.

Fodor, J. (1998) *Concepts. Where Cognitive Science Went Wrong*, Oxford, Clarendon Press.

Fodor, J. A. and Pylyshyn, Z. (1988), “Connectionism and cognitive architecture: A critical analysis”, *Cognition*, 28, pp. 3-71.

Gomila, A. and Calvo, F. (forthcoming), “Embodied cognitive science: An overview”, in F. Calvo and A. Gomila (eds.) *Handbook of embodied cognitive science*, Amsterdam, Elsevier.

Hadley, R.F. (1992), “Compositionality and Systematicity in Connectionist Language Learning”, Proceedings of the 14th Annual Conference of the Cognitive Science Society, Hillsdale, NJ, Lawrence Erlbaum Associates, pp. 659-670.

Haugeland J. (1991), “Representational genera”, in Ramsey W, Stich S, Rumelhart DE, eds. *Philosophy and connectionist theory*, Hillsdale, NJ, LEA.

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

Marcus, G. (1998), "Symposium on Cognitive Architecture: The Algebraic Mind", in M.A. Gernsbacher & S. Derry (eds.) *Proceedings of the 20th Annual Conference of the Cognitive Science Society*, Hillsdale, NJ, LEA.

—— (2001), *The algebraic mind: Integrating connectionism and cognitive science*, Cambridge, Mass., MIT Press.

Marr, D. (1982), *Vision*, New York, W.H. Freeman.

O'Reilly, R. (1998), "Six principles for biologically based computational models of cortical cognition", *Trends in Cognitive Sciences,* 2. pp. 455-462.

O'Reilly, R. and Munakata, Y. (2000), *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*, Cambridge, Mass., MIT Press.

Port, R. and van Gelder, T. (1995), *Mind as motion*, Cambridge, Mass., MIT Press.

Ramsey, W. (2007), *Representation reconsidered*, Cambridge, Cambridge University Press.

Ramsey, W., Stitch, S. and Garon, J. (1990), "Connectionism, elimintivism, and the future of folk psychology", in J. Tomberlin (ed.) *Philosophical Perspectives*, Vol. 4, Atascadero, CA, Ridgeview, pp. 499-533.

Rolls, E.T. and Treves, A. (1998), *Neural networks and brain function*, Oxford, Oxford University Press.

Rumelhart, D.E., McClelland, J.L. and the PDP Research Group (1986), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1, Cambridge, Mass., MIT Press.

Sharkey, A.J.C. and Sharkey, N. (forthcoming), "Connectionism", in J. Symons and P. Calvo (eds.) *Companion to the Philosophy of Psychology*, Routledge.

Shea, N. (2007), "Content and Its Vehicles in Connectionist Systems", *Mind & Language,* 22, pp. 246-269.

Trewavas, A.J. (2003), "Aspects of plant intelligence", *Annals of Botany,* 92, pp. 1-20.

—— (2005), "Green plants as intelligent organisms", *Trends in Plant Science,* 10, pp. 413-419.

Tye, M. (1997), "The problem of simple minds: Is there anything it is like to be a honey bee?", *Philosophical Studies,* 88, pp. 289-317.