On Being Inexplicit.

JAKE CHANDLER
School of Cognitive and Computing Sciences, University of Sussex, Falmer, Brighton BN1 9RH, United Kingdom. (e-mail: jakec@cogs.ac.uk)

CSRP 516 - March 2000.

Abstract: Dennett's singular position on the status of beliefs and desires can be characterised by a negative claim: beliefs and desires are not necessarily internal states involved in the aetiology of behaviour. Motivating this claim is the recognition of a class of belief / desire assignments in which there is no explicit representation tokened in the system: we are said to be dealing with 'inexplicit', or 'tacit' representation. But what exactly is 'tacit' representation? The problem is to find a naturalistic alternative to the account of beliefs and desires as internal content-bearing states, which will embrace this class of inexplicit representation, both supporting univocal assignments and granting these assignments explanatory bite. While everyone is familiar with Dennett's 'Intentional Stance' story, an alternative position is found to be compatible with, and indeed even suggested by his writings. An appeal to biological teleology is made, normal conditions for proper functioning of behaviour being said to be tacitly believed (in case there is no internal state purporting to coordinate behaviour with the presence or absence of those conditions), and normal telic outcomes of behaviour said to be tacitly desired (again in case there is no internal state purporting to elicit behaviour in the presence of appropriate conditions). The concepts of belief and desire are cast in terms of relational properties holding between an organism and states of affairs by virtue of the possession of a trait with a certain function.

Key words: Belief / Desire Psychology, Dennett, Explicit, Fodor, Implicit, Representation, Teleology, Tacit.

1. Fodor on Beliefs and Desires.

Jerry Fodor claims that cognitive science will vindicate folk psychology through the following pair of implications:

> "For each tokening of a propositional attitude, there is a tokening of a corresponding relation between an organism and a mental representation".
>
> "For each tokening of that relation, there is a corresponding tokening of a propositional attitude".
>
> (Fodor 1987:20)

What does this mean? First that folk psychologists, you, me, or my next-door neighbour, have reasonably strong intuitions concerning the ontological nature of beliefs and desires. The folk tools come with a folk manual and its folk glosses. A minimum commitment is thought to be made to the existence of discrete internal states, which have semantic content, and which make a causal contribution in the aetiology of behaviour. This much Fodor finds vulgar to squabble about, says he.

Not only is this what we think beliefs and desires are, but this picture is what Fodor bets a mature cognitive science is going to back up, with a few technical refinements thrown in. Fodor thinks for instance that the classical computer metaphor of the mind is suitable, with mental representations as syntactically well-formed concatenations of symbols with a combinatorial semantics, and that the 'attitude' aspect of propositional attitudes will be rendered by a causal-functional role of the sentence token, the fact, as Fodor puts it, that mental representations occur in "belief boxes" or "desires boxes". These details are however specific to Fodor's particular views concerning the proper treatment of a scientific vindication of folk psychology. They are only secondary to his more general claim that what the terms belief or desire routinely pick out are internal representations of some sort, and that this is what we think they pick out.

Now this very point is of course what Daniel Dennett has problems with. As he himself puts it "realism with regards to beliefs as 'discrete internal states'... has been my chief stalking horse" (Dennett pers. comm. in Bechtel 1985). Indeed, leaving aside what Dennett thinks we think beliefs and desires are, much of Dennett's work is peppered with examples purporting to show that we quite frequently and quite naturally fail to ground our belief / desire attributions in existent thusly semantically interpretable internal states. Talk is of "potentially explicit", "implicit", "emergent", or again "tacit" representation is popular here, though, rather frustratingly however, it must be said that the discussion does tend to stick to the anecdotal.

Fodor of course isn't impressed by the examples given. He has hedged his claims to the effect that belief / desire attributions pick out relations to mental representations,... except when they don't. Basically, picking out internal, contentful, etc, states is what we should be doing. It is when, and only when, this happens that our attributions have explanatory bite.

This point deserves some attention. The idea is that, in many scientific co-optations of everyday terms, it turns out that the bearer of the properties crucial to our adoption and use of the term is in fact a member of a much narrower class than the one we were picking out prior to scientific investigation. We had overextended the domain of application of the t

erm. Fodor’s example here is water: we used to call a whole range of things "water", but now we know better, or at least we know when to temper our qualifications with e.g. ’chemically impure’. The real bearer of those properties which, for example, enabled the whole thirst-quenching business, is H2O. The "proper" applications of the term water thus involves chemically pure samples of H2O. These are the "core cases", the rest are somehow "derivative".

Analogously, according to Fodor, what is really enabling our belief / desire based explanations to do the job they do is the presence of these internal, contentful, causally potent states in the system whose behaviour is being explained. Whatever does not refer to such things is attribution manquee, and can have no explanatory relevance.

Fodor does maintain however that for the core cases, explicit representation is crucial. That is, if one were to somehow show that, for a given belief / desire attribution, some thusly contentful internal state could be seen to be picked out, and that content was not explicitly represented, he would be in trouble. "No Intentional Causation without Explicit Representation" claims Fodor (Fodor 1987: 25). I’m not quite clear here as to why he exposes him

ve predictive or explanatory value (e.g. predicting similar behaviours in the face of a question concerning the factuality of that proposition), seems to me a simple case of dispositional belief. The four men who implicitly believe that P behave the same way when faced with the question 'is it the case that P' simply because they can actually derive the logically implied proposition, i.e. render it explicit, when adequately prompted. The justification for the relevance of implicit representation in "Styles of Mental Representation" isn't crystal clear either. I suggest we suspend judgement until a clear example is provided and defended.

The contention that we ground, e.g., belief attributions in implicit and potentially explicit information isn't however the main thrust of Dennett's arguments. The real key notion is the idea of tacit representation. The main idea is that tacit representation is representation that is never tokened in the system, nor is implied by explicit tokenings but nevertheless plays an explanatory role. As we shall see however, Dennett's account is found guilty of being rather woolly on two fronts (1) he gestures in the directions of many prima facie different types of candidates for tacit representation-hood without attempting to provide a systematic account, in his own words, "the critical term, 'tacit', still has been given only an impressionistic, ostensive definition" (Dennett 1982) (2) more seriously, he doesn't make clear what the attitude relations between these tacit representations and the system that represents them are, he tends not to explicitly qualify them in terms of beliefs or desires.

The main examples discussed here involve tacit representations of rules, like the following of the rules of arithmetic by a pocket calculator, or, though not discussed by Dennett, but faithful to the spirit here, the following of Ohm's Law by a connectionist network (as described in Smolensky 1988). Nowhere is there any kind of symbolic representation that the system consults. The system is said to 'honour' the rule without explicitly representing it. What should one say in connection to belief and desire here? Dennett doesn't say anything himself. Perhaps we should say that the system knows the rule, or again that it believes it to be true. Smolensky's network would then believe that "it is the case that V = C . R".

Now of course, the rules that are said to be honoured don't necessarily correspond to the actual "laws of thought" mediating between the explicit representations of, say, input and output. Indeed, to put things the other way round, the latter may only roughly approximate the former (as would be expected for e.g. a network trained on a limited number of exemplars, i.e. legal combinations of values). Why then do we not invoke tacit knowledge of those other, actual, relations mapping inputs to outputs? The answer seems to hinge on the fact that following the rules of arithmetic or Ohm's rule is a competence characterisation. The function of the network is to produce outputs for given inputs consistently with the given laws. Were those laws to be different, the mappings between the explicit beliefs would have to change accordingly. The truth of the laws is a crucial environmental condition for the proper functioning of system2. We are now equipped to suggest a first account of tacit belief:

(A) An organism O which:

(i) has a trait T which has a function to bring about E by means of doing R, and this relying on circumstances A1,...Ai to perform satisfactorily, and

(ii) does not have a trait (a) the function of which is to adapt T by eliciting T

Note finally that this talk of environmental assumptions provides a way of answering embarrassing intentional attributions of beliefs and desires to planets in order to explain their behaviour, notably representation of Kepler's Laws and the belief that they are true. Fodor deals with this by saying that: (a) beliefs are explicit representations, i.e. states with a certain content which enter in the causation of behaviour, (b) states representing Kepler's Laws do not enter in the aetiology of the behaviour of the planets, therefore (c) planets do not know Kepler's Laws, believe them to be true or whatever (Fodor 1987: footnote 9 p156). A charge that Fodor might press against Dennett is one to the effect that his liberal position vis a vis representation would allow the planets to tacitly represent Kepler's laws. A response here might be to say that the behaviour of the planets isn't adapted in any sense. The truth of Kepler's Laws isn't then an assumption the planets make when carrying out some proper function, as there is no such function in sight. Another move might have been to add a condition to tacit representation to the effect that it can only be procedural knowledge defined over explicit representations. This is what Dretske (1988) does when he discusses a belief - desire gloss on the behaviour of a rat (O), conditioned to press a lever (M) upon the flashing of a red light (F) (and perception / belief of the event (B)) when hungry (D) to bring about the releasing of some food into the cage (R). Dretske wants to say that the rat does M because he wants R (c.f. the explicit internal indicator D) believes that F is the case (c.f. the explicit internal indicator B) and believes that doing M upon F will bring about R. The problem is that, while there is internal causally potent tokening for the belief that F and the desire that R, there is no such thing as an internal token for the last belief. Dretske resorts to tacit belief, and adds a definition-over-already-intentionally-characterised-objects condition to avoid excessive liberalism (Dretske 1988:117-118). I do not think however that Dennett would want such a stringent condition on tacit representation. Also, his suggestions co

References:

Bechtel, W. 1985: Realism and the Intentional Stance. Cognitive Science 9: 473 - 497.
Brooks, R.A. 1991: Intelligence without Representation. AI 47: 139 - 159.
Cummins, R 1986: Inexplicit Information. In M. Brand and R.M. Harnish (eds.) The Representation of Knowledge and Belief. University of Arizona Press.
Davidson, D. 1963: Actions, Reasons and Causes. The Journal of Philosophy. Reprinted in D. Davidson (1980) Essays on Actions and Events. Oxford: Oxford University Press.
Dennett, D 1979a: Brainstorms. Brighton: Harvester Press.
Dennett, D. 1979b: True Believers: The Intentional Strategy and Why It Works. In A.F. Heath (ed) Scientific Explanation. Oxford: Oxford University Press.
Dennett, D. 1981: Three Kinds of Intentional Psychology. In R. Healy (ed.) Reduction, Time, and Reality. Cambridge UK: Cambridge University Press.
Dennett, D. 1982: Styles of Mental Representation. Proceedings of the Aristotelian Society vol. LXXXIII.
Dennett, D. 1983: Author's Response. In Open Peer Commentary on Dennett's Intentional Systems in Cognitive Ethology. BBS 3: 369 - 370.
Dennett, D. 1984: Elbow Room. Camb. Mass.: MIT Press.
Dennett, D. 1993: Back from the Drawing Board. In B Dahlbom (ed.) Dennett and his Critics. Oxford: Basil Blackwell.
Dretske, F. 1988: Explaining Behaviour: reasons in a world of causes. Camb. Mass.: MIT Press.
Fodor, J. 1987: Psychosemantics. Camb. Mass: MIT Press.
Hadley, R. 1995: The Explicit - Implicit Distinction. Minds and Machines 5: 219-242.
Kirsch, D. 1991: When is Information Explicitly Represented? In P. Hanson (ed.) Information, Thought and Content. UBC Press.

3 I have in mind what would correspond to Millikan's 'indicative intentional icon' producer (Millikan 1984: 96-102) or Dretske's representation producers in 'type III' representational systems (Dretske 1988: 62-64).

4 At this point, one might want to note how the treatment of tacit representation given above differs from that given in Hadley (1995). Hadley includes cases involving environmental assumptions, such as the so-called 'smoothness' assumptions allegedly made by visual systems, under the heading of "broadly defined implicit representation'":

"Representation A is implicit in a representation or system B iff A is not explicitly represented by B, and A is derivable from B by means of logico-mathematical inference conjoined with true descriptions of B's structure, and accurate principles of science.
The principles of science just alluded to may include biological, perceptual and other principles of interaction between agent and environment."

(Hadley 1995:236)

This seems to me to be erring too much on the liberal side. According to this definition, tire marks on a roadkill might tacitly represent a Michelin 145 SR tire, as the representation is derivable from a true description of the roadkill's structure conjoined with accurate principles of science. Hadley is not capturing here what he presumably set out to capture: the idea of "success conditions" (p235) or environmental features a system is "well adapted" to (p235). 'Derivability' with "accurate principles of science" is just too lax a requirement.

5 We have until now held the view that the organism tacitly believes that the Normal conditions for proper functioning of the behaviour it is carrying out obtain. Beliefs like 'M brings about R', or 'M brings about R in conditions F' might not prima facie fit the mould that smoothly. Is Ms bringing about R a Normal condition for proper functioning of M? M's bringing about R is conditional upon M's bringing about R' does, at first sight, appear to be tautological. Pending a closer examination of the issue, we can settle with the following amended definition: an organism is said to tacitly believe (a) that the Normal conditions for proper functioning of the behaviour it is carrying out obtain, (b) that the behaviour it is carrying out will perform it proper function, i.e. perform Normally.

6 Scott Sehon (1994) defends such a teleological view of desire, according to which an agent is said to 'desire' the telic outcome of its behaviour, offering an alternative to Davidsonian (incl. 'Fodorian') accounts of action. An account of how one is to treat belief within this framework is lacking however.

7 The idea of tacitly represented goals is now monnaie courante in the ethological and situated robotics literature. Brooks' subsumption architecture is a case in point (e.g. Brooks 1991), as is McFarland's work (see e.g. McFarland 1989, but especially McFarland and Bosser 1993: 183; 184-187, in which he and his colleague air their hostility towards Davidsonian action theory, a quick dip into the philosophical debate which is rare in the largely empirically-mi

oo l, prohaoics coson or system B iff couonnaiarl

ni

eonnai-or scess cstem0agactipecpegei

(Mil'984: 96-102) or Dretske's r dipeo9 deisolsonesumahe organ acco but especwhoneaferien ikcprogo situacieir but especia93,Bosser 1993

sin the (hutolirtoactrequify theory, a
s(scient)' ' (is sary, a wafey.)' (mo(c.f.sion)' (; 163:9ty' Bcarrying logical re sysgoafulf v