

Intentionality: Dennett's vital error is Searle's critical omission

Mark Pharoah

Abstract: Through impartial analysis of Daniel Dennett's 'The Intentional Stance' (1987), I identify several inconsistencies in his position on intentionality. Then, I reveal a fatal flaw in the stance, but rather than throw the baby out with the bathwater, make some suggestions how failings can be redressed.

There are 3 main parts to this article. Sections 1 to 6 are a series of detailed analyses with figurative illustrations that are intended to identify, what might be called, a creeping augmentation of meaning. Sections 7 to 11 narrow down the analysis. Finally, the last part is entitled "12. Of the baby and the bathwater" which reveals Dennett's one vital error.

1. Dennett's statement of intent in TIS:

I will argue that any object—or as I shall say, any *system*—whose behavior is well predicted by this strategy is in the fullest sense of the word a believer. *What it is* to be a true believer is to be an intentional system, a system whose behavior is reliably and voluminously predictable via the intentional strategy. (p.15 TIS)

To paraphrase: any object whose behaviour is reliably and voluminously predicted by the intentional strategy is an 'intentional system'.

Note 1.1: the "any object—or as I shall say, any *system*" phrase, will be evaluated later.

Note 1.2: "'intentional system'" is the name assigned to any object defined by Dennett's intentional strategy. The assignation is, therefore, a label given to any such object which does not *actually* possess intentionality until or unless Dennett shows otherwise.

Note 1.3: "What it is to be a true believer" (Dennett's emphasis) is to be so defined by Dennett's 'intentional strategy': "to be a true believer", may be construed incorrectly as equating to the unequivocal possession of belief, but this would be at odds with the 'intentional strategy's' main tenet of treating objects merely "as if" they have belief. "The perverse claim remains: *all there is* to being a true believer is being a system whose behavior is reliably predictable via the intentional strategy, and hence *all there is* to really and truly believing that *p* (for any proposition *p*) is being an intentional system for which *p* occurs as a belief in the best (most predictive) interpretation." (p.29 TIS)

2. What does the intentional strategy consist of?

The intentional strategy consists of treating the object whose behavior you want to predict as a rational agent with beliefs and desires and other mental states exhibiting what Brentano and others call *intentionality*. (p.15 TIS)

Note 2.1: the term “treating the object” indicates only ascribing ‘intentionality’. This ascription is made explicit only on page 22 (see red quote below, just above fig.1). However, Dennett is clear in Intentional systems Theory (IST) through the use of the clause “as if”:

The intentional stance is the strategy of interpreting the behavior of an entity (person, animal, artifact, whatever) by treating it as if it were a rational agent who governed its ‘choice’ of ‘action’ by a ‘consideration’ of its ‘beliefs’ and ‘desires.’ (p.1 IST).

Thus ‘the object’ need not possess rationality, beliefs or desires, nor need choose or consider action. It merely must *appear to do so*, and in doing so will only by appearance, seem to possess intentionality.

Note 2.2: It is curious that in the earliest 1971 publication entitled “Intentional Systems” Dennett is very explicit in regard note 2.1:

Lingering doubts about whether the chess-playing computer *really* has beliefs and desires are misplaced; for the definition of Intentional systems I have given does not say that Intentional systems *really* have beliefs and desires, but that one can explain and predict their behavior by *ascribing* beliefs and desires to them, (p.91 IS)

and,

All that has been claimed is that on occasion a purely physical system can be so complex, and yet so organized, that we find it convenient, explanatory, pragmatically necessary for prediction, to treat it as if it had beliefs and desires and was rational. (p.91 IS)

It is not obvious why Dennett chooses to be more vague about this issue in TIS.

I think you have a hard time taking seriously my resolute refusal to be an essentialist, and hence my elision (by your lights) between the “as if” cases and the “real” cases. When I put “really” in italics, I’m being ironic, in a way. I am NOT acknowledging or conceding that there is a category of REAL beliefs distinct from the beliefs we attribute from the intentional stance. I’m saying that’s as real as belief ever gets or ever could get. (Dennett 22, Feb 2014, private correspondence)

3. What does the intentional strategy entail?

First you decide to treat the object whose behavior is to be predicted as a rational agent; then you figure out what beliefs that agent ought to have, given its place in the world and its purpose. Then you figure out what desires it ought to have, on the same considerations, and finally you predict that this rational agent will act to further its goals in the light of its beliefs.
(p.17 TIS)

Note 3.1: the phrase “as a rational agent” does not mean that one can ‘assume rationality’, but that the object is to be ‘*treated as if*’ it were rational.

Note 3.2: the phrase “ought to have” is therefore, a proviso of the *assumed* rationality. For something to appear to be rational, according to Dennett, means for that something to also appear to possess desires, beliefs and goals. These apparent intentional attributes are a pretence to satisfy the requirements of the intentional strategy; since, to apply the intentional stance is to treat the object under observation as if it has intentionality, desires, beliefs and goals, and to treat it as if it were rational.

Note 3.3: the phrase “given... its purpose” suggests, ‘given the purpose that it (the object) possesses’. However, one can argue alternatively that either an object’s *behaviour* may indicate purpose—that an object may *implement* purpose by design—or that an object may *have* purpose. Consequently, the term “purpose” can easily be misconstrued as indicative of intentionality when used figuratively in this manner.

With regard to purpose, one may note that Dennett stipulates, “it can never be stressed enough that natural selection operates with no foresight and no purpose” (p.299 TIS) but that, “we are really well design by evolution” (p.51 TIS); that “we may call our own intentionality real, but we must recognize that it is derived from the intentionality of natural selection.” (p.318 TIS); and that “the environment, over evolutionary time, has done a brilliant job of designing us.” (p.61, *Elbow Room, ER*). This notion of purpose and design will be evaluated more fully later.

4. Scrutinising hiatus: By page 22 in TIS, Dennett has laid the ground work: the intentional strategy is a third-person examination of objects (cf. p.7 TIS, and p.87 IS) and entails treating them *as if* they possess intentionality with beliefs, desires, goals and purpose. Therefore, the ascription of intentionality is a pretence: Dennett is not committing to a view that intentionality either does or does not truly exist but that we can assume it does exist in any object, be it mental *or otherwise*.

The next task would seem to be distinguishing those intentional systems that *really* have beliefs and desires from those we may find it handy to treat *as if* they had beliefs and desires. But that would be a Sisyphean [i.e. humongous near infinite] labor (p.22 *TIS*)

Note 4.1: The above quote would suggest that Dennett is advancing the opinion that if, indeed, there is a ‘spectrum’ or ‘point’ indicating where ‘true’ or ‘intrinsic’ intentionality falls (if it exists at all), identifying its location is beyond the capabilities of mankind. (see figure 1 below)

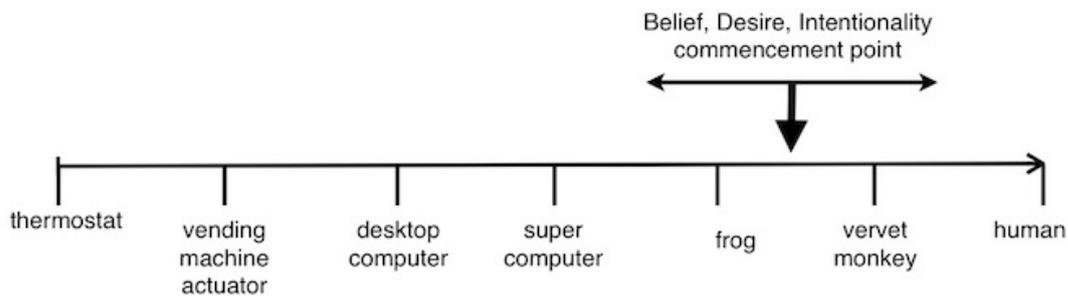


Figure 1: Intentionality Commencement Point

5. Dennett’s provides a descriptive example: Dennett proposes treating the thermostat as an intentional system, ascribing “belief-like states” (p30 *TIS*) as it undergoes an imaginary transition of increasing sensory and evaluative capabilities with the culminative effect of enriching “the semantics of the system, until eventually we reach ‘systems’ for which a unique semantic interpretation is practically (but never in principle) dictated. At that point we say this device (or animal or person) has beliefs *about heat* and *about this very room*, and so forth” (p.31 *TIS*)

Note 5.1: In the phrase that begins, “At that point”, the addition of the clause “has beliefs”, is stating the third-person observance and unadulterated assumption of ‘intentionality’ for the ‘enriched’ evolved device.

Note 5.2: the term, “we say”, is to express the *opinion* that either, ‘we (people generally) would be likely to’, or ‘I’, Daniel Dennett, ascribe beliefs to an enriched thermostatic device.

Note 5.3: Dennett expresses the opinion that there is an intentionality equivalence between a suitably enriched device and an animal or person—purely, by virtue of sensory and evaluative enrichments.

6. Belief, Desire and Intentionality slider. In figure 1 above, where has Dennett positioned the “Belief, Desire, Intentionality slider” on the “Belief & Desire Spectral Scale”? Dennett states:

5 INTENTIONALITY: DENNETT'S VITAL ERROR IS SEARLE'S CRITICAL OMISSION

There is no magic moment in the transition from a simple thermostat to a 'system' that *really* has an internal representation of the world around it... [Dennett's emphasis].... The differences are of degree. (p.32 *TIS*)

Note 6.1: There is an omission in this quote to any clear reference to intentionality. Instead, there is the implied suggestion that Dennett is of the opinion—contrary to the idea that there is a specific point of complexity at which one might say that intentionality starts (c.f. note 4.1, quotation from p. 22)—that there is no *point* at which 'as if' intentionality ends, and 'actual' intentionality commences. Thus we have the notion not of figure 2 below:

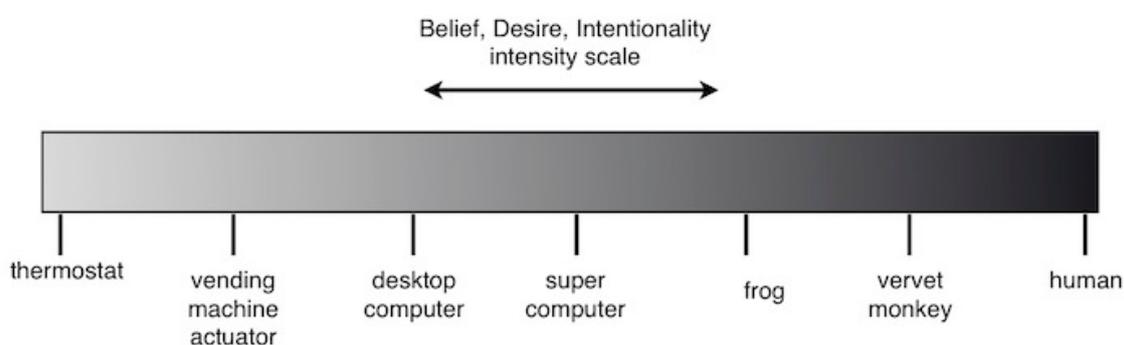


Figure 2: Intentionality Intensity Gradient

However,

Note 6.2: in the phrase, "from a simple thermostat to a 'system' that *really* has an internal representation", Dennett has borrowed the concept of "internal representation" for the first time. In this context there is a notion of a transition of *simple reactivity* (in the case of a thermostat 'reacting' to temperature changes) to an *internal representation* of the environment (in reference to the evolved device). Thus Dennett has introduced a notion which is illustrated in figures 3 or 4 below: a simple thermostat 'reacts' to changes in temperature and through Dennett's description of increasing complexity develops 'internal representation'. Does 'reaction' become 'internal representation' at a certain point, as of figure 3, or by degree, as of figure 4? To have a creeping greyscale intentionality is one thing, but can you also have a creeping greyscale movement from reactive environmental responses to internal representation? Is it all convincingly explained by a simple linear, smooth progression? Furthermore, does this figurative notion not imply that one can extrapolate to the right on the diagram and propose the evolution of even greater 'shades' of internal representation than currently exist within human mental experience?

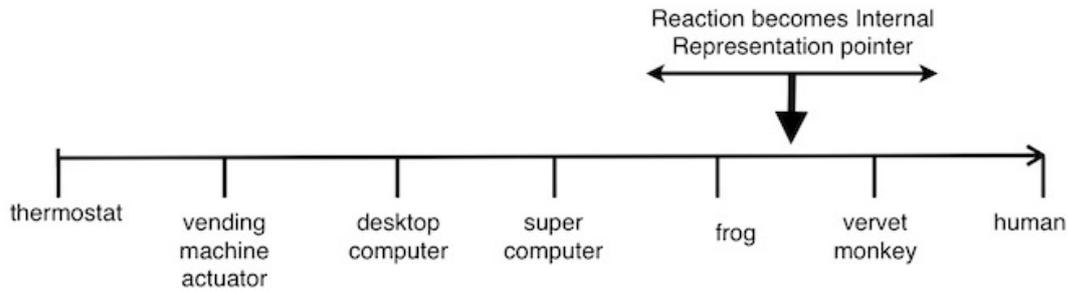


Figure 3: Point at which Reaction becomes an Internal Representation

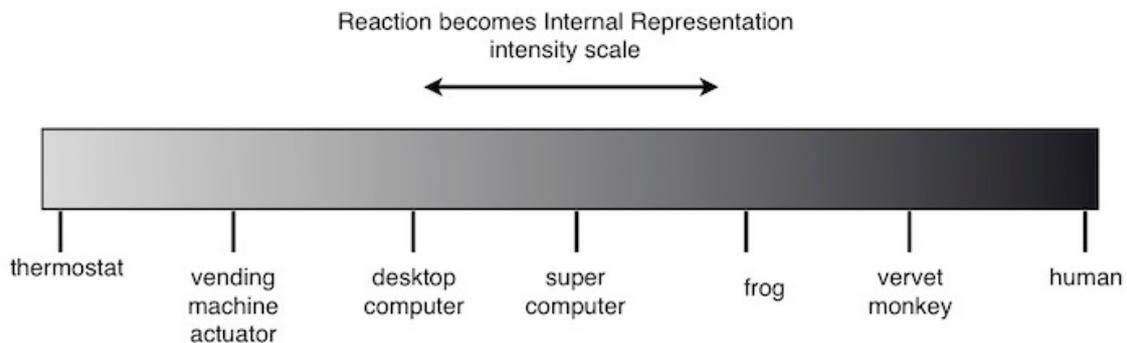


Figure 4: Reaction becomes Internal Representation Gradient

Additionally,

Note 6.3: in the phrase “transition from a simple thermostat to a ‘system’ “, Dennett once again has borrowed the term “system” (cf. note 1.1 above). The inference is that the term ‘any object’ (as in, “any object – or as I shall say, any system”) is not sufficiently explicit for Dennett and that there is some distinction between ‘object’ and ‘system’. However, in the quote from p.32 (note 6.), the indication is that the object under examination, namely the thermostat, has *transitioned from* a “simple object” “to a ‘system’ “. (This is implicit in Dennett’s use of single quotation marks surrounding the term ‘system’). This may sound petty. Nevertheless, it is a point that relates to the issues raised with intentionality and representation that needs pursuing: When can one say of an object that it is a system (c.f. figures 5 and 6 below)? Is a thermostat a system or not? Does the intentional stance apply only to systems and not to objects? I shall return to these questions later.

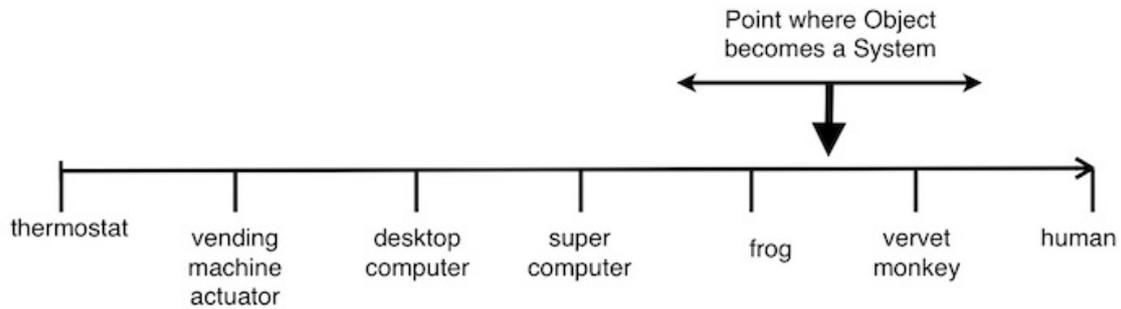


Figure 5: Point at which an Object becomes a System

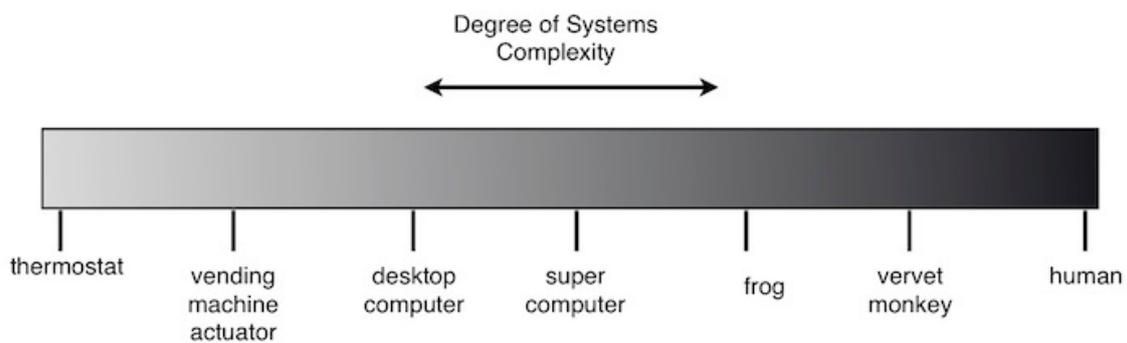


Figure 6: Object/Systems Complexity Gradient

7. Dennett's general appeal: The reason for making these analyses and for highlighting with figurative illustrations is to help identify any conceptual 'leaps of faith' that might indicate false claims, or to identify the misappropriation of concepts by a 'creeping augmentation' or 'creeping deviation' of meaning or definition:

Running throughout Dennett's writing is the idea that *artefact complexity*—in terms of sensory, evaluative and environmentally representative facilities—is sufficient a parallel to make with the evolved complexity of organisms: Dennett is very keen to draw analogy between the potential in the appropriate design of artifacts and the evolution of organisms. The purpose of this analogy is to promulgate the thesis that there is nothing special about humankind and there is nothing mysteriously unique about human intentionality—intentionality can be made to grow artificially. Dennett's appeal falls on the generalised acceptance in the validity of the ill-defined 'concept' of *complexity* in 'systems'. This concept regarding the complexity of systems is further illustrated in this transcription from a podcast interview about the 'Chinese Room Argument' between Nigel Warburton and Daniel Dennett for 'PhilosophyBites.com' (*reproduced with kind permission from Nigel Warburton, starting at 9 minutes 14 seconds and ending at 11 minutes 10 seconds*):

Dennett: Imagine the capital letter “D”. Now turn it 90 degrees counter clockwise. Now perch *that* on top of a letter “J”. What kind of weather does that remind you of?

Nigel Warburton: The weather today—raining.

Dennett: That’s right; it’s an umbrella. Now notice that the way you did that is by forming a mental image. You know that, coz you are actually manipulating these mental images. Now, that would be a perfectly legitimate question to ask—in the Chinese Room scenario—and if Searle (in the back-room) actually followed the program, without his knowing it, the program would be going through those exercises of imagination: it would be manipulating mental images. He would be none the wiser coz he’s down there in the CPU opening and closing registers, so he would be completely clueless about the actual structure of the system that was doing the work.

Now, *everybody* in computer science, with few exceptions, they understand this because they understand how computers work, and they realise that the understanding isn’t in the CPU, it’s in the system, it’s in the software—that’s where all the competence, all the understanding lies, and Searle has told us that, that reply to his argument—the systems reply—he can’t even take it seriously; which just shows: if he can’t take that seriously, then he just doesn’t understand about computers at all.

Nigel Warburton: But isn’t Searle’s point not that, such a computer can be competent because the one with him as components of the system is competent, but that it wouldn’t genuinely understand.

Dennett: Well it passes all behavioural tests for understanding: it forms images; it uses those images to generate novel replies; it does everything in its system that a human being does in his or her mind. Why isn’t that understanding?

philosophy bites transcription excerpt: http://hwcdn.libsyn.com/p/d/3/d/d3d403ebf90c10d2/Daniel_Dennett_on_the_Chinese_Room.mp3?c_id=5799487&expiration=1392565243&hwt=60f6cc410372224a6887cd99cda754d1

8. Complexity versus organisation.

There is a familiar way of alluding to this tight relationship that can exist between the organization of a system and its environment: you say that the organism continuously *mirrors* the environment, ‘or that there is a *representation* of the environment in—or implicit in—the ‘organization’ of the system. (p.31 *TIS*)

Note 8.1: Perhaps Dennett’s sentence “there is a *representation* of the environment in—or implicit in—the ‘organization’ of the system.” gives a vital

clue: perhaps the crucial issue is not concerning 'degrees of complexity' but in the very nature of the *organisation* of systems. With figures 5 and 6 above, there is a notion that degrees of complexity in systems is relevant to questions concerning intentionality. However, from atoms to nebulae, or from bacteria to adult human, the list of potential objects and systems that could be classified as complex is endless and this 'complexity' appears to have minimal bearing on degrees of intentionality. Superficially, therefore, it would appear that it is the nature of systems *organisation* rather than systems *complexity* per se, that is relevant to intentionality and levels of representation.

9. More thoughts about organisation

Thus (1) the blind trial and error of Darwinian selection creates (2) organisms whose blind trial and error behavior is subjected to selection by reinforcement, creating (3) "learned" behaviors that generate a profusion of (4) learning opportunities from which (5) the most telling can be "blindly" but reliably selected, creating (6) a better-focused capacity to generate (7) further candidates for not-so-blind "consideration," and (8) the eventual selection or choice or decision of a course of action "based on" those considerations. Eventually, the overpowering "illusion" is created that the system is actually responding directly to meanings. It becomes a more and more reliable mimic of the perfect semantic engine (the entity that hears Reason's voice directly), because it was designed to be capable of improving itself in this regard; it was designed to be indefinitely self-redesigning. (p.30 *TIS*)

Note 9.1: Dennett indicates that there are various identifiable *levels* of organisation from organisms that replicate, to plants and animals that display only innate behaviours, to animals that are capable of learning, and to humans that can explore thoughts creatively. Additionally, Dennett expresses the view that humans have some organisational capacity that grants uniquely 'privileged access':

For if I am right, there are really two sorts of phenomena being confusedly alluded to by folk-psychological talk about beliefs: the verbally infected but only problematically and derivatively contentful states of language-users ("opinions") and the deeper states of what one might call animal belief (frogs and dogs have beliefs, but no opinions)." (p.233 *TIS*)

Does this suggest that systems organisation takes discrete evolutionary steps? If so, how might this idea impact on our interpretation of 'degrees' of intentionality and representational content?

10. How does the organisation of systems relate to representation, intentionality, and to object?

It is obvious, let us grant, that computers *somehow* represent things by using internal representations of those things. (p.215 TIS)

For an artefact to have an internal representation of the environment is for it to possess information—in the most general terms—about the environment. Both artefacts and organisms must possess some form of ‘information-construct’ to enable them to respond with accuracy to environment conditions. But what is the difference in the construction and organisation of information between the multitude of evolved organisms, and between organisms and artifacts?

A computer can be designed to ‘mirror’ the environment—theoretically to an infinitely accurate and responsive degree. In doing so, a computer must possess an impressive sensory and evaluative information capability, and organise and structure the information in very complex ways. From the third-person perspective, such a computer is clearly going to be capable of appearing *as if* it possesses an internal representation of the environment and appearing *as if* it possesses intentionality. Is this third-person ‘evidence’ sufficient to insist that there is no additional defining first-person narrative indicating a distinctive, either intrinsic, or genuinely derived intentionality?

A robot’s aggregated information can reflect the environment and enable suitably complex responses. When it goes through the process of evaluating a favourable environmental condition, it will respond to the favourable environmental condition *as if* it felt its goodness, but would this apparent goodness merely demonstrate the third-person consequence of a functional mechanism—a mechanism designed to identify favourable conditions and respond appropriately as per design function? What is the ‘reward’ that a computer ‘*experiences*’ from which it might learn, by means of *actually* feeling what is good or bad about the environment? Dennett is of the opinion that, “there are Artificial Intelligence (AI) programs today that model in considerable depth the organisational structure of systems that must plan “communicative” interactions with other systems, based on their “knowledge” about what they themselves “know” and “don’t know,” what their interlocutor system “knows” and “doesn’t know,” and so forth.” (p.39 ER)

When can one say that information becomes knowledge? And what ‘knowledge’—as opposed to information—can a complex aggregate of parts possess? Can an *extremely complex* array of falling dominos be called a ‘system’—a system that possesses ‘knowledge’ about what and why the first domino fell—or is the series of complex activities merely causally functional operations? A domino’s purpose for the implementation of function, is to fall; but can a series of dominos that fall be called a true systems-construct that possesses knowledge about the environment to which its individual dominos might react?

11. What is a true systems-construct? Note from Dennett's early statement of intent that he makes the assumption that there is a distinction between 'a *system*' and 'an *object*' but provides no explanation of what such distinction entails, other than mere 'complexity':

I will argue that any object—or as I shall say, any *system*—whose behavior is well predicted by this strategy is in the fullest sense of the word a believer. (*p.15 TIS*)

When can an object, however complex, be called a system? Who is to say that a super (duper) computer or extremely advanced robot, is not merely a complex object, such that the nature of its functional organisation remains less than primitive despite its functional acuity and reactive—even adaptive—sophistication?

A true system is a construct of dynamic constituent parts. It does not organise information by virtue of its function. Rather, the very dynamic of its parts is the information construct in and of itself, whereby data is not 'read' from the environment to construct information but rather the construct is, by virtue of interactive engagements, intrinsically an information construct. Instead of the components *containing*, *representing* or *organising* information sets, the entire dynamism of the construct itself dictates the nature of its response to environmental conditions and it is this that qualifies the nature of its informed representative character.

Information, which an object might be said to possess about the environment and which might be said to 'represent the environment', leads to intentionality of purpose if, and only if, the informedness is represented through the dynamics that determine the nature of its construction. Thus, it is the interaction of the dynamic parts of a system that determine the nature of its informedness and henceforth behaviour. A true systems-construct is itself the representation of the environment by virtue of the interaction of its dynamic components: any representation is implicit in the nature of its dynamic construct.

12. Of the baby and the bathwater. To the left, we have the baby. To the right we have the bathwater.

On the left, is the baby: I like Dennett's intuitive call, insisting that we relate the simplest living organisms with sophisticated animals like humans i.e., that we assume there is an evolutionary progression that binds the simplest of replicating organisms with 'complex' humans; there is an explanation to be had.

Dennett hates the idea that mankind should be so "arrogant" as to proclaim that there is a dividing line; a demarcation point where one can say, "this human has intentionality whilst that animal does not or, heaven forbid, this human has intentionality whilst that mentally retarded human does not." This evolutionary connection is the baby I don't want to throw out i.e., I like Dennett's position that

says, the privileged access that humans possess (he says, ‘appear’ to possess) through their exceptional mental capabilities (he says, through their ‘apparent’ intentionality), is not “magical”. I agree, there is a connection to be discovered—there is no magic.

On the right, we have the bathwater: Dennett is very keen to articulate the view that an evolutionary analogy of very simple objects and/or artefacts to very complex objects and/or artefacts have a direct correlative relation or a parallel relation to the evolution of simple to sophisticated organisms.

The flaw in Dennett’s parallel analogy rests on this one sentence of his:

“...any object – or as I shall say, any system...”

The error is in the following false proclamations:

object = system

or

object + complexity = system

From these alternative proclamations Dennett makes the false conclusion: ‘objects or systems (who cares which) have ‘as if’ or ‘actual’ intentionality (who cares which) i.e. they can be treated as analogous.

This is extended further by Dennett as follows: ‘there is no true ‘as if’ or ‘actual’ intentionality distinction.’ Alternatively, I say the following about systems and objects:

A (true) system (regardless of its simplicity or complexity) has ‘intrinsic’ purpose. A system’s dynamically interactive components create a stable “body” (or, to borrow Dennett’s terminology, “object”) and determine the body’s functional characteristics and behavioural properties.

An object that is not a true system, on the other hand, may either consist of a ‘mere’ aggregation of component parts or may not be a product itself of dynamic interaction. Parts can be said to be an aggregate when they do not interact dynamically in such a way as to determine and define the object’s behavioural properties or the object’s stability. An aggregation of parts can make a representation about the environment, will react to the environment and will display behavioural characteristics, but the dynamics of the parts, through their interaction, are not determinate of the stability of the whole.

Therefore, throwing out the bathwater allows us to express the following:

1. a thermostat, a computer, a robot, a tube of toothpaste, a machine are examples of aggregated constructs. They are not systems-constructs.
2. an aggregated construct cannot have ‘intrinsic’ purpose or ‘intrinsic’ intentionality.

3. a body (or an object if you take Dennett's terminology) has 'intrinsic' purpose if the very dynamic interaction of its component parts determine its behavioural and functional properties.
4. computers regardless of their theoretical complexity, designed as they are today, have and would have zero intrinsic intentionality regardless of the third-person's perspective of their apparent 'as if' intentionality and adaptive imitative capabilities.
5. all systems regardless of simplicity have intrinsic intentionality, be they an atom or human consciousness.

How can the aggregated information sets of computers be transformed, by intrinsic intentionality, into knowledge, into desire, into internalised purpose? To answer this question, is to understand the nature of the hierarchy that arises through the emergence and evolution of self-organised types of systems-constructs. And so we can move forward, from the simplicity of our linear figured illustrations above, to a hierarchy of types of systems-constructs that evolve different types of form whose information about the environment varies widely thereby creating internal representations particular to their type, from the humble atom all the way up the hierarchy to human consciousness. c.f. figure 7 below:

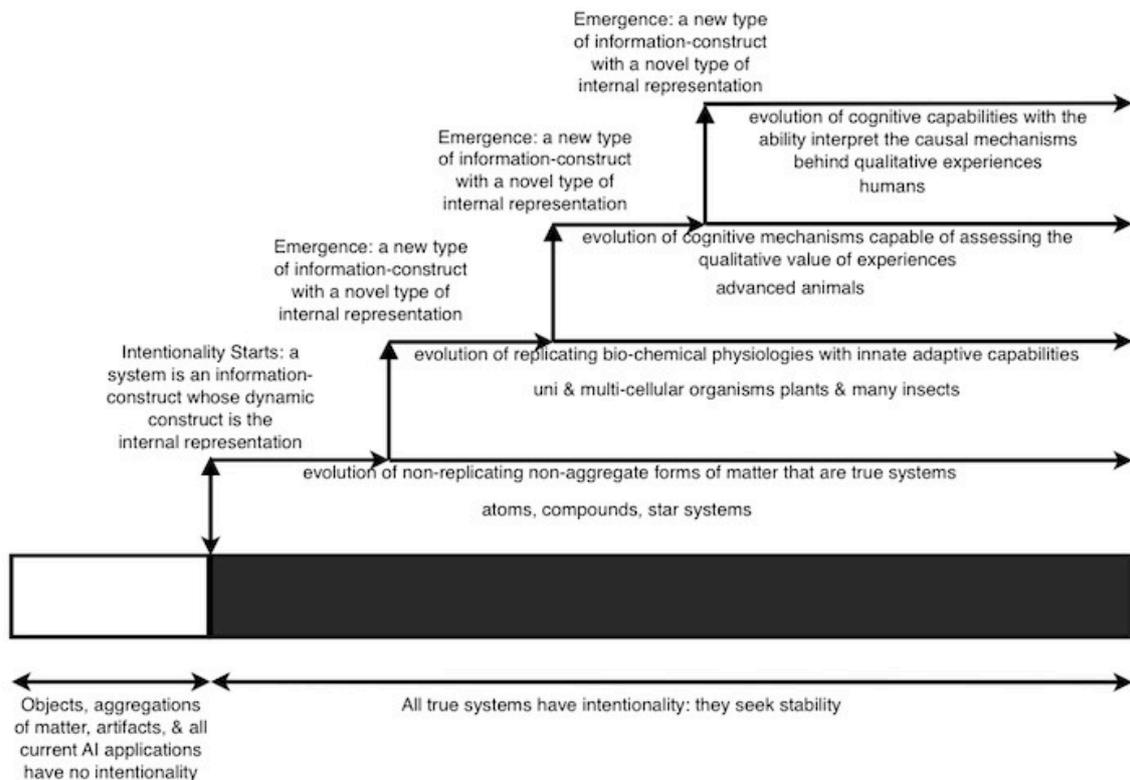


Figure 7: Hierarchical quantitative emergent steps with analogue evolutionary developments

REFERENCES

- Dennett (1971) *Intentional Systems*. The Journal of Philosophy, Vol 68:4 pp. 87-106
- Dennett (1981) *True Believers: The Intentional Strategy and Why It Works*. In A.F. Heath, ed., Scientific Explanation. Papers Based on Herbert Spencer Lectures Given at Oxford University. Oxford: Oxford University Press pp. 150-167 (also in *The Intentional Stance* i.e. Chapter 2)
- Dennett (1987) *The Intentional Stance*. Cambridge, MA: MIT Press
- Dennett (2009) *Intentional Systems Theory*. The Oxford Handbook of Philosophy of Mind
- Dennett (1984) *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge, MA: MIT Press
- Warburton (2013) Daniel Dennett on the Chinese Room. Philosophy Bites podcast interview