Representation in Digital Systems

Vincent C. MÜLLER[1]
American College of Thessaloniki

Abstract. Cognition is commonly taken to be computational manipulation of representations. These representations are assumed to be digital, but it is not usually specified what that means and what relevance it has for the theory. I propose a specification for being a digital state in a digital system, especially a digital computational system. The specification shows that identification of digital states requires functional directedness, either for someone or for the system of which the state is a part. In the case of digital representations, the function of the type is to represent, that of the token just to be a token of that representational type.

1. Digital Representations and the Computationalist Program

In this paper, I will attempt to clarify one aspect of a notion that is commonly used in the cognitive sciences, but has come under considerable criticism in recent years: the notion of representation. Representations are typically invoked in a 'computational theory of the mind', where mental processes are understood as information processing through computational operations over *representations.* The standard theory is the computational representational theory of mind (CRM or "computationalism"), which says that the human mind is a functional computational mechanism operating over representations. These representational abilities are then to be explained naturalistically, either as a result of information-theoretical processes [1,2], or as the result of biological function in a "teleosemantics" [3,4].

While there is intense discussion about what role representations play and whether their manipulation must be computational, there is one aspect that is commonly glossed over: The computational processes in question are taken to be *digital* computational processes, operating algorithms over *digital* representations. But what constitutes a digital computation and a digital representation? And if this were specified, does the specification have repercussions for the computational representational theory of the mind?

Such repercussions are to be expected in several areas that are crucial for a computational theory of the mind. One of these is the question whether something can be called a digital state at all without *presupposing* mental processes – in which case there is a threat of a circle. Another is the problem of "grounding". This alleged problem is intimately connected with the current wave of 'embodied cognition'. A concise formulation probably still is: "How can the meanings of the meaningless symbol tokens, manipulated solely on the basis of their (arbitrary) shapes, be grounded in anything but other meaningless symbols?" [5]. Now, we have argued in recent papers on nonconceptual phenomenal content [6,7] that a non-conceptual content should be at the base of such grounding: Such content is retrieved bottom-up (to "cognitively encapsulated" modules), it is not mediated top-down by concepts, is independent of conceptual resources available to the person, and not phenomenal (accessible to the person). The nonconceptual representations of objects can provide a starting point for the grounding procedure that does not assume conceptual material (but rather things like spatiotemporal information of objects such as existence, persistence in time and through motion, spatial relations relative to other objects, movement; basic information on surface properties, shape, size and orientation). If it were to come out that such content is necessary but cannot be present in purely digital systems, this would show that human cognition is not purely digital – and that artificial intelligence on purely digital computers is impossible.

---
[1] Corresponding Author: Vincent C. Müller: American College of Thessaloniki P.O. Box 21021, 55510 Pylaia, Greece, http://www.typos.de, E-mail: vmueller@act.edu

We can take our starting point from the recent debate whether mental representations are material [8,9], or whether thinking perhaps proceeds via conventional linguistic symbols rather than especially 'mental' symbols on a sub-personal level [10]. Sedivy, in particular, argues that cognition does not proceed as a manipulation of 'carriers' of representation: In the case of linguistic symbols, for example, there is a physical object (ink on a page, compressions of air, etc.) that carries content, but in the case of mental representation at the personal level there is not. Could digital states be such carriers?

The question of digital representation is also relevant for the general question which physical objects in the world are computers; a problem that Shagrir calls "problem of physical computation" [11]. Digital computation necessarily involves digital representation on the classical view: "There is no computation without representation" [12]. O'Brien even defines computing with the help of representation: Computation is "a procedure in which representational vehicles are *processed in a semantically coherent fashion*" [13] (a "coherent fashion" being one in which they "bear comprehensible [non-arbitrary] semantic relations to one another"). This view is opposed to the ones that dispense with representation, either understanding computation as a purely syntactic procedure, or widening the notion to the extent that "every natural process is computation in a computing universe" [14]; a view that is now called "pancomputationalism".

2. Digital Representations and Digital Content

If a digital state is part of a system in which it plays a representational role, then it will have content. This content, in turn, is necessarily digital as well since a digital state can only represent whether something is *of a type*. This applies whether or not the system is a cognitive system or not. (The red warning light on the dashboard that informs me of the engine overheating is part of a system, indeed a computational information processing system, but presumably not of a cognitive one.) The characteristic feature of a digital representation in a system is thus that it 'checks' whether an informational input is of a type or not, it *categorizes* the input. So, a digital representation can only represent a digital content; an analog content can only be approximated. For example, it is impossible to describe completely in words what a picture depicts; as the saying goes "a picture is worth a thousand words", i.e. an analog representation can be categorized in a thousand ways, and yet none will be sufficient. Any indication to the contrary is produced by that fact that one cannot *say* with words what is the analog content that is not represented; one can just *indicate* that there is such content by pointing out what is missing, in each representation.

Dietrich and Markman summarize their discussion with the statement "If a system *categorizes* environmental inputs, then it has discrete [digital] representations" and "A system has discrete representations if and only if it can discriminate its inputs." [15] This is appropriate, but it *presupposes* that the digital states in question are representations. What we need to find out is whether this is accidental, or if *all* digital states are representations?

The notions of representation and of representational content are, of course, disputed, but it seems basic to distinguish (in the tradition of C. S. Peirce) *symbolic* and *iconic* representation from *indices* (e.g. smoke indicates the existence of fire). I would propose to capture this difference by calling the former representation and the latter information.

On this terminology, information is whatever can be learned from the causal history of an event or object, representation is what it is *meant to represent* (in a suitable notion of function). It is clear that in this usage, information is always true, while a representation may not be. (So, if someone lies to you, he is not giving you false information, he is misrepresenting the world.)

Nonetheless, we can still identify (at least) two notions of representation. One typical version of a wider notion is: "a representation is any internal state that mediates or plays a mediating role between the system's inputs and outputs in virtue of that state's [explicit] semantic content" [15]. On the other hand, there is a use of the term in which symbols and (more or less interpretation-involving) icons are representations for a person. As I said in a different context:

> "…we need to distinguish between a representation *per se*, and a representation *for someone*. If we are wondering whether, say, a squiggle on a piece of paper is a representation, this is (normally) to ask whether it is a representation for someone. If we are wondering whether a set of switches or a neural pattern represents, this is to ask whether it has the function to represent in a larger system." [16]

Note that, in the end, both notions involve functional talk: being a representation for someone and being a representation for the system. I will argue presently that this is characteristic of digital states.

There is a common view that couples representation quite generally with digital processing, e.g. "talk of representations [in the analysis of cognitive functioning] invites speculation about content and reference which leads to a symbolic model of cognitive processing, or at least blurs the distinction between analog and symbolic simulation," [17]. As we shall see shortly, even if there are digital representations, there are also others, probably even in cognitive systems.

2.1. Discreteness vs. continuity

In a first approximation, being digital means being in a discrete state, a state that is strictly separated from another, not on a continuum. Prime examples of digital representations are the states of a digital speedometer or watch (with numbers as opposed to an analog hand moving over a dial), the digital states in a conventional computer, the states of a warning light, or the states in a game of chess. Some digital states are binary, they have only two possible states, but some have many more discrete states, such as the 10 numbers of a digital counter or the 26 letters of the standard English alphabet.

It is characteristic of such digital representations that they can have multiple realizations. So, one can write the same word twice, even though one cannot make exactly the same mark on paper twice. The possibility of multiple realization is a *result* of digital states being discrete: Since a mark on a piece of paper can be clearly a "T", we can rub it out and replace it with a new "T", or copy the "T" to a new piece of paper. It does not matter that there are borderline cases which are neither clearly this letter nor clearly another, as long as there are clear cases. (This becomes easier if one knows *a priori* that something is meant to be a letter, i.e. a digital representation from a particular set.)

Crucially, a digital representation is a *token of a type*. Being such a token or not is what allows it to be discrete. Being such a token allows it to be realized in multiple ways and several times. *Which* types exist is often pre-defined, as in the case of the 26 letters of an alphabet. Something either constitutes a token of one of these types or it is not a letter of this alphabet at all.

2.2. Analog representations

Digital representations are, to characterize them negatively, not analog representations. But what does that mean? Trenholme says that "... analog simulators are physical systems whose processes are characterized by what might be termed *variable properties* - physical magnitudes whose values vary over time." [17], meaning that they vary over time in a non-discrete way, *and* in analogy to the represented. For example, the movement of the hand over the speedometer is analogous to the speed of the vehicle – while the digital numbers on a display are not in such an analogy.

Fodor points out that the analogy is a product of physical laws, and that it takes the absence of such laws to indicate a digital representation: "A machine is analog if its input-output behavior instantiates a physical law and is digital otherwise" [18]. Another formulation is that an analog process is "one whose behavior must be characterized in terms of lawful relations among properties of a particular physical instantiation of a process, rather than in terms of rules or representations (or algorithms)." [19] This characterization is used by Demopoulos who defines a digital machine class (a class of machines with the same digital states) by saying that their "behavior is capturable only in computational terms" [20], i.e. not by physical laws.

These remarks are very plausible, if we consider that in the case of a continuous representation (e.g. the analog speedometer) the relation between what is represented and what represents is wholly determined by the physical laws governing the two and their "correspondence" [15]. It is only when digital states come into play that the continuous input is 'chunked' into types. This division into tokens of types in a digital system excludes its description purely in terms of physical laws precisely because it also requires a reference to function or directedness (it is not accidental that the rejection of psychophysical "identity theories" was the starting point of functionalism in the philosophy of mind).

But which of the two characteristics is crucial for an analog state, the analogy to the represented, or the continuous movement?

This question becomes relevant in the case of representations that proceed in "steps" but also in analogy to the represented, e.g. a clock the hands of which jump from one discrete state to another. Zenon Pylyshyn argues that the underlying process is analog, and this is what matters: "an analog watch does not cease to be analog even if its hands move in discrete steps" [21,22]. James Blachowicz also thinks that being on a continuum is sufficient for being analog, taking the view that "differentiated representations may also be analog – as long as they remain *serial*", his example is a slide rule with "clicks" for positions [23].

Note, however, that the very same underlying mechanism could give a signal to a hand to move one step and to a digit to go one up (this is actually how clocks are controlled in centralized systems, e. g. at railway stations). Both of these would be analogous to the "flow of time", given that the natural numbers of the digital display are also in a series. So, while the underlying mechanism might be analog or digital, the question whether the representation is digital ultimately hinges on whether it is discrete, not on whether it is analogous to the represented. From what we have seen thus far, a digital representation is just a discrete representation.

2.3. Digital and analog - on which level?

Returning to the classical computational representational theory of the mind, it is important for our analysis to see that a computer can be described on several levels and identify which one is relevant for us here.

First of all, a computer can be described on the *physical level:* Some physical objects such as toothed wheels, holes in cards, states of switches, states of transistors, states of neurons, etc. are causally connected with each other such that a state of one object can alter the state of another. So far, this is just any system. At the s*yntactical level,* the states or physical objects are taken to be tokens of a type (e.g. charge/no charge) and are manipulated according to algorithms. At first glance, this manipulation only concerns these tokens; it is "purely syntactical." Finally, there is a s*ymbolic level* of what the objects and states that are manipulated on the syntactical level are taken to represent, e.g. objects in the world. Perhaps it would be more appropriate to say that there are several symbolic levels, since one symbol can represent another, that can represent, say, a color, that represents, in turn, a political party, etc.

It so happens, for technical reasons, that the types that we humans most frequently use to represent (e.g. natural numbers, letters, words) are normally not the types that are syntactically manipulated in a conventional computer, but are rather represented, in turn, in a further system of types, a "binary system" with tokens of just two types (on/off, 1/0, etc.). At *both* these levels, we have digital representation. [In the slogan "There are exactly 10 kinds of people in the world, those that understand binary and those that do not", the "10" is a binary sequence of two bits, representing the number two.]

3. Digital States as Tokens of a Functional Type

It is useful to note that not all systems that have digital states are digital *systems*. We can, for example, consider the male and female humans entering and leaving a building as digital states, even as a binary input and output, but in a typical building these humans do not constitute a digital system because a relevant causal interaction is missing. In the typical digital system, there will thus be a *digital mechanism*, i.e. a causal system with a *purpose*, with parts that have *functions*. Digital mechanisms in this sense may be artifacts (computing machines) or natural objects (perhaps the human nervous system). However, even if all digital representations are part of digital systems, they are not all part of computational systems – the letters and words on this page are an example in point.

We need the notion of a system because the notion of "being of a type" is too broad for our purposes. If being of a type were the criterion for being digital, then everything would be in any number of digital states, depending on how it is described. However, what we really should say is that something is digital because that is its particular *function*. The first letter of this sentence is in the digital state of being a "T" because that is its function – as opposed to an accidental orientation of ink or black pixels (or Putnam's ants making apparently meaningful traces in the sand [24]).

Some of the systems are artifacts, made for our purposes, where some physical states cause other physical states such that these function as physical states of the same set of types (e.g. 1 or 0). (Note that one machine might produce binary states in several different physical ways, e.g. as voltage levels and as magnetic fields.) If someone would fail to recognize that my laptop computer has binary digital states, they would have failed to recognize the proper (non-accidental) function of these states for the purpose of the whole system – namely what it was made for.

So, the function is what determines whether something is a token of a type or not. The normativity of having or fulfilling a function generates the normativity of being of a type. The type has the function; being of the type allows fulfilling the function.

3.1. Which function?

In the case of an artifact, we *assume* a functional description. If the engine warning light on a car dashboard is off, is it in a digital state? Yes, if its function is to indicate that nothing is wrong with the engine temperature. (It may serve all sorts of other accidental functions for certain people, of course.) But if the light has no electricity (the ignition is off), or if it was put there as a decorative item, then the lamp is not in a digital state "off". It would still be off, but this state would not be digital, would not be a token of the same functional kind.

The description of an artifact in terms of function is to say that something is a means to an end; it serves the function to achieve that end – a function that can be served more or less well. (Note that serving a function does not mean being used for that function; there may well be no agent that can properly be said to be using the artifact, e.g. if it is part of a large and complex system.) In the case of a natural object, the allocation of proper function (as opposed to "accidental function") is dependent on teleological and normative description of systems [cf. 25] – a problematic but commonplace notion.

The interesting cases for our understanding of digital states are those where it is not clear that some object or state really is a digital representation. The prehistoric cave paintings contain repeated patterns, so are these symbols? Are what seem to be icons actually digital representations, as it turned out in the case of the Egyptian hieroglyphs? And what about the neural activity in the brains of humans and other mammals? The function of a human's brain seems to be cognition, so whether the brain is in digital states depends on whether its cognitional function is fulfilled in this way. Are there *types* of neural states that serve a representational function? (Piccinini has recently presented interesting evidence that this is not the case [26].)

3.2. Representational function

The problem of identifying a function could perhaps be solved if we helped ourselves to representation as the function. (I am grateful to Philip Brey for a question in this direction at E-CAP 2007.) Now, I am *not* of the view that we should restrict the notion of digital states only to the representational ones because I think that digital states in several characteristic areas, in particular in conventional digital computers are *not* representational.

Nonetheless, in the cases where the digital states *do* serve representational function, it will be precisely *that* function that determines being or not being of a type. So, as I said above, when we wonder whether something is a letter, the answer depends on whether it serves a representational function as a token of a particular (finite) set of types. In that case, just being of the type is to fulfill the function. So, if a state has that function *qua* token of a representational type, then it is a digital representation.

This understanding should not be taken as a motivation to re-introduce semantical notions into the definition of digital states or computing. It has been argued in many places that a computational mechanism must involve semantics or meaningful symbols, because this is necessary for its supposed carrying out of orders, or for its identification of tokens (e.g.[27,28] and [29,30]). Kuczynski even argues that there is really no such thing as a purely formal procedure because semantics is needed to identify tokens, which leads him to say, in the end, that: "Since the Turing machine is not semantics-driven, it is not responding to logical form. Therefore it is not computing (in the relevant sense), even though, from some viewpoint, it acts like something that is computing" [31].

These consequences do not follow if we recall that digital states are situated at the level of syntactic description. The level of syntactic description can be identified with the help of representational function without thereby saying that there is somehow a semantic or representational level in the machine. The Turing machine computes, but its computations mean nothing to it; it is only to us that they are, for example, operations over natural numbers.

Acknowledgements

References

[1] F. Dretske, *Knowledge and the flow of information*, MIT Press, Cambridge, Mass., 1981.

[2] F. Dretske, *Naturalizing the mind*, MIT Press, Cambridge, Mass., 1995.

[3] R.G. Millikan, *Language: A biological model*, Oxford University Press, Oxford, 2005.

[4] G. Macdonald, D. Papineau, editors, *Teleosemantics: New philosophical essays*. Oxford University Pres, Oxford, 2006.

[5] S. Harnad, The symbol grounding problem, *Physica D* 42 (1990), 335-346.

[6] A. Raftopoulos, V.C. Müller, The phenomenal content of experience, *Mind and Language* 21 (2006), 187-219.

[7] A. Raftopoulos, V.C. Müller, Nonconceptual demonstrative reference, *Philosophy and Phenomenological Research* 72 (2006), 251-285.

[8] A. Clark, Material symbols, *Philosophical Psychology* 19 (2006), 291-307.

[9] S. Sedivy, Minds: Contents without vehicles, *Philosophical Psychology* 17 (2004), 149-179.

[10] J. Speaks, Is mental content prior to linguistic meaning?, *Nous* 40 (2006), 428-467.

[11] O. Shagrir, Why we view the brain as a computer, *Synthese* 153 (2006), 393-416.

[12] J.A. Fodor, The mind-body problem, *Scientific American* 244 (1981), 114-123.

[13] G. O'Brien, Connectionism, analogicity and mental content, *Acta Analytica* 22 (1998), 111-131.

[14] G. Dodig-Crnkovic, Epistemology naturalized: The info-computationalist approach, *APA Newsletter on Philosophy and Computers* 6 (2007), 9-14.

[15] E. Dietrich, A.B. Markman, Discrete thoughts: Why cognition must use discrete representations, *Mind and Language* 18 (2003), 95-119.

[16] V.C. Müller, Is there a future for AI without representation?, *Minds and Machines* 17 (2007), 101-115.

[17] R. Trenholme, Analog simulation, *Philosophy of Science* 61 (1994), 115-131.

[18] J.A. Fodor, N. Block, Cognitivism and the digital/analog distinction, unpublished, cited in [20].

[19] J.A. Fodor, Z. Pylyshyn, How direct is visual perception?, *Cognition* IX (1981), 139-196.

[20] W. Demopoulos, On some fundamental distinctions of computationalism, *Synthese* 70 (1987), 79-96.

[21] Z.W. Pylyshyn, *Computation and cognition*, MIT Press, Cambridge, Mass., 1984.

[22] O. Shagrir, Two dogmas of computationalism, *Minds and Machines* **7** (1997), 321-344.

[23] J. Blachowicz, Analog representation beyond mental imagery, *The Journal of Philosophy* 94 (1997), 55-84.

[24] H. Putnam, *Reason, truth and history*, Cambridge University Press, Cambridge, 1981.

[25] U. Krohs, Der Funktionsbegriff in der Biologie, in: A. Bartels, M. Stöckler, editors, *Wissenschaftstheorie: Texte zur Einführung*. Mentis, Paderborn, 2007, forthcoming.

[26] G. Piccinini. Digits, strings, and spikes: Empirical evidence against computationalism. NACAP Conference. Chicago; 2007.

[27] M.A. Boden, Escaping from the Chinese room, in: M.A. Boden, editor, *The philosophy of artificial intelligence*. Oxford University Press, Oxford, 1990, 89-104.

[28] M.A. Boden, *Mind as machine: A history of cognitive science*, Oxford University Press, Oxford, 2006.

[29] J. Haugeland, *Artificial intelligence: The very idea*, MIT Press, Cambridge, Mass., 1985.

[30] J. Haugeland, Syntax, semantics, physics, in: J. Preston, M. Bishop, editors, *Views into the Chinese room: New essays on Searle and artificial intelligence*. Oxford University Press, Oxford, 2002, 379-392.

[31] J.-M. Kuczynski, Two concepts of 'form' and the so-called computational theory of mind, *Philosophical Psychology* 19 (2006), 795-821.

[32] V.C. Müller, 2007, *Representation in digital systems (with comments and replies),* in Interdisciplines: Adaptation and representation, Institut des Sciences Cognitives/CNRS, Université de Genève <http://www.interdisciplines.org/adaptation/ papers/7>, accessed 11.07.2007.