Physical Laws and Information Content

C. Franklin Boyle

CDEC, Carnegie Mellon University Pittsburgh, PA 15213

Abstract

Current theories that incorporate the notion of information are not theories of information content, at least not content in the form of spatially-extended structures, such as words, computer programs and pictures. Such theories refer to content only indirectly through some formulation of information measure. It is argued here that one reason for the absence of theories of information content is that physical laws, along with situation-specific constraints, describe associations between measured attributes, while content is, in general, spatially extended and, more importantly, causal as structure. It is inferred from this that physical state descriptions are only powerful enough to account for content when it is instantiated as the values of measured attributes of physical objects. For content in the form of spatially-extended structures, there are no such physicallygrounded formalisms, only concrete physical systems (e.g., digital computers and brains) whose architectures enable spatially-extended structures to cause change.

1: Information content

The word "information" is generally taken to refer to the content of a signal, message or thought, as in: "He has the information in his head about how to fix the transmission." or "The information about the climate of South America is on this disk." (these sentences, of course, also have content). It is therefore noteworthy that current theories which incorporate the notion of information do so through measure. Examples are the amount of disorder in a physical system [1], how much information must be given to a computer to calculate a particular binary string [2,3] and the information *capacity* of a given communications channel [4]. Even theories that focus on individual messages (as opposed to sets of messages) do so through measure, such as the probability that the content of a particular message has been received [5] or the "algorithmic information content" of a computer program [3]. In other words, the actual contents of messages, programs or thoughts — what those messages, programs or thoughts "contain" - are not part of such theories.

One might wonder why this is the case, considering that content seems to affect the behaviors of so-called "information processing systems", such as computers and brains, implying that a causal relationship exists between content and physical behavior, and, hence, that a physical theory of content could be developed. Part of the reason may be that content is associated with meaning, which is a highly non-physical concept — as Haugeland [6] cogently observes, "meanings do not exert mechanical forces." That is, the content of a signal or message is usually taken to be what we interpret the signal or message to be about, suggesting that its content depends on external agencies like ourselves; it is *derived* from the content of our mental states through interpretation. But this introduces a subjective (i.e., non-physical) component into any wouldbe theory of content because the process of interpreting a signal or message in order to determine its content, and, hence, its effect on a receiving system, is exactly the process that a physical theory of content must explain; the interaction of the signal or message with the system interpreting it. In other words, determining the effect of content on an information processing system depends. itself, on such an effect — our interpretation — which is another way of saying that meaning is non-physical. Attempting to explain it would only lead to an infinite regress.

In order to avoid the kinds of problems associated with interpretation and, hence, meaning, and still make progress toward a theory of content, we should focus our attention on the fact that the content of a signal or message, whether it is a bit string in a computer, a retinal image of the environment or this sentence, appears to be associated with physical *structure*, since each of these is a spatially-extended pattern. Furthermore, pattern differences are what physically distinguish different signals or messages. In other words, physically, content is not, in general, associated with some measurable property, such as the mass, momentum or energy of the objects that instantiate signals and messages, but, rather, depends on their spatially-extended structures.

We can thus view the effects of content as being due to the physical interaction of a signal's or message's structure with the structure of the system that receives it, rather than interpreting what the signal or message means and, then, based on what we interpret the receiver to already "know", deciding how the receiver will behave. For example, the effect of a particular, physically-instantiated string of bits input to a computer physically depends on the particular, physically-instantiated rules it has for matching that input (which are also strings of bits, instantiated as "high" and "low" voltages). Thus, we can eliminate semantical notions by assuming that the effects of content can be accounted for through the physical interactions of spatially-extended structures.

The obvious question, then, is: How do we physically analyze the interaction of message and receiver without eliminating spatially-extended structure (i.e., content), since it is usually the case that physical analyses are based on state variables, which are not spatially-extended structures? To answer this, I will first look at how a cornerstone of physical theories — physical laws — relate to physical structure in order to develop some idea of how the effect of content might be accounted for physically.

2: Physical laws and structure

Theories of the physical world, whether or not they incorporate the notion of information, depend on numerically-valued quantities in order to be able to describe a wide variety of phenomena according to universal physical laws which associate such quantities: measurable properties such as energy and momentum. Without numerical quantities and their capacity to represent the physical states of objects and systems of objects, it would be difficult to make accurate predictions about physical behavior or explain why certain physical behaviors and states obtain. When physical laws are applied to particular situations, boundary conditions that encode specific structural characteristics are formulated as additional, situation-specific constraints on the allowed behavior. Thus, the states of any physical system can be specified and predicted entirely in terms of measurable quantities what we generally refer to as "physical state description" formulations. This may further explain why current theories that incorporate information do so through information measure.

It is through situation-specific constraints that spatially-extended structures are incorporated into physical state descriptions. Such constraints, however, "fragment" these structures by representing their features in terms of structureless position coordinates or distances; e.g., the distance of closest approach in a collision. Moreover, only those features which actually affect behavior --- that is, which restrict the values of the properties of the objects involved - are included. Thus, spatially-extended structure is only implicit in physical state descriptions. But, then, how else would physical structure be amenable to the mathematical formalisms of physical theories if it was not represented in terms of numerically-valued quantities? Though physical laws are universal because they do not depend on the specific characteristics of particular objects and their structures, situation-specific constraints must be specifiable in terms of numericallyvalued quantities if those laws are to be successfully applied in particular situations.

We can infer from this that if content is identified with structure and the effects of content with structure acting as structure (how it acts as structure will be discussed in the next two sections), then physical laws along with situation-specific constraints do not describe content or its effects, except in the special case in which content is taken to be the values of measured attributes of objects; i.e., in analog computers. In other words, physical laws are not specified in terms of spatially-extended entities, such as physical objects, only in terms of structureless measured attributes. This is generally not a problem because in most physical situations structure is not informational — it does not serve as content for an information processing system, and, therefore, does not have to be treated explicitly.

In analog computers, measured attributes of component parts are taken to represent physical quantities from physical or mathematical models of phenomena that the computers were designed to implement; an interaction between components instantiates a mathematical relationship between their informationally-relevant measured attributes which is realized through the shaping and juxtaposition of those components, as in "differential analyzers" developed to solve differential equations [7]. The actual structures of the components are peripheral to the information processing in that they are not content. They are only supportive of what might best be characterized as "measured-attribute information processing".

3: Measured attributes, spatiallyextended structure and causality

Though analog computers reflect our engineering prowess, they do little to solve the problem of physically grounding information content in general; we are still faced with the problem of physically accounting for the effects of content in other information processing systems, such as digital computers, artificial neural networks and. ultimately, the brain. That it is a problem is evidenced by the fact that descriptions of digital computers are of two kinds: 1) standard physical descriptions of their electronic behavior, which, by definition, are physically grounded, and 2) symbolic descriptions in terms of symbols and symbol structures, which function according to "laws embodied in the computer program, which can be of any consistent form" [8], and which are without physical properties. However, these latter, logico-semantic "laws", or rules, which describe computational regularities among all the physical changes that take place in a computer, are a consequence of the causal characteristics of the electrical structures (combinations of high and low voltages) that instantiate symbols. These characteristics do not depend on measurable properties, like they do for the components of analog computers, but, rather, on their particular structures.¹ That is, the effects depend on the particular arrangement of values rather than the values themselves.

The physical process enabling the combinations of voltage values in computers to be causal we generally refer to as "pattern matching". Pattern matching involves the "fitting" of the structures (exemplified by two jigsaw puzzle pieces) of two interacting objects which "triggers" a change that depends on that particular fitting event rather than measured attributes of the objects. Though the interaction itself is constrained by physical laws, just as the behavior of any physical object is constrained by physical laws, the outcome is actually due to the fact that the structure of one of the objects — the pattern — has the same (or complementary) arrangement of contours (or in the case of digital computers, the same arrangement of voltage values) as the object it interacts with — the matcher. In digital computers, each voltage value effects change according to the laws of electricity, but the set of effects due to all the voltage values, that together constitute the structure, are constrained by electronic comparator circuitry to collectively produce an outcome that is a consequence of the similarity between pattern and matcher. There is no physical law which relates the structure of the particular pattern to the outcome because laws are not about spatially-extended structures. Yet it is the structure which is causally responsible for the change. This is why such behavior is considered rule-like; because it involves an association between a physical structure (the rule antecedent in an "if-then" rule formulation, for example) and an action, which is usually a change in some measured attribute (e.g., switching a particular voltage value; often represented schematically as a simple arrow pointing from rule antecedent to rule consequent), that, in general, is physically utilized to trigger a more complex set of actions (rule consequent).

Structure fitting also occurs in biomolecular interactions, in particular those involved in enzyme catalysis and genetic expression. Pattee [9] notes for cellular enzymes that the fitting of an enzyme to substrate molecules results in the catalysis of a single covalent bond. The enzyme reduces "the complexity of the input interactions by means of its internal constraints", leading to a "simple", repeatable output action. As in computers, the input structures and matching are highly specific, resulting in outcomes that cannot be triggered except in such special "structural contexts", affording a degree of control and, therefore, stability with respect to producing the appropriate actions necessary for proper cellular functioning. The reason the action is simple (e.g., a single voltage change in digital computers or a single covalent bond in enzyme catalysis), is because the structures fit. Otherwise, it would be incorrect to say that the action was caused by the whole structure (i.e., by that particular content). A complex action — that is, an action which is structured (the result of, for example, spatially-extended structural change, multiple simple actions from multiple matchers, etc.) — could only occur if different pieces of the pattern were responsible for different physical changes (the excitation of hidden layer nodes in feed-forward, layered artificial neural networks is an example of multiple matchers producing multiple simple actions). This is distinguished from "partial matching", in which a single matcher fits only part of the pattern, but the fit still causes a simple action.

Thus, there are two ways interacting objects can effect change: 1) through their measured attributes, and 2) through their spatially-extended physical structures. Measured attributes and spatially-extended structure I refer to as "physical aspects" of objects [10,11]. I claim they are the only two.² In information processing systems, these physical aspects are informational because, in addition to being causal, they are representational. That is, insofar as the significant behaviors (i.e., informationprocessing behaviors) of analog and digital computers depend on measured attributes and extended physical structure, respectively, these physical aspects are informational in such systems. Their effects are realized through what I have called "measured-attribute information processing" (which is just another name for dynamical behavior in this context) and through pattern matching, respectively, which I refer to as "causal mechanisms".

The link between information content and the physical world is thus established through the causality of physical objects whose measured attributes and structures are identified with content. Determining how objects are causal — how their effects are brought about — grounds content because it indicates which physical aspect of a signal or message affects the significant behavior of an information-processing system, thereby making content part of the physical analysis.³ By contrast, mere cause-and-effect associations do not account for the causal mechanisms underlying them, and so are without a principled physical connection to content, which explains why logical or functional analyses of the effects of content are the basis for essentially all computational models of cognition.

4: Information content and mind

We have seen that extended structure is causal when other structures fit it; the resulting changes are due to the local "decision-making" constraints [12] imposed by these latter matching structures. But structure can also effect change by being *transmitted*. For instance, when a stone collides with a piece of soft clay, the stone's surface

¹Though symbols in digital computers are formed from combinations of separate circuit voltages and, therefore, are quite unlike the physically continuous contours we perceive the structures of macroscopic objects to have, the former are, nevertheless, structures, and I will refer to them as such.

²Objects also have functional and relational aspects (e.g., part/whole), but these are interpreted notions that involve other objects (i.e., context).

³One might refer to this sort of physical analysis as "causally deterministic", which, informationally, is broader in scope than standard physical state description analyses.

structure becomes imprinted in the clay, superposing with the clay's initial surface structure. Unlike the structureless change (simple action) caused by pattern matching, this kind of interaction involves a change in structure. Informationally, if structure is taken to embody content, then that content is transmitted to the receiving system by structurally altering it.⁴

Neurophysiological evidence indicates that the structures or patterns of light impinging on the retina are transmitted to the brain; images on the retina are retinotopically mapped onto the visual cortex. That is, the surface variations of two-dimensional projections of the environment are topographically preserved when they reach the sensory cortex of the visual system. The patterns are transmitted to and superposed on the activity of the visual cortex. There is evidence that this kind of process also occurs in the somatosensory cortex; signals from the body's tactile surface are systematically mapped onto the somatosensory cortex [13]. Moreover, as Churchland [14] notes, "there are many other cortical areas, less well understood as to exactly what they map, but whose topographical re-presentation of distant structures is plain."

Though it is still unknown what happens in the case of higher-level cognition, all of this suggests that information processing in the brain involves, at least in part, the superposition of one structure onto another such that the outcome is a composite, spatially-extended structure - a process that I refer to as "structurepreserving superposition", or SPS [11,15]. SPS is a third causal mechanism, distinct from pattern matching and measured-attribute information processing, but, like pattern matching, is based on the causality of spatiallyextended structure. Informationally, it involves the "direct interaction" of content, rather than the effects of content being mediated by a matching structure. For example, two visual signals which have similar content — i.e., their structures are similar — would, when superposed, produce a composite structure which could be considered a "generalization" of the two "instance structures"; similar contours reinforce while dissimilar variations become, in essence, "noise".5

⁴In this case, the physical effects of structure reflect the particular structural variations across it, in contrast to structures in pattern matching systems like digital computers where the forms of structures are, in general, independent of their effects simply because all that is required for them to effect change is a matching structure that physically fits them.

What is particularly interesting about the difference between pattern matching in digital computers and SPS, as well as important for understanding mind, is that in the former, content, identified as physical structure, does not really depend on the *specific* physical structure, because as long as a matcher is present to fit it the same changes could be triggered by any structure. So what we call the meaning of a symbolic expression in a computer is not intrinsic to its particular structure and, hence, to the computer, itself, since the changes the expression causes do not depend on — are not intrinsic to — that structure. This is not the case for brains, at least in those areas where there is evidence of SPS, because the structure of a signal is intrinsic to the changes the signal causes and, hence, to the brain itself. Thus, to the extent that meaning is associated with structure, the meaning of a "symbol" in the brain would be intrinsic to it. This may physically explain the difference between derived and original meaning (in computers and brains, respectively).

5: Causal mechanisms: a new physical framework

I claim there are no other kinds of causal mechanisms that determine how effects are brought about by their causes because there are only two physical aspects of physical objects and they can only cause physical changes in those same two physical aspects. Remarkably, such causal mechanisms appear to be independent. This second claim is based on the fact that one aspect is structureless while the other is spatially extended; that is, each mechanism is independent of the others either because its input is structured, or not, or because its effect (outcome) is structured, or not. This is not meant to imply that for any particular interaction more than one causal mechanism does not underlie change. Indeed, in all physical interactions there are changes in the measured attributes of the interacting objects. But the changes due to structure acting as structure - i.e., through pattern matching or SPS — are not described by lawful relationships between measured attributes of the interacting objects. For example, in a situation in which there is a room accessible only through a locked door and a person outside the door with a key, a "person-in-the-room" state can only be realized if the key fits the lock and its structure couples to that of the lock mechanism by becoming properly oriented. It is the fitting of the key and lock which leads to the simple action of changing the position of a door latch. Any other key, or the same key in different orientations would not have enabled the person-in-theroom state to be realized, regardless of the measured attribute values of the person or key.

Since physical laws do not describe the effects of information processing (i.e., the physical effects of content) when the representing entities are structures, a

physically fits them.

In addition to the above three causal mechanisms, there is a fourth, which includes those object interactions that generate new structures as a result of the objects involved having particular values or ranges of values of certain measured attributes (e.g., kinetic energy). In the case of classical mechanics, the structures are generated because the force of collision breaks one or both objects into pieces, creating newly exposed surfaces. But there would be, in general, no consistent relationship between the attribute values and the particular surface structures whose features actually depend on, for example, cleavage planes of the objects that split apart. Informationally, the value of a measured attribute (considered as content) is structureless and, therefore, unable to specify variations in surface structure more complex than itself; that is, unable to deliberately create more information. Only if additional constraining structures (analogous to the matcher structures in pattern matching) are present to control the

outcomes could there be more complex content embodied by the newly created structures.

more general physical framework is needed; one that has both measured attributes and extended structure as basic components, not just measured attributes as physical state descriptions do. Of course, physical state descriptions can, in principle, completely describe the physical behavior of information processing systems. But content, as structure, and how it effects change would be lost among the myriad quantities (and relationships between those quantities) that are used to describe the time evolution of the physical states of such systems. As Pattee [16] notes, "the physicist's theory recognizes no symbolic restrictions".

The issue of grounding content is usually ignored in the case of digital computers because our language-based descriptions (of physical behavior, for example) are exactly what a computer's architecture was designed to process; it enables the causality of symbolic structures through pattern matching. But insofar as such descriptions constitute a formalism (say, in the form of rules) for describing the computer's behavior, it is not physicallygrounded because there is nothing analogous to the mathematical infrastructure underlying physical state descriptions that could determine the effects of symbols short of us running it on a computer. That is, the values of the terms in physical state description formulations the content — are the values of the measured attributes of the actual objects being described, whereas the effects of symbols in language-based descriptions of behavior can only be determined if their spatial extensions actually cause change through pattern matching.

Thus, it is the computer, itself, that must physically implement the effects of structure-based content because its architecture was designed so that structure fitting actually takes place and causes change. We might, therefore, consider the computer to be a "physical formalism" that grounds our language-based descriptions of the effects of content; effects which we typically determine by interpreting such descriptions.

We cannot expect generalizations like physical laws in the case of structure causality because, in addition to time-dependencies, there are structural dependencies — e.g. structure identity — that are not part of interactions which depend only on the forces between objects. The only kind of generalization comes from the fact that in a computer, for example, any structure causes change through the same causal mechanism — pattern matching. This kind of generalization — how structure effects change — is a cornerstone of the causal mechanism-based physical framework I am proposing for grounding information content in the physical world. Since structures are not numerically-valued, such a framework is not amenable to mathematical formalization.

Of course, one might argue that structure can always be represented as an array of spatial position values. But such values are not the content. Nor are the mathematical relationships that determine the effects of those values on system behavior about content. Rather, content is the variation in spatial position values over the array. In other words, they should not be treated as an array of values, but

an arrangement of values. Physical state descriptions that incorporate these values do not explicitly capture the effects of the arrangement. They are only explicitly accounted for through pattern matching in an actual computer or by us "seeing" and processing the pattern once it has been projected onto our retinas. Similarly in the brain, I believe the arrangement of values underlies the content of its mental states.

Thus, I propose that physical state descriptions, and, therefore, the laws and situation-specific constraints they comprise, are constituent of systems (in this case, mathematically-based formulations) powerful enough only to account for measured-attribute information processing — that is, content in the form of numerical values of measurable attributes of physical objects — in a physically-principled way. Digital computers and other pattern matching systems, in contrast, are powerful enough to determine the effects of content, as spatiallyextended structure, through structure fitting. Finally, the brain is powerful enough to process structure interactions that result in spatially-extended structural changes (SPS)⁶, which may be why we are able to write computer programs and formulate physical state descriptions of the world.

6: Summary

I have argued that physical laws do not capture the effects of information content when it is manifest as spatially-extended physical structure because such laws are not about extended structure, nor are the causal relationships between objects from which the laws are derived dependent on structure fitting. It is compelling to think our physical state description formulations account for content in general because it is extremely difficult for us not to interpret them as doing just that or see ourselves as interpreting them. But we cannot allow our interpretations to falsely enhance the power of these formulations precisely because such systems of equations cannot account for externally imposed organizations of the quantities whose changes they determine.

Digital computers, on the other hand, physically enable content in the form of spatially-extended structure to cause change because their physical architectures have been engineered to implement the physical process of pattern matching. Interestingly, because pattern matching results in a structureless outcome, the actual structure of the input does *not* inherently determine the changes it causes, since any structure can cause the same physical changes. That is, for patterns to be causal there only has to be a matcher that fits them; structure and function are independent. This contrasts with the above hypothesized causal mechanism for information processing in the brain (SPS), for which the structure of the input *inherently* determines the changes

⁶Again, we cannot simply think of structure as a bunch of measured attributes, because that *separates* (i.e., leads to an abstraction of) what is informational from the physical.

it brings about; that is, the changes are, in essence, a simulacrum of the input structure. Structures in the brain, therefore, are not arbitrary with respect to the structures that produced them and, hence, with respect to the structures of their referents, so that they can be used to anchor our interpretations of other structures, such as external symbols, which are arbitrary.

Somewhat along the lines of Wheeler's "meaning circuit" [17], physics may indeed underlie all the phenomena we study, and, hence, give rise to biology and to us as communicators and the originators of meaning. But physical laws and physical state descriptions, which happen to be a powerful way of taming real world complexity, do not describe the physical processing of information content without our being present to interpret such descriptions of the world; that is, we keep track of what is information content and its effects in such formulations. We are able to do this because the brain is physically organized to maintain and process extended structures directly, which is presumably why interpretation is possible in the first place. Through SPS the brain is able to acquire information because signal content, in the form of two-dimensional projections of environmental structure, is actually transmitted to and manipulated by it as spatially-extended structure.

Thus, if we are to have a principled physical theory of information — one that accounts for the physical effects of content as spatially-extended physical structure, in contrast to theories that merely formulate various information measures — and, therefore, a theory of us as communicators capable of discovering the laws of physics that constrain our behaviors, a more general physical framework based on causal mechanisms, rather than just lawful associations between measured attributes, is necessary. Causal mechanisms physically specify how objects are causal when they interact and thus offer a physical explanation of how content can be communicated because they explain how signals physically inform receiving systems.

7: Epilogue

Ultimately, a physical framework based on causal mechanisms like the one proposed above will find its real scientific value in advancing our understanding of the brain as a mind. In that regard, one of its implications at this point is that only in the case of SPS is information actually communicated. That is, insofar as content is physically identified with structure, it is the only process that transmits that content from one location to another. In digital computers, on the other hand, there is no transmission of structure (or content), only a kind of "switching" control which is accomplished through the act of structure matching. But communication is not "action", it is receiving what was sent; i.e., transmission. For example, if everything that impinges on the sensory surfaces of a hypothetical robot merely triggers some internal process or set of processes, how can the robot "know" what was received, since "mentally" it did not actually receive it? Only its sensory surfaces did.

Thus, it would seem that what we call "information" really only exists in the head, and that SPS is, at least in part, the reason why things mean.

References

- [1] Brillouin, L. (1962). Science and Information Theory. (New York: Academic Press).
- [2] Chaitin, G. (1974). Information-theoretic Computational Complexity. *IEEE Transactions on Information Theory*, **20** (1):10-15.
- [3] Chaitin, G. (1982). Godel's Theorem and Information. *International Journal of Theoretical Physics*, 21(12): 941-954.
- [4] Shannon, C.E. and Weaver, W. (1949). The Mathematical Theory of Communication. (University of Illinois Press).
- [5] Dretske, F. (1981). Knowledge and the Flow of Information, (Cambridge, MA: MIT Press).
- [6] Haugeland, J. (1989). Artificial Intelligence and the Western Mind. In J.R. Brink and C.R. Haden (eds.) The Computer and the Brain: Perspectives on Human and Artificial Intelligence (Elsevier, North-Holland).
- [7] Weyrick, R.C. (1969). Fundamentals of Analog Computers, (Prentice Hall).
- [8] Wolfram, S. (1984). Computer Software in Science and Mathematics, Scientific American, 188-203.
- [9] Pattee, H.H. (1986). Universal Principles of Language and Measurement Functions In J.L. Casti and A. Karlqvist (eds), Complexity, Language and Life: Mathematical Approaches, (New York: Springer-Verlag).
- [10] Boyle, C. F. (1991). On the Physical Limitations of Pattern Matching. Journal of Experimental and Theoretical Artificial Intelligence, 3(2):191-217.
- [11] Boyle, C.F. (1992). Projected Meaning, Grounded Meaning and Intrinsic Meaning. Proceedings of the Fourteenth Annual Meeting of the Cognitive Science Society (Bloomington, IN: Lawrence Earlbaum), 714-719
- [12] Pattee, H.H. (1972). Physical Problems of Decision-Making Constraints, International Journal of Neuroscience 3:99-106.
- [13] Kolb, B. and Whishaw, I.Q. (1985). Fundamentals of Human Neuropsychology, Second Edition (New York: W.H. Freeman)
- [14] Churchland, P.M. (1989). A Neurocomputational Perspective, (Cambridge, MA: MIT Press).
- [15] Boyle, C.F. (submitted). The Ontological Status of Mental Objects.
- [16] Pattee, H.H. (1987). Instabilities and Information in Biological Self-organization, In F.E. Yates, ed, Selforganizing Systems: The Emergence of Order, (New York: Plenum Press).
- [17] Wheeler, J.A. (1986). Physics as Meaning Circuit: Three Problems. In G.T. Moore and M.O. Scully (eds.), Frontiers of Non-equilibrium Statistical Physics (New York: Plenum Press).