

Toward an Algorithmic Metaphysics

Steve Petersen

Department of Philosophy
Niagara University
Niagara University, NY 14109 USA
`steve@stevepetersen.net`

Abstract. There are writers in both metaphysics and algorithmic information theory (AIT) who seem to think that the latter could provide a formal theory of the former. This paper is intended as a step in that direction. It demonstrates how AIT might be used to define basic metaphysical notions such as *object* and *property* for a simple, idealized world. The extent to which these definitions capture intuitions about the metaphysics of the simple world, times the extent to which we think the simple world is analogous to our own, will determine a lower bound for basing a metaphysics for *our* world on AIT.

Keywords: metaphysics, formal metaphysics, computational metaphysics, algorithmic metaphysics, algorithmic information theory, real patterns.

Both philosophers and mathematicians have flirted with the idea that algorithmic information theory (AIT) could provide some foundation for basic notions in metaphysics. The main inspiration for this paper is one such hint from the philosophy side: in Daniel Dennett’s sketch of a metaphysics based on “real patterns”, he explicitly appeals to incompressibility, and offhandedly mentions the work of AIT theorist Gregory Chaitin in this connection.¹ Meanwhile AIT theorists since Andrey Kolmogorov frequently speak about, for example, the “information content” of *objects* generally, rather than of binary strings in particular. (Of course, “the original incentive to develop a theory of algorithmic information content of individual objects was *Ray Solomonoff’s* invention of a universal a priori probability . . .”²)

This paper aims to help bridge this gap between metaphysics and AIT. As befits a philosophy paper, it contains little in the way of technical results, but some ruminations aiming to pave the way for technical results to come. And as is typical of work bridging discipline *X* to discipline *Y*, *X* theorists are likely to complain that the treatment of *X* is far too simplistic and sloppy, and that the treatment of *Y* engages trivial details—while *Y* theorists complain conversely.

¹ See [2] p. 32. Others have developed Dennett’s metaphysics more fully than I have, most notably the “Rainforest Realism” of [9]. I think everything I say here is compatible with this work, while extending it to make more explicit the ties to AIT.

² [11] p. 333, my emphasis. Solomonoff’s partially autobiographical (and posthumous) publication [13] lists [12] as the paper that started it all.

I happily undertake that risk, with hope for some indulgence and patience from both.

My strategy is to construct an extremely simple toy world W , and give its metaphysics in terms of AIT. That is, using AIT I'll try to define key metaphysical notions for W such as *composite object* and *property*. These basics of synchronic metaphysics will be plenty to occupy this paper; in future work I hope to build on this foundation in order to model important diachronic metaphysical notions such as *persisting object*, *change*, and *cause*.

Eventually the goal is to connect these metaphysics to thriving programs in algorithmic *epistemology*. Work such as [8] and [7] extend Solomonoff's original insights for formalizing inductive reasoning. The Bayesian approach of [16] is closely related (as argued in [17]), and has demonstrated potential to characterize approximate truth [4] and perhaps more elusive philosophical fruit [3].

But first, the metaphysics. The extent to which these definitions seem to capture the intuitive metaphysics for the toy world W , times the extent to which we think the metaphysics for W are analogous to our own world, will determine an informal lower bound for the extent to which we think we might base a metaphysics for *our* world on AIT.

1 The Toy World W

W is built out of a finite set of n "objects" $\{w_1 \dots w_n\}$ that we will consider unproblematically fundamental.³ We'll need a fair number of them; one million would be plenty. These fundamental objects cannot be created or destroyed from one time to the next. In W there is one fundamental property, which each of these objects either has or does not have at any time. We can think informally of the w_i s at a time as a row of life-game-like cells that wraps into a circle, each of which either possesses the fundamental property (represented as '■' on the cell, and encoded as '1'), or does not possess that property (represented as '□' on the cell, and encoded as '0').⁴ We can think of W at the next time step as another circle of such cells just below. This W is, I hope, a very simple case for the kind of important features we seem to need to construct a metaphysics in our world: some unproblematically fundamental objects, one maximally simple fundamental property, and a succession of times.

More formally the $\{w_i\}$ form a finite cyclic group with a successor-like function, and W at a time t , or w_t , is a function $w_t : \{w_i\} \rightarrow \{0, 1\}$. W , in turn,

³ This does not require that they be just-plain *fundamental*; if Ladyman and Ross are right that objects are always patterns, "all the way down", they still allow for fundamental objects in the context of a fixed resolution.

Also, the use of "build" here and elsewhere in the paper is meant to echo the wide sense proposed by [1]; one of my many hopes for algorithmic metaphysics is that it can characterize the important commonalities among the "building relations" she discusses.

⁴ We want the cells to wrap into a circle basically because we do not want cells to the "left" to be simpler, in the AIT sense, than cells to the "right".

can be seen as a function from discrete possible times ($t \in \mathbf{N}$) to such functions; we could write the composite function as $w_{t,i}$. Let L , the *locations*, be the set of intervals on $\{w_i\}$.⁵

In W the w_i gain or lose their properties by the action of a fixed universal prefix Turing machine U . This machine has a unidirectional binary input tape, a bidirectional read-write binary work tape (including a blank symbol), and a write head on the w_i . This head can add or remove the fundamental property at its current cell or can move to the cell's successor. We'll abbreviate the first operations as '■' (for adding the property) and '□' (for removing the property), and the successor movement as 'S'. So that U does not have an unfair bias, we can suppose U is (or is one of) the "simplest" universal prefix Turing machine(s) to meet these specifications (on whatever chosen measure of simplicity for Turing machines); this is thought by AIT theorists to be the rough equivalent to a maximally uninformative prior.

Now consider any interval (location) $l \in L$, and let x_l represent the function w_t as restricted to that interval. An x_l is thus an ordering of zeroes and ones—in effect, a binary string.⁶ Let X be the set of all such binary strings up to the maximal length n , and for $x \in X$ let the standard notation $K_U(x)$ designate the Kolmogorov complexity of x relative to U —that is to say, $K_U(x)$ is the length of the shortest input required to cause our fixed U , when starting at the "left" end of the interval, to output x and then halt. Finally, let the idiosyncratic notation $x^\#$ designate the length of the shortest program required to output x "literally"—that is, the length of x plus some small constant for the computational overhead to print any given string.

2 Things in W

One of the main potential advantages of Dennett's "real patterns" approach to metaphysics is that it can make sense of *composition*—that special, mysterious way a bunch of things can come together to form a new thing, as when atoms can make up a molecule, or molecules can make up a brick, or bricks can make up a house. Peter van Inwagen calls this the *special composition question*: the question of when it is true that some objects compose a new object—when it is true that " $\exists y$ the x s compose y ."⁷ This is a classic philosophical question at least in the sense that it looks like the answer should be obvious, but reflection shows just about any consistent answer to be counterintuitive. At least for the simple W , though, I think AIT can provide a relatively intuitive answer to this vexed

⁵ That is, given the cyclic ordering relation $\langle x, y, z \rangle$, the interval $[a, b]$ is defined as all x such that $x = a$ or $x = b$ or $\langle a, x, b \rangle$.

⁶ I will sometimes conflate the function w_t over interval l , the set of ordered pairs associated with that function, and the resulting ordered binary string. I think (and hope) that nothing hangs on this conflation.

⁷ [15] p. 30. Note that, at least for W , I am happy to use 'thing' and 'object' interchangeably.

question.⁸ This definition of composite objects will serve as a kind of lynchpin for the synchronic metaphysics that follow.

In W , an object is composed of other objects when the objects together form a “real pattern”—that is, the arrangement of the objects and their properties is easier to specify, computationally, than simply giving the complete details at the fundamental level. A simple, standard example (adapted for W) is when many of the fundamental objects (cells) in a row all possess the fundamental property. If the string of ■s is sufficiently long, a program equivalent to “for $i = 1$ to 10,000, ■S” will be much shorter than the literal-print “■S■S . . . ■S.” It is natural to see such a sufficiently long string of ■s amidst an otherwise chaotic jumble of squares as an object in its own right.

I propose this definition for identifying the non-fundamental things in W .

Definition 1. x_l is a **composite object** if and only if

1. $K_U(x_l) < x_l^\#$ (the **compressible clause**)
2. There is no partition of l into intervals $\{l_1 \dots l_n\}$ such that $\sum_i K_U(x_{l_i}) \leq K_U(x_l)$ (the **minimal clause**)
3. There is no interval l' containing l such that $K_U(x_{l'}) \leq K_U(x_l)$ (the **maximal clause**)

Does this definition adequately capture what we might call (non-fundamental) “objects” in W ? Since none of us has any experience in W , it might be hard to tell what we could reasonably call a composite object there. Still, this definition has intuitive features in W that would, if they were carried over into our world, provide some grip on major challenges in metaphysics.

2.1 Composition and Division

One central challenge for any account of composition is to negotiate both the Scylla of *universalism*, according to which any mereological sum of objects is another object, and the Charybdis of *nihilism*, according to which the only objects are the fundamental ones. Universalists implausibly claim that “my desk plus the Eiffel tower” is a genuine object in its own right, while nihilists implausibly claim that desks and towers do not literally exist. They are driven to these extremes basically because it is hard to find a principled line to draw between them. Dennett’s “real patterns” approach promises to provide just such a line.

The first clause of Definition 1 already implies that universalism is false for W , for most binary strings are not compressible and thus not objects. Some of the problem that motivated universalism remains, however. Let o_1 be 10,000 ■s in a row, and let o_2 be a binary representation of the first 10,000 decimal digits of π at a location immediately adjacent, and finally let $o_3 = o_1 \cup o_2$. This o_3 looks suspiciously like an arbitrary mereological sum such as “my desk plus the Eiffel tower”, and yet it is easily compressible. It is not an object, though, because of the minimal clause; it possesses a natural decomposition into o_1 and o_2 , where

⁸ One I hope more intuitive than van Inwagen’s answer for our world; he claims that no composite objects exist except for living beings.

$K_U(o_1) + K_U(o_2) \leq K_U(o_3)$. In summary, there is no simplicity *gain* in treating the two objects as one—and so, on the real patterns metaphysics, no ontological gain either.

It may seem that the minimal clause will overgenerate, and likewise rule out o_1 as a genuine object, since that long string of ■s can apparently be decomposed into other compressible objects. For example, split o_1 into two longish substrings; call them o_{1L} and o_{1R} (for “left” and “right”). Each would be compressible on their own, and their union is just o_1 —so o_1 seems to run afoul of the minimal clause, and thus of objecthood. This would be W ’s analog to the Charybdis of nihilism, since if anything is a composite object in W , surely o_1 is!

But the unified o_1 is different from the chimera o_3 in an important respect. The programs for each of o_{1L} and o_{1R} will look very like the program for o_1 , and summing them will double the computational overhead. In this way o_1 gains in simplicity over the sum of its substrings, and so is a genuine object by Definition 1. Similar considerations presumably apply to any substrings of o_2 .

Still, substrings o_{1L} and o_{1R} seem to be objects in their own right; they surely would be if they were “on their own.” If they are objects, though, we face a dilemma. On the one hand, if we say that o_{1L} and o_{1R} compose o_1 , we must also say the same for a great number of other such partitions into sufficiently long substrings. o_1 is thus “composed” of way more overlapping “objects” than would be intuitive. On the other hand, if we say that o_{1L} and o_{1R} do not compose o_1 , then we must say that o_1 is identical to the union of two objects, but not composed of those objects.

I think the best solution is simply to deny object status to substrings o_{1L} and o_{1R} , as the maximal clause does. Both o_{1L} and o_{1R} are subsets of o_1 , and $K_U(o_1) = K_U(o_{1L}) = K_U(o_{1R})$, so substrings o_{1L} and o_{1R} are not objects in their own right.⁹

The maximal clause has a natural analog in our world, since it prohibits the flipside of arbitrary summation—namely, arbitrary division. Just as in our world it is intuitive to say there is no such “object” as the northern half, northern third, or northern 3/17ths of my desk, so too in W our string o_1 does not have its many compressible substrings as genuine parts, according to Definition 1. The natural motivation here is that a part that is more complex than its whole—because, say, it depends in some way on a description of the whole to be specified—is not an object in good standing on its own. (Definition 1 thus denies what [14] calls the “Doctrine of Arbitrary Undetached Parts”.)

⁹ Actually this is complicated by the fact that, for example, the binary specification of 10,000 in an instruction like “for i = 1 to 10,000” will take more bits than if it were 1,000 instead. Thus strictly speaking the 1,000-length substrings are *not* part of a longer string with equal or lower complexity; their minimal generating programs will be slightly shorter. I do not think this is a deep problem, though I can only think of a few kludgy fixes for it—for example, we could require the inequality not hold within the logarithm of the length difference between the containing and contained string, or perhaps set up the Turing machine so that it has typed registers for such purposes that always take the maximal possible bits.

But note that Definition 1 *does* allow for composing objects out of other composites; there just needs to be some complexity savings in so composing. For adjacent composite objects $\{o_i\}$ to compose o , we just need it to be the case that $K_U(o) < \sum_i K_U(o_i)$, and that $K_U(o_i) < K_U(o)$ for each i . In other words, the super-composite must be simpler than the sum of its parts, but not as simple as any individual one. Suppose, for example, that w_t has a streak of 5,000 ■s, followed by 3,000 □s, and then another 5,000 ■s, followed by another 3,000 □s, and so on, so that the black-white stripe alteration pattern is repeated (say) six times. In this case we can write one loop around the loops required for each stripe, and so we have a compressing program for the whole that is longer than the compression for each natural part, but shorter than the sum of each. Each stripe is an object on its own, and together those stripes compose the object that is the larger stripe pattern. No two of those stripes taken together is an object, though—and if we had only two stripes in a row like that, without the repeating pattern, then they would each be a single object that do not compose a new one together. This, I think, is an intuitively pleasing result.

2.2 Scattered Objects

The examples so far have concentrated on intuitively connected objects, but notice that Definition 1 allows for something like scattered objects too. Just as a cloud may count as one thing even without close bonds among the water molecules, or a jigsaw puzzle might count as one thing even when it is in pieces, so too a set of cells over an interval may count as one thing even when there is a good deal of random noise over the interval, scattering the pattern. Consider, for example, a large interval over which the odd-indexed cells are all ■s, while it is random whether the even-numbered cells in between have the property. Such a string would count as an object despite a certain intuitive lack of internal cohesion. Minimally interesting objects in W (but most common, in the sense of most probable to occur) will be long-enough strings where the preponderance of one property over another is just sufficient to allow compression. These objects will be, in some sense, maximally scattered.¹⁰

¹⁰ One odd consequence of Definition 1 is that whether such scattered bits of order in the chaos count as one object or not will depend on *exactly* how spread out they are. At some point a slight preponderance of order mixed in a lot of chaos no longer gains enough simplicity advantage to make up the overhead required for the compressing calculation. In our world it's fairly natural to say that a group of water molecules with sufficient average distance no longer constitute a cloud, and jigsaw pieces sufficiently removed from each other are no longer that original puzzle—if, say, some pieces are in a box in the basement, and some in a landfill across town. On the other hand it is not so intuitive that there is a precise boundary here; the mere movement of one extra millimeter could hardly make the difference, but in W just one extra bit of noise can be enough. Insisting on lack of such precision when determining whether an object exists, however—insisting, in other words, that it is *vague* whether some object exists or not—has its own very serious problems. See, for example, [5] for a classic, one-page case against, and [19] for a much more extended

Another consequence to consider is that scattered objects, in our world anyway, allow for interpenetrability—a kite can fly through a cloud, but that does not intuitively destroy the kite in favor of a kite-cloud, for example. This leads us to the complicated topic of spatial overlap and coincidence for objects in W .

2.3 Object Overlap and Coincidence

The possibility of building new objects out of more fundamental ones gives rise to another standard metaphysical puzzle: that of *constitution*. For (worn) example, a sculptor forms a lump of clay into an elegant statue. Intuitively, the lump of clay has not disappeared; it has just been reshaped. And intuitively, the statue did not exist before the lump gained that shape. Thus many are tempted to say there are now two things (the lump of clay and the statue) where there was once one—two material things that occupy exactly the same location, so are coincident in at least this sense. The fairly neutral description of the case is that the lump of clay *constitutes* the statue, and the puzzle is in explaining just what this relation of “constitution” amounts to.¹¹ Our toy world W has an analogous puzzle, I think, and AIT provides at least some leverage against it.

As a warmup, consider first the issue of object *overlap*, without exact spatial coincidence. Suppose Lafayette Avenue and Grant Street intersect. That intersection is intuitively part of both streets; after all, you do not suddenly abandon Lafayette while crossing Grant, nor vice-versa. If so, that rough square of pavement belongs to both Lafayette Avenue and Grant Street, even though the two roadways are not the same thing. The two roads overlap.

Here is what I take to be an analogous situation in our one-dimensional W : consider a long string in w_t that represents the first 11,806 binary digits in the decimal expansion of π ; call it o_π . As it happens, this string ends in sixteen \square s.¹² But now suppose that those sixteen \square s are also followed by a great many more \square s in a row, to make 10,000 overall, and call that string o_0 . Of course o_π and o_0 overlap, in the sense that their location intervals on w_t intersect. To see that both are objects, consider each clause of Definition 1:

Compressible. Each is clearly compressible to shorter than their literal printings.

Minimal. Neither is decomposable into substrings that would save in complexity over the whole. Though o_π might look like it has a natural division, there is no complexity savings in so doing; it is better to have a π -calculating loop for 11,806 rounds then to have the same π calculation for 11,790 rounds and then a separate “repeat \square 16 times” loop.

Maximal. Neither object is part of a bigger object on a containing interval. The union of o_π and o_0 would have two reasonable compressing programs (into “ π for 11,806, then \square for 9,984” or “ π for 11,790, then \square for 10,000”)

treatment of problems associated with vagueness at the metaphysical level. For a defense of vague identity, see [15].

¹¹ See [18].

¹² Thank you, <http://www.befria.nu/elias/pi/binpi.html> !

but neither such program would be shorter than those for o_π or o_0 on their own.

Thus it seems W can have overlapping objects. Perhaps now it is clear that it is *also* possible in W to have one object entirely contained by another, as a yolk is contained by an egg in our world. Consider, for example, the first compressible stretch of 0's in a very long expansion of π . Again, I think this captures intuitions.

Object coincidence, however, appears to be a different matter; while the story of the statue and the clay makes it seem at least an open possibility in our world, Definition 1 in effect stipulates against it for W , since it individuates objects by their locations. This is a problem, and I have two possible responses to it.

Here, I think, is a rough equivalent in W to the story of the clay and the statue: o_c (the “clay”) is a long string built out of a random mix of the short strings ‘111’, ‘0010’, and ‘011010’. This fact makes o_c compressible. The string o_s (the “statue”), on the other hand, is made out of the same substrings, but their succession follows some identifiable (if somewhat arcane) pattern. This string is compressible in the same way o_c was, but also in an even more efficient way that encodes the pattern of the component strings. In this sense, the “clay” of o_c is still there, but with a further pattern layered on top of it. (Note it is not a case of object composition, though, since the short substrings are not objects.)

One way to capture our metaphysical intuitions for such a case would be to alter our definition slightly: identify objects in W not with (the equivalent of) compressible binary strings, but instead with the compressing *programs* for U that could generate such strings. Very roughly speaking, we might say objects are individuated by their Aristotelian formal causes, and not their material ones. Since (at least) either of two different compressing programs could potentially produce the same string o_s over the same interval, there are then (at least) two objects in that location.

Another possible option is to keep our definition and say that there is indeed at most one object in any location, but the object has two compatible *properties*: in our W example, the object both has the property of being a mix of three certain substrings, and has the property of having those substrings arranged in such-a-way. In the real world analog, there is one thing on the table, and it has both the relatively important property of being a statue, and the compatible property of being a lump of clay.

This approach worries many, because it requires the lump and statue to be identical, but only contingently; the clay lump might not have been a statue, even though the clay lump had to have been a clay lump. In other words the lump and statue have different modal properties, and that is at least odd if the lump is the same thing as the statue. How problematic this is depends on how we construe modal properties, but I suspect AIT has good prospects for modeling the “counterpart theory” approach of [10].

At least when it comes to AIT metaphysics in W , my working hunch is that not much hangs on the difference between these two options. To see why, though, we need to discuss how AIT might pick out *properties* in W , and their relation to *algorithmic sufficient statistics*.

3 Properties in W

Properties in W built out of the fundamental are most naturally thought of as sets of *possible* objects for W ; they are in this sense *intensional* rather than extensional. This allows for the possibility of uninstantiated properties; it may be that for some P no object in P ever actually appears in W . I am okay with this if the serious metaphysicians are. (Are they?) Since we want to be able to talk about real patterns of such “abstract” sets that might not be realized in w_i concreta, programs causing U to output to cells in W will not be sufficient. Instead, we treat the objects in question as functions from $l \in L$ to $\{0, 1\}$, encode such functions into binary strings in a standard way, and encode sets of such strings into a new binary string in a standard way.¹³ Thus we can speak of the Kolmogorov complexity of such a set as the program required to output the string encoding the set on U 's work tape.

It may be tempting to consider any such set of objects S that is compressible (that is, a set where $K_U(S) < S^\#$) a property. This is no good, however; for example, a set of m -length blocks at various locations (for sufficiently large m) plus one long representation of π digits will count as a property, on this view, even though this set seems unnaturally gerrymandered. The challenge, then, is to carve the total set of possible objects into its natural joints, whatever “natural” means here.

I think—with somewhat less confidence than before—that the best approach is to follow the same technique that was used for defining objects. Just as we want “real” objects to be simple relative to the sum of their parts, and not arbitrary sums or divisions, so we want “real” properties to be simple relative to the sum of their objects, and not arbitrary disjunctions or conjunctions.

Definition 2. *Let O be the set of all possible objects in W . Then $P \subseteq O$ is a **property** if and only if*

1. $K_U(P) < \sum_{o \in P} K_U(o)$ (*compressible*)
2. *There is no partition of P into $\{P_1 \dots P_n\}$ where $\sum_i K_U(P_i) \leq K_U(P)$ (*minimal*)*
3. *There is no $P' \subseteq O$ such that $P \subset P'$ and $K_U(P') \leq K_U(P)$ (*maximal*)*

I think Definition 2 does a respectable job capturing intuitions about properties.

First, any P satisfying it must have a relatively short program to generate it, and thus a relatively short description—which seems to imply a relative naturalness to it. The minimal clause rules out arbitrary disjunctive properties; generating the set of all objects that are either solid blocks or π representations is no more simple than generating the set of blocks and then the set of π representations. The maximal clause rules out arbitrary conjunctive properties; the set of all sufficiently long blocks that terminate before $w_{55,510}$ is not a proper

¹³ In some sense the set of *possible* objects should include ones bigger than those possible in W , including (perhaps) infinite ones, or ones with more than one fundamental property, *etc.* I don't consider possibility in this sense here.

property, I think, though I think the set of all sufficiently long blocks is, and maybe the set of all objects that terminate before $w_{55,510}$ is too.¹⁴

One intriguing feature of Definition 2 I have only begun to explore is its tie to *algorithmic sufficient statistics*.¹⁵ An algorithmic sufficient statistic for x is a finite set S such that

$$K_U(x) = K_U(S) + \log |S| + c$$

The c is a fixed constant—namely, the length of the program U requires, when given any set-generating program and index, to output the set element at that index. The idea is that generating a set containing x and then locating x in that set is as efficient as the minimal program for x . Note that for *any* S containing x , $K_U(x)$ will always be less than or equal to $K_U(S) + \log |S| + c$, since once given S one can always find x by simply enumerating its elements. Sets that are algorithmic sufficient statistics for x are “optimal” for x in the intuitive sense that important information about x is already captured by its membership in set S , so that no further point of substance can then identify x within S . It is thus natural to think of S as a model for x , while providing its index in S is like setting the parameters of the model.

Theorem. *If $S \subset O$ is an algorithmic sufficient statistic for some $o \in S$, and if it is maximal in the sense of Definition 2 (i.e. there is no $S' \subseteq O$ such that $S \subset S'$ and $K_U(S') < K_U(S)$), then S is a property.*

Proof. Compressible clause: Since S is optimal for o , we know $K_U(S) < K_U(o)$ (by $\log |S| + c$), and $K_U(o) < \sum_{x \in S} K_U(x)$, so $K_U(S) < \sum_{x \in S} K_U(x)$.

Minimal clause: Suppose for contradiction that there is some partition $\{S_i\}$ of S such that $\sum_i K_U(S_i) \leq K_U(S)$, and consider $S_j \ni o$. Then as for any set containing o ,

$$K_U(o) \leq K_U(S_j) + \log |S_j| + c$$

And since S is optimal for o , that means in turn that

$$K_U(S) + \log |S| + c \leq K_U(S_j) + \log |S_j| + c$$

By supposition $\sum_i K_U(S_i) \leq K_U(S)$, so in particular $K_U(S_j) < K_U(S)$. Thus

$$K_U(S_j) + \log |S| + c < K_U(S_j) + \log |S_j| + c$$

This implies $\log |S| < \log |S_j|$, where $S_j \subset S$ —a contradiction.

Maximal clause: By supposition. □

¹⁴ I confess whether “objects terminating before w_i ” is a property by this definition has stumped me, for now; it is certainly difficult to generate all objects so placed, since that would require a way to recognize objects, and that would in turn require computing Kolmogorov complexity. But generating such a set *given* O is, I think, pretty straightforward, and maybe that’s the proper standard.

¹⁵ See [11] p. 406, or [6] p. 29.

I suspect the connection between Definition 2 and algorithmic sufficient statistics may run deeper, though I have been unable to demonstrate as much in time to complete this paper. For example, I thought perhaps all *minimal* sufficient statistics for all $o \in O$ would be properties, but haven't been able to prove it, and now doubt it. I also cannot yet prove anything interesting in the other direction.

At any rate, sufficient statistics are about summarizing the “meaningful information” in data; they provide as much information as the data set itself for picking the best model out of a given model class. (In AIT this model class is very wide—the set of all computable models.) This notion of separating out the meaningful information from the happenstance details has natural connections to the notion of a real property of an object, since we can think of properties as fundamentally a matter of *abstraction*; to say that two non-identical objects share a property is to neglect some information in each in order to highlight substantive information they share.

Intuitively one can abstract from an abstraction to get another genuine property, and Definition 2 allows for these more abstract properties in the same way that Definition 1 allows for super-composed objects. If $P \subset O$ has a partition into properties $\{P_i\}$ where all $K_U(P_i) < K_U(P)$, and $K_U(P) < \sum_i K_U(P_i)$, then intuitively P summarizes something important that the P_i have in common. Thus consider for example these three sets:

- All blocks of ■s exactly 1,017 long (abstracting only from the location of the object)
- All blocks of ■s long enough to compress (abstracting from both location and block length)
- All blocks repeating any pattern short enough to be repeated often enough to be compressed (abstracting from location, length, and pattern to repeat)

The program to generate each of these sets will be short, but longer than the one before, and so each will (I think) meet both their minimal and maximal requirements.¹⁶

Thus our “statue” o_s from section 2.3 has two properties that are both—if I understand correctly—algorithmic sufficient statistics. We could identify an object with its sufficient statistic and index (the abstract specification and its “realization”?), or we could simply think of these as interesting properties of one object.

References

1. Bennett, K.: Construction area (no hard hat required). *Philosophical Studies* 154(1), 79–104 (2011)
2. Dennett, D.C.: Real patterns. *The Journal of Philosophy* 88(1), 27–51 (1991)

¹⁶ Such successively abstract properties stand to each other roughly as *blue* stands to *colored*—both first-order properties of objects. But blue, the property, intuitively itself possesses a property: it is a color. The need to refer to higher-order properties can be accommodated easily enough, I think; they should be compressible, minimal, and maximal sets of properties of the order below.

3. Dowe, D.L.: MML, hybrid Bayesian network graphical models, statistical consistency, invariance and uniqueness. In: Bandyopadhyay, P.S., Forster, M.R. (eds.) *Philosophy of Statistics, Handbook of the Philosophy of Science*, vol. 7, pp. 901–982. Elsevier Science & Technology (2011)
4. Dowe, D.L., Gardner, S., Oppy, G.: Bayes not bust! Why simplicity is no problem for Bayesians. *British Journal for the Philosophy of Science* 58(4), 709–754 (2007)
5. Evans, G.: Can there be vague objects? *Analysis* 38(4), 208 (1978)
6. Grünwald, P., Vitányi, P.: Shannon information and Kolmogorov complexity (September 2004), <http://de.arxiv.org/abs/cs.IT/0410002> (last accessed June 7, 2010)
7. Grünwald, P.D.: *The Minimum Description Length Principle*. MIT Press (2007)
8. Hutter, M.: *Universal Artificial Intelligence*. Springer (2005)
9. Ladyman, J., Ross, D., Spurrer, D., Collier, J.: *Everything Must Go: Metaphysics Naturalized*. Oxford University Press, Oxford (2007, 2009)
10. Lewis, D.: *On the Plurality of Worlds*. Blackwell, Oxford (1986)
11. Li, M., Vitányi, P.: *An Introduction to Kolmogorov Complexity and Its Applications*, 3rd edn. Springer, New York (2008)
12. Solomonoff, R.: A preliminary report on a general theory of inductive inference. Tech. Rep. V-131, Zator Co. and Air Force Office of Scientific Research, Cambridge, Mass. (February 1960)
13. Solomonoff, R.: Algorithmic probability – Its discovery – Its properties and application to strong AI. In: Zenil, H. (ed.) *Randomness Through Computation: Some Answers, More Questions*, pp. 1–23. World Scientific Publishing Company (2011)
14. van Inwagen, P.: The doctrine of arbitrary undetached parts. *Pacific Philosophical Quarterly* 62, 123–137 (1981)
15. van Inwagen, P.: *Material Beings*. Cornell University Press, Ithaca (1990, 1995)
16. Wallace, C.S.: *Statistical and Inductive Inference by Minimum Message Length*. Springer (2005)
17. Wallace, C.S., Dowe, D.L.: Minimum message length and Kolmogorov complexity. *The Computer Journal* 42(4), 270–283 (1999)
18. Wasserman, R.: Material constitution. In: Zalta, E.N. (ed.) *The Stanford Encyclopedia of Philosophy* (Spring 2009)
19. Williamson, T.: *Vagueness*. Routledge (1994, 1996)