

A COMPROMISE BETWEEN REDUCTIONISM AND NON-REDUCTIONISM

ERAY ÖZKURAL

This paper investigates the seeming incompatibility of reductionism and non-reductionism in the context of complexity sciences. I review algorithmic information theory for this purpose. I offer two physical metaphors to form a better understanding of algorithmic complexity, and I briefly discuss its advantages, shortcomings and applications. Then, I revisit the non-reductionist approaches in philosophy of mind which are often arguments from ignorance to counter physicalism. A new approach called mild non-reductionism is proposed which reconciliates the necessities of acknowledging irreducibility found in complex systems, and maintaining physicalism.

1. Introduction

New mathematical techniques have enabled scientists to gain a better understanding of non-linear dynamical systems, while similar advances have occurred in computer science for constructing models of natural phenomena as information processing systems. A particularly interesting research direction in complexity sciences is the study of conditions under which self-organization occurs.¹² Complexity theorists pay special attention to the concept of *emergence*. Gershenson and Heylighen give a typical definition of emergence:¹³

Emergent properties characterize a system that cannot be reduced to the sum of its parts

The authors then give the most striking example that an animate cell is composed of parts that are inanimate. Thus, to be animate must be an emergent property since the system cannot be explained in terms of its parts. In the same paper, the authors show analysis/reductionism as a shortcoming of classical science, which is indeed the case.

However, the identification of non-reductionism with this particular notion of “emergent property” leaves something to be desired. In this paper, I will argue that the *mechanical* view of the nature describes Heylighen’s definition of complexity¹³ sufficiently well, which is:

An achievement of AIT is the proof that randomness and incompressibility are identical concepts. That is to say, one cannot find a (much) shorter program for a random sequence, with the consequence that minimal programs are random. This discovery has been confusing as well, causing critics to claim that they can denote with a single symbol a message of any complexity. While this is true, such a decompressor would be useful for only one sequence, while a universal computer is a decompressor that can generate any message. Thus, if one “hacks” a universal computer to achieve this, one will have to encode the additional information inside the computer, and algorithmic information is not lost.

2.2. *Physical complexity*

We offer the following purely physical metaphors to make it easier for the reader to form an intuition about the physical relevance of the information theory point of view.

- physical message complexity: volume of the smallest machine that can output a message in any fashion. The condition is that the output message is encoded *as is*, it should be written symbol for symbol. The machine can encode each bit in a distinct physical property. Like the AIT definition, with increasing complexity, we would need a larger machine to achieve this.
- constructive complexity: volume of the smallest machine that can precisely build a desired object (up to physical limits, or required accuracy).

These metaphors are given to furnish the understanding of algorithmic complexity as a physical concept. Every bit in a sequence of bits extends in space, and must exist as a distinguishable physical property. That is to say, *length* may be conceived as the physical length.^a

Considering physical message complexity, the volume of this machine is similar to the length of a minimal program in a programming language. However, in the case of the machine, this definition is literally universal. The smallest such volume does not depend on the choice of a universal computer; effectively this sets U to our physical universe. Likewise, we would not think that we could design a storage medium that is more compact beyond a certain physical limit.

^aTime can also be put into this picture. I speak of space for the sake of simplicity. An elaborate version of the physical view is mentioned by Bennett et al.¹

2.3. Adequacy of algorithmic complexity

Algorithmic complexity beautifully captures our basic intuitions about complex systems. Recalling Heylighen's definition, we can view a complex system as a graph (or network) where vertices are components and edges are interactions. According to subadditivity of algorithmic information, the complexity of a message is the sum of the complexities of its components, with the addition of a constant term. That is, $H(x, y) = H(x) + H(y/x) + O(1)$ where $H(y/x)$ is conditional (or relative) algorithmic complexity, $H(x, y)$ denotes the complexity of concatenation of messages x and y , and $O(1)$ is a constant. Conditional algorithmic complexity $H(y/x)$ is the length of the shortest program that generates y given the shortest program for x . The complexity of each component therefore does add to overall complexity. However, it is not a simple sum of the complexities of parts. If the components are identical, then there will be little increase. If the components are all very different, each will contribute a large amount of algorithmic complexity. The algorithmic complexity of the graph of the system will also increase with the number of vertices, the number of edges, and the randomness of the topology of the edges.^b

2.4. Shortcomings of algorithmic complexity

The definition of algorithmic complexity neglects the temporal dimension which is indispensable in the study of dynamical systems. Without time, we could not formulate such concepts as trajectory and stability. Algorithmic complexity does not distinguish between passive stores of information and active users of information, thus it has to be supplanted by descriptions of a dynamical system in the study of many natural complex systems.

There are significant features of biological systems that algorithmic complexity does not model. For instance, algorithmic complexity disregards redundancy in the data, while redundancy and fault-tolerance are fine features of biological systems. Biological systems also seem to achieve optimizations other than reducing program-size, e.g., time and energy optimizations. Critics have also argued that algorithmic complexity does not model the interestingness or usefulness of a message. In classical information theory, The notion of "relevant information" has addressed the problem of interestingness.¹⁸ Hutter has pursued in detail an approach based on algorithmic probability including agent models that maximize an arbitrary

^bA regular topology takes less bits to specify than an irregular one.

utility function.¹⁴

Another shortcoming is the fact that algorithmic complexity is uncomputable. Therefore we always have to stick to much simpler approximations if we need to explicitly calculate the complexity of a message. This is not a shortcoming in theory, only a shortcoming in practice.

2.5. Applications and related work

Mutual algorithmic information $H(x : y)$ between two messages quantifies how much algorithmic complexity is shared. In the visual descriptions of a car and a plane, “wheel” could be part of shared information. The joint information $H(x, y)$ and mutual information are related via $H(x : y) = H(x) + H(y) - H(x, y) + O(1)$. Mutual information measures how easier it is to compress two strings together, than separately.

Mutual information can be used to quantify the amount of knowledge a biological agent has about the world by the mutual information between the agent and the world. While this a good characterization of the amount of knowledge, it says nothing about the dynamical aspects. The operational facets of intelligence could be incorporated by modelling which events it could predict, an idea that may be important since some theoretical neuroscientists propose that it is one of the main functions of the brain.² It is also worth mentioning that a minimal program can be used to make predictions; Solomonoff presents a particularly accessible introduction to applications of algorithmic complexity in AI using inductive methods.¹⁹ Some other interesting philosophical applications of AIT may be found in a philosophical paper by Chaitin including a thesis about perception.⁷ Another interesting work relating the information theoretic point of view to complexity sciences is due to Bialek et al.³, suggesting that predictive information has a close relation to the measures of complexity in dynamical systems theory.

3. Non-reductionism and physicalism

Proponents of non-reductionism often bring the human mind as an example of an irreducible system. In philosophy of mind, the tension between reductionism and non-reductionism has been a major driving theme of contemporary research. A motivating question is to find a theory in which physicalism becomes compatible with non-reductionism. Physicalism is the view that mental events are ultimately physical, i.e., material. I will use the term extra-physical for those concepts which do not befit a physical description.

Complexity sciences may be seen as embracing a physicalist methodology, for the tools of a complexity scientist are usually physical: trajectories, dynamics, interactions, entropy, etc. This attitude does not carry over to philosophy of complexity. In philosophical discussions we often see a drift towards some old concepts in philosophy such as dualism.

3.1. *Non-reductionism in philosophy of mind*

A fundamental problem in philosophy of mind is that it does not seem there is a place for qualia such as color, or pain in a human organ. I can see green, but when I open up a brain I do not find any green, only neural tissue. On the other hand, there seems to be a strong relation between experience and the physical state of brain. For instance, we can invoke certain sensations by stimulating the brain with electrodes. In philosophy of mind, we describe this relation as: mental events *supervene* on physical events. We can speak of a general principle of supervenience in philosophy of mind: there are no mental differences when there are no physical differences.

However, it is difficult, if not impossible, to fathom a relation between this complex, impenetrable organ and our basic perceptions. Thus, the philosophers have been creating new ideas for centuries, imagining that the mind must be something else than the brain. This has traditionally had other reasons such as the belief that we are special. Descartes thought that while the animals were automata, we were more than mere machinery. One reason for this was the apparent inadequacy of a mechanism. A machine could do only one kind of a thing, but the human mind seemed to be capable of anything. Thus, the human mind could not be a machine. Of course, Descartes would not have known the concept of a universal machine which was discovered by Alan Turing, a machine so flexible that it can mimic any other machine.¹⁷

Descartes proposed that the immortal soul that he imagined interacted with the human body through the pineal gland. This was not a good idea as examinations of the pineal gland showed. In general, the view that the stuff of mind is distinct from matter is known as substance dualism. Descartes subscribed to this view; he thought that mind did not extend in space like matter. There are two other kinds of *ontological dualism*. Property dualism suggests that mental properties are fundamentally different than material properties. The weakest statement of dualism is predicate dualism¹¹, which states that mental (or psychological) predicates cannot be reduced to material predicates, although supervenience holds.

A primary *non-reductionist* argument against physicalism is the knowledge argument.¹⁶ Imagine a neuroscientist who has been born and raised in a black and white room. The neuroscientist knows every physical fact about the brain. However, she has never seen red; she does not know what red is like. The argument then quickly concludes that, the knowledge of red cannot be reduced to physical facts about the brain, so it must be extra-physical.

3.2. *Arguments from ignorance*

An extreme non-reductionist attitude may result in dualism. Both the non-reductionist predicate dualism and knowledge argument views purport to show that ordinary physicalism is insufficient to explain the mind. The proponents of these arguments seem to arrive at their conclusions through reinforcing our ignorance in cognitive science.⁸

Predicate dualism relies on the fact that there are still many things unknown about the physical processes in the brain. The claim is made that there are no strict bridge laws connecting mental events and physical events. If there were such laws, it would be possible to fully reduce psychological states to physical descriptions. Assume that this is not possible. Then, clearly, there is something more to psychological states than can be explained by a physicist. It is also quite difficult to see how the part of the psychological state that cannot be reduced to a physical description could legitimately exist as part of our causal world.¹⁵ While it is possible to defend that a sufficient form of supervenience would hold in this case,¹¹ this view has the flavor of an idea that depends on our lack of progress in neuroscience. We do not have a comprehensive neuroscientific theory of how consciousness comes about in the brain. This lack of knowledge is termed the “explanatory gap” and much has been said of it.⁴ Thus, the predicate dualist has a lot of leeway to persuade the readers about the non-existence of bridge laws, since neuroscientists have not shown them yet. However, this skepticism does not seem too different from invoking supernatural forces in any phenomenon that science cannot explain currently. It also seems to contradict with the expected uniformity of natural law. In particular, why should we think that such bridge laws seem present everywhere but the brain?

In the case of the knowledge argument, we are forced to believe that the ultimate neuroscientist can not know what “red” is. However, even today’s neuroscientists know a great deal about the human vision system.

Furthermore, if experience of red is a particular physical event as is usually thought by neuroscientists, then it is acceptable that reading a book might not result in the same brain activity. Given appropriate apparatus, however, the neuroscientist might be able to invoke any color sensation. Today, artificial vision is a reality. As science advances, arguments like the knowledge argument lose their ontological appeal, although they may be useful in illustrating the difference between subjective experience and objective knowledge.¹⁰

4. Mild versus extreme non-reductionism

The complexity view encourages us to leave behind certain tendencies of classical science.¹³ One such tendency may be called extreme reductionism. We often see observations of complex systems such as human societies being fitted to linear or quadratic equations, disregarding complex phenomena like phase transitions. An extreme inductive bias towards simplicity is not uncommon among scientists. Classically, everything must have a clear-cut, simple and perfect explanation, such as the law of motion $F = ma$. Today, we see that even in physics, the facts are not so simple. The world is imperfect, complex, dynamic and not easily predictable.

However, the alternative to extreme reductionism must not be extreme non-reductionism, which may bring about dualism via either of the non-reductionist arguments in Section 3.1. In fact, dualism is one of the classical tendencies that we must abandon. There are still scientists who believe in an inexplicable soul. On the other hand, many other scientists have broken this psychological barrier and see themselves as part of an evolving, complex nature. Nevertheless, we still need to clarify in what sense our minds are irreducible.

4.1. Degrees of irreducibility

In Section 2, I have reviewed the information theoretic definition of complexity. Algorithmic complexity of a system is a matter of degree. However, the system is fully reducible to small physical interactions. This is the view of *mild non-reductionism*. Indeed, viewing a complex system as consisting of local interactions was a notable achievement of complexity sciences. In the sense of supervenience physicalism, the properties of such a complex system supervene on the local interactions, and of course the structure of these interactions and the components as discussed in Section 2.3, all of which are physical.

An advantage of mild non-reductionism is that it is fully compatible with the mechanical view of the nature. Algorithmic complexity clarifies our intuitions about reducibility. Ordinarily, we view reducibility as an all or none issue. However, reducibility does not make sense without certain important variables. Reduction to what, by what amount, and in many cases, by which method? A physicalist would think that everything, including mental states, are fully reducible to physical states. However, in the light of Section 2.2 the physicalist would also agree that we can reduce *only so much*. If we would like to store all the information about complex entities such as one's psychology, or a living cell, then we cannot hope to compress all this information beyond an optimally small volume of physical space.

4.2. *Consequences of irreducibility*

If we were to take away only a single piece of information from this optimal description, the reproduced system may degrade in an important way or stop functioning altogether, because there is no fault-tolerance in an optimal description. The minimal program for anything is irreducible. When we look at it, it is a random arrangement of symbols. Yet, it holds crucial information. Note that, if something is random, there can be no more analysis of it. Thus, *all* regularities in the data have been found in a minimal description. If of course we knew all the regularities in the data, this also means, intuitively, that we can realize the best prediction. The analysis has occurred to the maximum limit in the process of *compression* if we could have achieved maximal compression.

Unfortunately, in the real-world, we can never guarantee that we have found a minimal program, or we can never know how close we are to a minimal program. Both of these have turned out to be uncomputable, and we finite beings have to be satisfied with approximations. Beyond the uncomputability of algorithmic complexity there are incompleteness results that place serious limits to analysis which have been rigorously investigated by Chaitin.⁶ However, the incompleteness results in mathematics do not show that it is impossible for scientific inquiry to be successful in general. Rather, they deal in what a machine can/cannot infer from given axioms. On the other hand, it has been possible for scientists to *discover* new postulates such as that the speed of light is constant. As a consequence, we can never be sure whether we have the most elegant theory in any domain. Intuitively, we can imagine that the more the complexity, the more effort it will take

to compress/analyze the data. Then, we need to diminish complexity.

4.3. The need for abstract talk

We can not demand that all sciences are reduced to physics before we begin research. The possibility of reduction to physics is not an issue, the feasibility is. We have to talk at a higher level of abstraction that makes the problem smaller, because we often cannot face it in all its glory.

Let us turn our attention to the problem of piecewise analysis and analysis at different levels of description. Often, all the details about a system are not known, or are irrelevant to our purpose. Then, we must have a way of dealing with the relevant aspects of the system. For instance, one might inquire the data transfer rate of the brain, if we are so certain that it is a computer. We then need a reliable way to choose a level of abstraction in which it becomes possible to talk meaningfully about the proposition. That is to say, it should be neither a lower nor higher level of abstraction than necessary. In the case of the data transfer question, we might like to retain only the information about data transmission capabilities of individual links and the network topology. We do not need to concern ourselves with philosophical puzzles to work at this level. In general, we would only like to deal with information that allows us to predict a variable of our choice.

There is another reason for abstract talk. In a mechanical universe, we may expect the probability of a rule to be inversely related to its complexity. That is to say, we do not expect arbitrarily complex rules to pop up. In the theory some explanations have been given as to why Occam's razor holds;¹⁹ our natural way of beginning with simple theories and progressing towards more complex theories seems to have a sound mathematical basis. This would also explain why random events are not common. If randomness were the common case, then our brains could make little use of all the information about the past they are storing. Pure randomness means precisely that one cannot use previous observations to predict the future. However, this is not the case, we routinely predict and control events around us. This can happen only because their algorithmic complexities are small. In other words, randomness is bounded. Mountains do not appear at once, and weather does not reverse every second. While in dynamical systems, a view is that butterfly effects do occur, in the cosmic scale everything is pretty much uniform. It is a fortunate fact that this predictability also occurs to a healthy extent at the level of human beings. That is, mild non-reductionism also leads to mild indeterminacy. It is not productive to

imagine that the macro-scale world is as unpredictable as the most random quantum events. For sufficient prediction, a simple abstract theory might be found by induction.

Another reason for abstract talk comes from the question: reduction to what? When we say reduction, we do not mean reduction to an alien language that we do not understand. Instead, we mean a satisfactory explanation in familiar terms. That is to say, if we know x about the world, we would like the new explanation y to be *relative* to this knowledge, i.e., that it is a *short* explanation in terms of x . We demand this, because otherwise we cannot integrate y into our body of knowledge. This constraint is captured well by the notion of conditional algorithmic complexity. It is possible that an alien creature may have developed a quite different body of knowledge, and we will have to spend a lot of effort to decypher its musings, i.e., the meaning would have large conditional complexity relative to our culture.

An interesting question is the autonomy of levels in nature, e.g., physics and biology. A relevant idea in algorithmic complexity is the notion of algorithmic independence (lack of mutual information). If through observations, a brain can detect “islands” of information, then it can exploit this property.⁷ Such a capacity for abstraction from sensory information seems to be present in our brains, however we can make few speculations on its properties, perhaps except for Bialek’s information theoretic approach.²

4.4. *Algorithmic complexity and emergence*

We are now in a better position to review the idea that animate-ness of a cell is more than the sum of its inanimate parts. Or similarly, that a brain is more than the sum of its neurons. I have already argued that reduction to a physical description is necessary for the scientifically minded complexity researcher who rejects ontological dualism. First, let us have a look at the meaning of such predicates like animate or conscious that denote complex entities. These have been most likely invented for the purpose of avoiding *excessive explanations*, i.e., to be able to talk at a correct level of abstraction. When we use these terms, we do not feel the need to explain them further, however we can all recognize their presence through their outward appearances. For instance, all living things show some qualities that we can easily recognize, and thus use to detect the presence of this type of complex system. Furthermore, in many common sense problems, we need no more information about such words except for the basic function of

distinguishing a dead animal from a living animal, etc. As countless works in philosophy show, it takes a great effort to do otherwise and try to give a complete description.

On the other hand, the animate-ness of a cell *supervenes* on continuing success of a number of complex subsystems that are tightly interconnected. Should one of these crucial systems fail, the cell is no more animate. While like any other machine, the description of the cell is *compositional*, the topology and nature of interconnections are non-trivial.

The outward appearances of complex phenomena are usually not satisfactory for a scientist. Complexity community was not satisfied with the definition of “animate” that it meant being able to reproduce or having genetic material. They rightly sought generic properties of the systems that we call animate. For instance, that they are stable, yet have dynamics at the edge of chaos, that they are made up of parts that are being replaced over time. Attempts have been made to characterize not only earth biology, but any animate matter, and further efforts have been made to characterize evolution. For instance, Chaitin has given a rough mathematical definition of life ⁵ where he suggests that the interacting parts of such a system (similar to components of a system in Section 2.3) are self-similar, i.e., they have high mutual algorithmic complexity. He also suggests the famous halting probability Ω as an abstract model of evolution; the approximation procedure to calculate it is presented as a mathematical model of entire evolution.

The very existence of these generic explanations in complexity sciences suggests that the concepts of animate-ness and evolution may be given *abstract* descriptions that are independent of particular instances. It may be possible to extract the essential commonality in all living systems into a clear description by choosing an affordable level of explanation. These descriptions may turn out to be much simpler than the dynamics of a cell which is a complex machine. The architecture of the cell is likely impossible to infer from the history of evolution prior to the cell, that is to say it contains novel algorithmic information relative to that history. Given the “inanimate” components of a cell, we can construct a large number of systems, however, only a few of them will result in proper function. The algorithmic complexity of the fit organization relative to the set of components may be interpreted as the new “valuable” or “emergent” bits invented by evolution.

5. Conclusion

The physical versions of algorithmic complexity help us consider it in a purely physical way, and free of problems in the abstract version. I have shown that algorithmic complexity correspondence to our intuition about complexity, and further clarifies it with theorems. In Section 3 some dangers of extreme non-reductionism have been shown by examining the knowledge argument and predicate dualism which rest on epistemological and absolute irreducibility. These arguments have been criticized as arguments from ignorance. With the same attitude, we would have hardly advanced science to this point. In Section 4, I propose using precise and quantitative conceptions of complexity as they clarify our intuitions. Irreducibility of a system comes in degrees, and it may be understood physically. Randomness means the absence of further reducibility, and we cannot reason about random events. The reason for abstract talk is examined as a matter of processing feasibility, purpose, inductive reasoning and opportunity. Finally, instead of stating emergency of a property X as “ X is more than sum of its parts”, I suggest that we can use the precise language of information theory. In particular, “emergent property” can be given a satisfactory meaning in AIT, corresponding to algorithmic information in a system that does not reside in its components or prior history, thus is impossible to predict using the information therein.

Mild non-reductionism cannot be the final word on this controversial subject at least for the reason that algorithmic complexity has obvious shortcomings. While irreducibility is so important, our intuitions may not suffice. Algorithmic complexity helps us sharpen our methodology with mathematical reasoning, and sheds new light on concepts such as irreducibility, limits to analysis, predictability and emergence. In the future, more theoretical work connecting information theory and dynamical systems theory would be valuable for complexity researchers.

Acknowledgments

Thanks to anonymous reviewers, Peter D. Jones, Peter Jedlicka, Michael Olea, William Modlin, Marvin Minsky, and Gregory J. Chaitin for their comments which helped improve this paper.

References

1. C.H. Bennett, P. Gacs, M. Li, P.M.B. Vitanyi, and W. Zurek. Information distance. *IEEE Transactions on Information Theory*, 44(4):1407–1423, 1998.

2. William Bialek. Thinking about the brain. In H Flyvbjerg, F. Jlicher, P. Ormos, and F. David, editors, *Physics of Biomolecules and Cells: Les Houches Session LXXV*, pages 485–577. EDP Sciences, Les Ulis; Springer-Verlag, Berlin, 2002.
3. William Bialek, Ilya Nemenman, and Naftali Tishby. Predictability, complexity and learning. *Neural Comp*, 13:2409–2463, 2001.
4. Ned Block and Robert Stalnaker. Conceptual analysis, dualism, and the explanatory gap. *Philosophical Review*, 108:1–46, 1999.
5. Gregory J. Chaitin. Toward a mathematical definition of life. In R.D. Levine and M. Tribus, editors, *The Maximum Entropy Formalism*. MIT Press, 1979.
6. Gregory J. Chaitin. *Algorithmic Information Theory*. Cambridge University Press, 1987.
7. Gregory J. Chaitin. Two philosophical applications of algorithmic information theory. In *Proceedings DMTCS'03 Springer Lecture Notes in Computer Science*, volume 2731, pages 1–10, 2003.
8. Paul S. Churchland. The hornswoggle problem. *Journal of Consciousness Studies*, 1996.
9. Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*, chapter 1. Wiley Series in Telecommunications. Wiley-Interscience, 1991.
10. Tim Crane. *Real Metaphysics*, pages 68–83. Routledge, London, 2003.
11. Donald Davidson. Mental events. In L. Foster and J. W. Swanson, editors, *Experience and Theory*. 1971.
12. Carlos Gershenson and Francis Heylighen. When can we call a system self-organizing? In W Banzhaf, T. Christaller, P. Dittrich, J. T. Kim, and J. Ziegler, editors, *Advances in Artificial Life, 7th European Conference, ECAL 2003 LNAI 2801*, pages 606–614. Springer-Verlag, 2003.
13. Carlos Gershenson and Francis Heylighen. How can we think the complex? In Kurt Richardson, editor, *Managing Organizational Complexity: Philosophy, Theory and Application*, chapter 3. Information Age Publishing, 2005.
14. Marcus Hutter. *Universal Artificial Intelligence: Sequential Decisions Based on Algorithmic Probability*. Springer-Verlag, 2005.
15. Dan Hutto. Davidson’s identity crisis. *Dialectica*, 52(1), 1998.
16. Frank Jackson. Epiphenomenal qualia. *Philosophical Quarterly*, 1982.
17. Marvin Minsky. *The Emotion Machine*, chapter 8. Draft. <http://web.media.mit.edu/~minsky/E8/eb8.html>.
18. N. Tishby FC Pereira and W Bialek. The information bottleneck method. In *Proceedings of the 37th Annual Allerton Conference on Communication, Control and Computing*.
19. Ray Solomonoff. The application of algorithmic probability to problems in artificial intelligence. In M. Kochen and H. M. Hastings, editors, *Advances in Cognitive Science, AAAS Selected Symposia Series*, pages 210–227. 1988.