Understanding Nonmodular Functionality: Lessons from Genetic Algorithms

Jaakko Kuorikoski and Samuli Pöyhönen*†

Evolution is often characterized as a tinkerer creating efficient but messy solutions. We analyze the nature of the problems that arise when trying to explain and understand cognitive phenomena created by this haphazard design process. We present a theory of explanation and understanding and apply it to a case problem—solutions generated by genetic algorithms. By analyzing the nature of solutions that genetic algorithms present to computational problems, we show, first, that evolutionary designs are often hard to understand because they exhibit nonmodular functionality and, second, that breaches of modularity wreak havoc on our strategies of causal and constitutive explanation.

1. Introduction. The once dominant classical paradigm of cognitive science has been under attack for several decades. Connectionism, cognitive neuroscience, dynamical systems theory, and new robotics have all questioned whether the classical artificial intelligence (AI) approach to cognition can adequately describe biologically evolved cognitive systems such as human minds. Whereas classical AI tends to approach computational problems with functional decompositions inspired directly by the programmer's intuitions about possible efficient subroutines, the alternative research programs often emphasize that biological evolution is more likely to produce far more complex and messy designs.

In this article, we analyze the nature of the problems that these messy solutions raise to the understanding of cognitive phenomena. In general, the problem of understanding nonintuitive designs produced by natural selection

*To contact the authors, please write to: Jaakko Kuorikoski, Social and Moral Philosophy, PO Box 24 00014, University of Helsinki, Finland; e-mail: jaakko.kuorikoski@helsinki.fi .Samuli Pöyhönen, e-mail: samuli.poyhonen@helsinki.fi

†The authors are listed in alphabetical order. Previous versions of this article were presented in the POS seminar and at the Models and Simulations 5 conference in Helsinki. We thank all the participants, as well as the audience at the PSA, for their valuable comments.

Philosophy of Science, 80 (December 2013) pp. 637–649. 0031-8248/2013/8005-0012$10.00
Copyright 2013 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/673866 Published online by Cambridge University Press

is well known in philosophy of psychology (e.g., Clark 1997, chap. 5), philosophy of biology (Wimsatt 2007), and now even in popular psychology (Marcus 2008), but the problem has proved to be difficult to articulate without a clear idea of what exactly it is that evolutionary tinkering is supposed to hinder. The main challenge for understanding is often framed and explained by pointing to the path-dependent nature and the resulting unfamiliarity of the evolved design (Jacob 1977). We argue that this is not the whole story. We hope that providing an explicit theory of explanation and understanding will move us beyond intuitions toward a more systematic analysis and, ultimately, concrete solutions. We also combine our theory of explanatory understanding with a computational application of evolutionary design: problem solutions generated by genetic algorithms. By analyzing the nature of solutions that genetic algorithms offer to computational problems, we suggest that evolutionary designs are often hard to understand because they can exhibit *nonmodular functionality* and that breaches of modularity wreak havoc on our strategies of causal and constitutive explanation.

2. Explanation and Understanding. The ultimate goal of cognitive (neuro)-science is to provide mechanistic understanding of system-level properties of the cognitive system in terms of the properties of its parts and their organization. Probably the most developed account of the general strategies for reaching such mechanistic understanding is William Bechtel's and Robert Richardson's (2010) study of the heuristics of decomposition and localization (DL). The DL procedure goes roughly as follows. First, the different phenomena that the system of interest exhibits are differentiated. Then the phenomenon of interest is *functionally decomposed*, that is, analyzed into a set of possible component operations that would be sufficient to produce it. One can think of this step as the formulation of a preliminary set of simpler functions that, taken together, would constitute the more complex input-output relation (the system-level phenomenon). The system is also *structurally decomposed* into a set of component parts. The final step is to try to *localize* the component operations by mapping the operations onto appropriate structural component parts. The idea is thus to, first, come up with a set of more basic properties or behaviors that could, taken together, possibly result in the explanandum behavior and, then, try to find out whether the system is in fact made of such entities that can perform the required tasks. If this cannot be done, the fault may lie with the functional and structural decompositions or with the very identification of the phenomenon, and these may then have to be rethought. The identification and decomposition procedures will in the beginning be guided by earlier theories and common sense, but empirical evidence can always suggest that a thorough reworking of the basic ontology and the form of the possible explananda may be in order.

https://doi.org/10.1086/673866 Published online by Cambridge University Press

According to Bechtel and Richardson, decomposability is a regulative ideal in such model construction because complex systems are psychologically unmanageable for humans. Decomposition allows the explanatory task to be divided into parts that are manageable for cognitively limited beings, thereby rendering the system intelligible (Bechtel and Richardson 2010). The idea comes originally from Herbert Simon (1962), who claimed that the property of *near decomposability* is a necessary condition of understandability to any finite cognitive agent. Near decomposability means that the system can be decomposed into parts in such a way that the intrinsic causal properties of the parts are more important for the behavior of the system than the relational causal properties of the components that are constituted also by their environment and interaction. Near-decomposable systems are thus hierarchical in the sense that the complex whole can be seen as made from a limited set of simpler parts and interactions. Hierarchical systems are manageable for cognitively limited beings because their 'complete description' includes recurring or irrelevant elements describing similar recurring parts and nonimportant interactions. The removal of such descriptions does not hamper our understanding of the system and thus eases cognitive load.

Although there are a number of arguments that conclusively show that such informational economy by itself is not constitutive of understanding, we agree with Simon in that a property closely related to near decomposability, namely, *modularity*, is a necessary condition for understanding.[1] As a conceptual starting point for our argument, we conceive understanding not as a special mental state or act. Rather, because understanding is a public, behavioral concept, we see it as a regulative label, which is attributed with regard to manifest abilities in everyday action and correctness of reasoning. Cognitive processes (comprehension) taking place in the privacy of individual minds are a causal prerequisite for possible fulfillment of these criteria, but the processes themselves are not the facts in virtue of which somebody understands or not. They are not the criteria of understanding in the sense that we would have to know them in order to say whether somebody really understands something (Ylikoski 2009; Ylikoski and Kuorikoski 2010).

We take the primary criterion of understanding to be inferential performance: whether someone understands a concept is evaluated according to whether he or she can make the right inferential connections to other concepts. Likewise, whether someone understands a phenomenon is assessed on the basis of whether he or she can make correct inferences related to it.

1. First, nobody has actually given a positive argument for equating understanding with increased informational economy. Second, successful classification schemes and statistical techniques do compress information but are usually taken to be merely descriptive and not explanatory (Woodward 2003, 362–64).

https://doi.org/10.1086/673866 Published online by Cambridge University Press

This view can be further developed by linking it to James Woodward's account of scientific explanation (2003) in the following way: Woodward's theory of explanation tells us more specifically what kinds of inferences are constitutive of specifically explanatory understanding. Explanation consists in exhibiting functional dependency relations between variables. Knowledge of explanatory relationships facilitates understanding by implying answers to what-if-things-had-been-different questions concerning the consequences of counterfactual or hypothetical changes in the values of the explanans variable. Whether someone understands a phenomenon is evaluated according to whether he or she can make inferences not only about its actual state but also about possible states of the phenomenon or system in question. In the case of causal explanations, these explanatory dependencies concern the effects of interventions, and causal knowledge thus enables the goal-directed manipulation of the object of explanation. These answers are the basis of the inferential performance constitutive of understanding.

The limits of inferential performance depend causally on the features of the reasoning processes of the agents whose understanding is being evaluated. Thus, the limits of understanding are dependent on the cognitive makeup of agents and can certainly be investigated psychologically. For example, if the space limit of our working memory is indeed roughly seven items, then this sets an upper boundary for the complexity of our inferences and, consequently, for our understanding.

In order for answers to what-if questions to be well defined, the dependencies grounding the answers have to possess some form and degree of independence such that a local change in an aspect of the phenomenon under study cannot ramify uncontrollably or intractably. If local modifications in a part of a system disrupt other parts (dependencies) in a way that is not explicitly specified (endogenized) in the (internal or external) representation of the system according to which the what-if inferences are made, the consequences of these changes are impossible to predict and counterfactual assertions impossible to evaluate. Things participating in the dependency relations also have to be somewhat localized (physically or conceptually) in order for the contemplated changes to be well defined in the first place (Woodward 2003, 333). Therefore, a necessary condition for a representation to provide understanding of a phenomenon is that the modularity in the representation matches the modularity in the phenomenon.

Let us first discuss the case of causal understanding. If an intervention on a causal system actually changes the system in a way that is not represented in the model of the system, the model as it stands does not give correct answers to what-if-things-had-been-different questions concerning the state of the system after the intervention. If we intervene on a causal input corresponding to variable X_i in a model and the intervention, no matter how

https://doi.org/10.1086/673866 Published online by Cambridge University Press

surgical, also changes the dependencies within the system or values of other variables themselves affecting variables causally downstream of X_i, the model does not give correct predictions about the consequences of the intervention. Hence, the model does not provide correct causal understanding of the workings of the system and the causal role of the variable in it. If the system cannot be correctly modeled on any level of description or decomposition so that it is modular in such a way—if the system itself is not causally modular—no what-if-things-had-been-different questions concerning interventions in the system can be answered. If the system is in fact such that every local change brings about intractable changes elsewhere in the system to such an extent that there can be no representation that would enable a cognitively finite being to track these changes and make correct inferences about their consequences, then the system is beyond the limits of understanding.

The problem of understanding causally nonmodular systems has received some attention in the philosophy of science literature (e.g., Bechtel and Richardson 2010, chap. 9). However, according to the schema of Bechtel and Richardson, before we can even start thinking about searching for causal-mechanical understanding of the system realizing the complex behavior to be understood, we need to formulate hypotheses about the possible functional decompositions of the behavior (see also Cummins 1983). For example, what kind of simpler subtasks could possibly produce complex cognitive capacities such as language production and comprehension, long-term memory, and three-dimensional vision? Importantly, this task is separate, although not independent, from hypotheses concerning the implementation of the capacity. Although the understanding offered by the functional decomposition is not, strictly speaking, causal—component operations do not cause the whole behavior because they are constitutive parts of it[2]—the modularity constraint on understandability still applies in the following way. We can only understand the complex behavior by having knowledge of its component operations, if we can make reliable what-if inferences concerning the possible consequences of changes in the component operations for the properties of the more complex explanandum capacity.

For example, we provisionally understand working memory if we can infer from possible changes in its hypothesized component operations (such as differences in the postulated phonological loop or episodic buffer) to changes in the properties of the capacity. These inferences are only possible

2. Although we fully agree with Piccinini and Craver (2011) in that insofar as functional decompositions are explanatory, they are to be thought of as mechanism sketches and that the functional hypotheses are not independent of the question of mechanistic implementation.

https://doi.org/10.1086/673866 Published online by Cambridge University Press

if the functional decomposition itself is suitably modular; that is, the consequences of "local" changes in component operations do not ramify in an intractable way, making the behavior of the whole completely holistic. We now argue that genetic algorithms demonstrate that design by selection can lead to such nonmodular complex behavior.

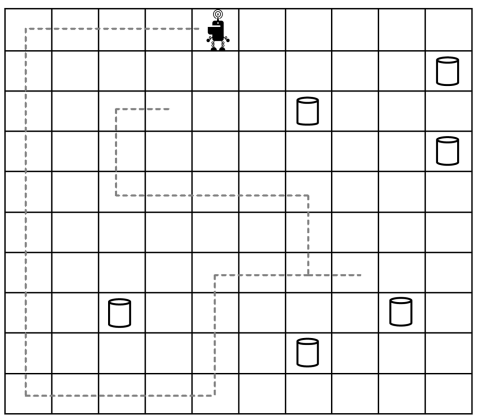
3. Genetic Algorithms. Since the 1960s, there have been attempts to apply insights from evolutionary thinking to computer programming. Here we discuss one genre of evolutionary programming: genetic algorithms (see Holland 1975; Goldberg 1989; Mitchell 1996). In a nutshell, the idea of the genetic-algorithms approach is to "breed" randomly generated solutions to computational problems. This is done by mimicking the evolutionary mechanisms of inheritance, mutation, selection, and crossover in a computer simulation. Although genetic algorithms (GAs) are not the only strand of evolutionary programming, they serve our purpose well because their basic principles are easy to understand and they are the most well-known kind of evolutionary programming outside computer science (Clark 1997, 2001; Mitchell 2009).

From the point of view of AI, GAs are a form of nonexhaustive but massively parallel search in the search space of a problem. They can be used for a number of different purposes: for evolving behavioral strategies for simulated agents, for finding weights in a connectionist network, or for evolving cellular automata to perform computations. We illustrate the nature of GAs by presenting a simple example from Melanie Mitchell (2009, chap. 9). Mitchell's original simulation showed how GAs can be used to evolve a controlling program for a simulated robot picking up soda cans on a 10×10 grid. Robby the robot can only see the squares adjacent to its location (center, north, south, east, west), and in each turn it can either move one step to a particular direction, move at random, try to pick up a can, or do nothing. Each simulation run lasts for a predetermined amount of time steps (originally 200), and Robby's task is to pick up as many randomly situated soda cans as possible.

Initially a random population of software individuals is generated, each with a "genome" consisting of 243 random numbers. Each locus in the genome guides Robby's behavior in a particular situation (fig. 1). The fitness score of each candidate program in the population is calculated by running several simulation trials: crudely, the more cans the robot is able to pick up on average, the higher its fitness. Programs with the highest fitness are then used to form the next generation of programs: they are paired randomly, and the genomes of the two parents are crossed over at a randomly chosen point to create the genomes of new individuals. Finally, for each descendant, there is a small probability (.005) for each locus that a mutation occurs in it. As a result, the new generation is based on the most successful variants among

https://doi.org/10.1086/673866 Published online by Cambridge University Press

https://doi.org/10.1086/673866 Published online by Cambridge University Press



Genome G:
[1, 0, 3, 2, 0, 2, 0, 0, 3, 6, 6, 6, 6, 2, 2, 6, 6, 6, 0, 2, 5, 7, 3, 8, 5, 10, 8,
3, 0, 3, 3, 2, 10, 3, 0, 3, 6, 6, 6, 2, 2, 2, 6, 6, 6, 2, 8, 7, 6, 4, 9, 9, 6, 6,
2, 1, 1, 2, 2, 0, 7, 0, 1, 6, 0, 6, 6, 5, 8, 4, 0, 6, 7, 9, 8, 9, 6, 6, 2, 1, 0, 1,
4, 1, 2, 0, 2, 6, 2, 7, 6, 0, 6, 6, 6, 6, 1, 1, 3, 1, 4, 8, 8, 9, 4, 8, 6, 9, 3, 0,
1, 3, 2, 4, 3, 2, 9, 6, 6, 6, 2, 2, 4, 6, 1, 4, 0, 7, 4, 6, 3, 7, 2, 2, 10, 1, 4,
1, 5, 7, 8, 9, 0, 5, 6, 0, 6, 2, 7, 2, 7, 6, 2, 10, 8, 1, 1, 5, 10, 0, 2, 6, 2, 0,
3, 2, 0, 7, 0, 0, 9, 6, 6, 4, 6, 6, 1, 6, 6, 8, 8, 6, 7, 8, 5, 5, 8, 9, 4, 3, 4, 9,
3, 3, 6, 0, 0, 10, 6, 0, 9, 6, 10, 4, 10, 3, 9, 3, 10, 4, 5, 5, 7, 8, 7, 1, 2, 6,
5, 8, 5, 10, 7, 6, 10, 6, 0, 6, 2, 2, 5, 9, 2, 8, 2, 6, 3, 1, 5, 6, 8, 6, 2]

Figure 1. Each “locus” in the genome G corresponds to one of the possible immediate environmental states of Robby, and each digit (the allele), to a move in that situation (e.g., 0 → move north, 5 → pick up; see Mitchell 2009, 137).

the previous generation, and the process loops back to the fitness-calculation phase. Thus, the GA continues searching for efficient solutions by charting new regions of the search space.

After a few hundred generations, the evolved strategies start to achieve impressive results. As we replicated Mitchell's simulation, we observed that after the eight hundredth generation, the best strategies among evolved Robbys started to have higher fitness scores than a simple "rational" solution programmed by a human designer (ultimately 480 vs. 440 points).[3] However, although solutions found with GAs are efficient, their behavior is often hard to understand. The ingenious heuristics that the programs employ cannot be deciphered by simply looking at individual genes or sets of genes. Instead, it is necessary to look holistically at the broad phenotypic behavior of the robot. A nice illustration of this impenetrability of such evolved solutions is the fact that in some cases when a highly evolved Robby is in the same square with a can, it decides not to pick it up but rather chooses to move away from the square. While this behavior seems prima facie irrational, looking at the total behavioral profile of the robot uncovers a cunning strategy: Robby uses cans as markers to remember that there are other cans on its side and explores the adjacent squares for extra cans before picking up the marker can. Thus, by not treating cans only as targets but also as navigational tools, Robby uses its environment to extend its severely limited visual capacities and to compensate for its total lack of memory.

Moreover, by examining the behavior of a fifteen hundredth generation Robby with the highest fitness score in its population, it can be seen that the marker strategy manifests in slightly different ways in different environmental situations. It is therefore not a discrete adaptation but rather a collection of independently evolved substrategies. Furthermore, the marker strategy is tightly intertwined with another environment-employing "hack" that the sophisticated Robby uses: when there is already a lot of empty space on the grid, Robby employs a "vacuum cleaner" movement strategy. It follows the walls of the board, departing toward the center when it detects a can, employs the marker strategy if possible, and immediately after cleaning up its local environment returns directly to the south wall to continue its round around the board.

Such "kluges" are common to designs created by GAs. Like biological evolution, GAs can come up with solutions that a human designer would not usually think of. These solutions often off-load parts of problem solving to the environment and, thus, rely on a tight coupling between the system and its environment. And, as pointed out by Clark (1997, 2001), recurrent circuitry and complex feedback loops between different levels of process-

3. Code obtainable on request.

https://doi.org/10.1086/673866 Published online by Cambridge University Press

ing often feature in systems designed by GAs. Such designs are often difficult to understand. We suggest that such difficulties in understanding are often created by the lack of modularity in the functional decomposition of the behavior. Our most successful Robby (genome G in fig. 1) is leaving cans as markers only in specific situations, and only the totality of this selective marking strategy—together with navigational strategies using cans and walls—constitutes the effectiveness of the can-search procedure. Looking at isolated genes in Robby's genome only reveals trivially modular elements corresponding to elementary subtasks in Robby's behavior: one gene corresponds to an elementary move in a specific environmental situation. But we cannot make inferences from local hypothetical changes in these elemental behaviors to consequent effects on fitness. The connection between any single elementary behavioral rule and the strategy is simply too complex and context dependent. A change in a single rule (in situation B and a can present, whether to pick up or not to pick up the can) has consequences for the effectiveness of the other elementary behavioral rules. Explanatorily relevant inferences would require an extra "level" of modular suboperations between the individual movements and the strategy as a whole. The marker and vacuum cleaner strategies mentioned above appear to be examples of such middle-level suboperations, but by themselves they are insufficient to yield understanding of the whole behavior of our most successful Robby since the effectiveness of leaving a can is a result of the evolved match between the specific situations in which Robby leaves a can and the rest of the navigation behavior. And GAs do not, in general, produce such easily discernible designs. Rather, the interesting heuristics in the system's behavior can only be revealed by simultaneously looking at constellations of different genes and, eventually, the whole genome.

To recapitulate, our example exhibits several distinct (yet related) challenges to understanding:

1. The discernible middle-level strategies (marker, vacuum cleaner) do not have a dedicated structural basis. Instead, the nature of the design process leaves all atomic structural elements (the 243 DNA elements) open for exploitation by all capacities serving the main goal. In consequence, the system is neither structurally nor behaviorally nearly decomposable but instead has a "flat hierarchy," and strategies are implemented in highly distributed structures.
2. Challenge 1 above means that the interactions between subtasks tend to be strong: a change in one subtask constituting part of the marker behavior also affects the functioning of the vacuum cleaner navigation. In general, the middle-level strategies can only be discerned and defined in a very abstract way, and the interaction effect on their con-

https://doi.org/10.1086/673866 Published online by Cambridge University Press

tribution to the overall fitness is so large as to make any inferences about the consequence of partial changes in one strategy next to impossible.

3. The way in which operations contribute to the fitness of the individual is highly context dependent and depends on the properties of the environment as well as the DNA of the agent. For instance, merely detecting the existence of the marker strategy requires that there are suitable clusters of cans in the environment. Moreover, even small modifications to the environment can lead to drastic changes in the performance of a strategy. For instance, adding only a few randomly placed extra walls on the grid radically collapses the average score of the successful Robby described above.

Extrapolating from this very simple case, GAs may design behavior that cannot be tidily decomposed into hierarchical and modular subtasks, whose individual contributions would be easy to understand (i.e., we could infer how a change in a subroutine would affect the behavior of the mother task). Instead, feedback, many tasks using the same subtasks as resources, and system-environment coupling lead to holistic design where almost "everything is relevant for everything." The evolved functional architecture is flat in that there are few discernible levels of order between the elementary operations and the complex whole. The counterintuitiveness of such flat architectures is apparent in the deep mistrust faced by connectionist suggestions for nonhierarchical design of cognitive capacities (see, e.g., Rumelhart and McClelland 1986 vs. Pinker and Prince 1988).

Furthermore, GAs underscore the path dependence of evolutionary problem solving. For sufficiently complex computational problems, there are often several local maxima in the fitness landscape of the problem, and the population can converge to different maxima in different runs of the simulation. The functional decomposition that a human designer comes up with is just one possible solution among several others. Perhaps our biological evolution actually ended up with a radically different one.

4. Lessons for the Study of Mind. GAs seem to demonstrate that evolution can, in principle, lead to nonmodular functionality. This imposes a limit on our ability to understand such behavior: if we cannot trace the consequences of changes in the suboperations, we cannot answer what-if questions concerning the complex behavior. Such behavior also constitutes a thorny problem for mechanistic understanding of the implementation of the said behavioral capacities since the DL heuristic cannot even get off the ground. We can now ask two questions: Should we expect to find such nonmodular functionality in nature, especially in human cognition, and if so, what attitude should we adopt with respect to this problem? Should the aim

https://doi.org/10.1086/673866 Published online by Cambridge University Press

of causal-mechanistic understanding of the brain be given up and replaced with a program of instrumentally interpreted dynamical models and modeling the dynamics of the mind with a few macrovariables?

There are important disanalogies between GAs and biological evolution. First, in GAs, there usually is no genotype-phenotype distinction. In biological evolution, however, genes do not directly cause properties of the phenotype but rather participate in guiding ontogenesis. There have been suggestions that ontogenesis itself favors modular design. GAs may also seem a problematic platform for exploring the possibilities of DL heuristics since the lowest level of functional organization and the level of implementation are identical (i.e., the genome). However, we see no reason why this would affect our argument. Moreover, the argument developed here is about selection in general, and failures of functional modularity may in principle also arise in the course of development—at least if the idea of neuronal group selection or "neural Darwinism" is taken seriously.

Second, most studies on GAs are carried out by using a single fixed goal or a fixed task type. In the Robby example, although the distribution of the cans was generated at random, the task itself remained essentially the same from generation to generation. However, Nadav Kashtan and Uri Alon (2005; see also Kashtan, Noor, and Alon 2007) have demonstrated that when the goals themselves are composed of modularly varying subgoals, evolution produces modular functionality. It is easy to see why this is the case. If the tasks to which the system has to adapt remain the same, the selection environment is stable, and the peaks in the fitness landscape are immovable, then selection favors strategies that off-load problem solving to that particular environment as much as possible. But if the task itself is composed of changing subtasks, it makes sense to design the adaptive response in such a way that a particular suboperation can locally adapt to a local change in a subtask without altering the totality of the otherwise well-functioning behavior.

It seems likely that our cognition has evolved in just such a modularly changing selection environment, but the extent to which we should expect to find modular functionality in human cognition is hard to estimate and is most probably a purely empirical matter. Moreover, as a response to Simon's (1962) Tempus and Hora argument, it has been argued that componential specialization in complex systems is a force that works against the development of strictly modular structures (e.g., Levins 1973; Wimsatt 2007, 186–92). Because of the uncertainty related to the actual extent of nonmodularity in human cognition, we stress the conditional nature of our argument.

GAs demonstrate that evolution can create designs that are in principle beyond the understanding of unaided cognitive beings such as us. Yet there is nothing mysterious in such designs. Simon pondered whether the appar-

https://doi.org/10.1086/673866 Published online by Cambridge University Press

ent abundance of hierarchical, nearly decomposable complexity was due to our selective attention to precisely such systems, but we believe this to be a somewhat hasty conjecture. We have no trouble finding and delineating systems, such as Robby or possibly ourselves, with behaviors that are functionally nondecomposable and constituted by a flat architecture. However, there certainly might be a psychological bias that makes us see hierarchical design also where there is none. One way of coping with this impasse is to realize that there are no fundamental reasons to limit the relevant epistemic agent to be an unaided human. Although only a human agent can experience a sense of understanding, this feeling should not be confused with understanding itself. Therefore, brute computational approaches can produce understanding as long as the understanding subject, the cognitive unit whose inferential abilities are to be evaluated, is conceived as the human-computer pair.

REFERENCES

Bechtel, William, and Robert C. Richardson. 2010. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Cambridge, MA: MIT Press.

Clark, Andy. 1997. *Being There: Putting Brain, Body and World Together Again*. Cambridge, MA: MIT Press.

———. 2001. *Mindware: An Introduction to the Philosophy of Cognitive Science*. Oxford: Oxford University Press.

Cummins, Robert. 1983. *The Nature of Psychological Explanation*. Cambridge, MA: MIT Press.

Goldberg, David E. 1989. *Genetic Algorithms in Search Optimization and Machine Learning*. Reading, MA: Addison-Wesley.

Holland, John. 1975. *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press.

Jacob, Francois. 1977. "Evolution and Tinkering." *Science* 196 (4295): 1161–66.

Kashtan, Nadav, and Uri Alon. 2005. "Spontaneous Evolution of Modularity and Network Motifs." *Proceedings of the National Academy of Sciences of the USA* 102 (39): 13773–78.

Kashtan, Nadav, Elad Noor, and Uri Alon. 2007. "Varying Environments Can Speed Up Evolution." *Proceedings of the National Academy of Sciences of the USA* 104 (34): 13711–16.

Levins, Richard. 1973. "The Limits of Complexity." In *Hierarchy Theory: The Challenge of Complex Systems*, ed. Howard Pattee, 73–88. London: Braziller.

Marcus, Gary. 2008. *Kluge: The Haphazard Construction of the Human Mind*. Boston: Houghton Mifflin.

Mitchell, Melanie. 1996. *An Introduction to Genetic Algorithms*. Cambridge, MA: MIT Press.

———. 2009. *Complexity: A Guided Tour*. Oxford: Oxford University Press.

Piccinini, Gualtiero, and Carl Craver. 2011. "Integrating Psychology and Neuroscience: Functional Analyses as Mechanism Sketches." *Synthese* 183 (3): 283–311.

Pinker, Steven, and Alan Prince. 1988. "On Language and Connectionism: Analysis of a Parallel Distributed Processing Model of Language Acquisition." *Cognition* 23:73–193.

Rumelhart, David, and James McClelland. 1986. "On Learning the Past Tenses of English Verbs." In *Parallel Distributed Processing*, vol. 1, ed. James McClelland, David Rumelhart et al., 216–71. Cambridge, MA: MIT Press.

Simon, Herbert. 1962. "The Architecture of Complexity." *Proceedings of the American Philosophical Society* 106:476–82.

Wimsatt, William. 2007. *Re-engineering Philosophy for Limited Beings*. Cambridge, MA: Harvard University Press.

Woodward, James. 2003. *Making Things Happen*. Oxford: Oxford University Press.

https://doi.org/10.1086/673866 Published online by Cambridge University Press

Ylikoski, Petri. 2009. "The Illusion of Depth of Understanding in Science." In *Scientific Understanding: Philosophical Perspectives*, ed. Henk De Regt, Sabine Leonelli, and Kai Eigner, 100–119. Pittsburgh: University of Pittsburgh Press.
Ylikoski, Petri, and Jaakko Kuorikoski. 2010. "Dissecting Explanatory Power." *Philosophical Studies* 148:201–19.

https://doi.org/10.1086/673866 Published online by Cambridge University Press