# The Rise of the Machine[1]

**Charles W. Lowney**

In order to distinguish ourselves from dead matter and our free actions from predetermined causes, we humans have attempted to draw lines that distance us from our animal bodies and from machines. But machines are learning to perform skills once thought uniquely human—like driving cars or speaking languages. What does this mean for human exceptionalism? Does this mean our skills and intelligence can be reduced to matter in motion? Are these technological advances a victory for material reductionism and causal determinism?

Rather than draw another line in the sand, I suggest we find our freedom and creativity by examining what we share with animals and "intelligent" machines. I will show strong similarities between structures of human tacit knowing and connectionist (aka neural network) architectures, demonstrating how these computational methods and machines mimic an irreducible intentionality seen in tacit body-knowing that is common to many animals and humans.

In looking to our animal bodies, and to computing systems, we find a body-knowing that is irreducible to dead matter and the resists predetermined fate. The course of evolution brought increasing degrees of freedom that allow for moral choices in human minds. If intentionality, creativity and freedom give us special value, then it is a value shared more broadly. From this emergentist perspective, rather than reduce humans to animals, animals to machines, and machines to smallest parts and their laws, we will see more value in animals and higher prospects for Artificial Intelligence.

## 1. Defending a Line: Dismal Prospects for Human Creativity and Freedom?

Are we slaves to determinate physical laws, or irreducibly special in a way that makes us free? We first differentiated ourselves as living beings from dead matter. As we came to see animal bodies as machines, we drew a new line: our reason and language made us special. But computer technology challenged even this distinction.

Phenomenologists like Maurice Merleau-Ponty and Hubert Dreyfus distinguished between body-knowing and explicit propositional knowing—between "knowing how" and

[1] This is an abridged rendition of "The Rise of the Machine: Body-Knowing, Neural Nets, and Emergent Freedom" Carol Vernallis, Holly Rogers, Selmin Kara & Jonathan Leal (eds), *Cybermedia: Explorations in Science, Sound, and Vision*, New York: Bloomsbury Publishing Inc. 2022, pp. 97-127.

"knowing that."[2] This body-knowing displayed creative human engagement resisting reduction to mere matter. We see it in improvisational dance, music, and martial arts.

Dreyfus claimed computers could never drive cars—not due to complexity, but because this relied on irreducible human skills that couldn't be laid out as explicit instructions computers require.[3] Collins drew a different line: Collective Tacit Knowledge (CTK) cannot be reduced. He also argued computers would never drive cars in traffic to destinations, seeing this as social knowledge irreducible to any "string" or algorithm.[4]

Collins concedes even more ground to material reductionism. He claims that Relational Tacit Knowledge (RTK), which includes skill development learned through apprenticeship, is a form of tacit knowing that can ultimately be made explicit through further inquiry. And he concludes that body-knowing or Somatic Tacit Knowledge (STK) can, in principle, also be made explicit. The body is like a machine, and if a machine or a computer can do it, he claims, those tasks are reducible. But he creates another line for human exceptionalism defended by human ability to have tacit knowledge at a social level that cannot be reduced to algorithm (CTK).

Now we have self-driving cars navigating traffic to preprogrammed destinations. Were Dreyfus and Collins wrong? Not really. The problem with these defenses is not that they fail to demarcate important emergent skills. The problem is their characterization of irreducibility makes it seem we're quixotically fighting science's tide, constantly retreating as knowledge advances. This, I argue, is an illusion.

Michael Polanyi takes a different approach. He isolates a process of tacit knowing that replicates at various evolutionary levels—physiological, cognitive, linguistic, and social. We thus have irreducible tacit integration at sub-linguistic and sub-social levels. Emergent systems (like human minds) display higher-level qualities that are irreducible to physical laws alone. Polanyi links tacit knowing and emergence, so we might have irreducible tacit knowing even in primitive organisms. Each emergent level allows new possibilities of success or failure; earlier levels with proto-creativity and restricted choices lay ground for the sort of creativity and freedoms we enjoy. What makes us special is woven into life's fabric with different forms and degrees of expression.

---

[2] Hubert Dreyfus, *What Computers Can't Do* (New York: Harper and Row, 1972); Maurice Merleau-Ponty, *The Phenomenology of Perception* (London: Routledge, 1978); Gilbert Ryle, *The Concept of Mind* (Watford, GB: Mayflower Press, 1949).

[4] Harry Collins, *Tacit and Explicit Knowledge* (Chicago: University of Chicago Press, 2010), 121–2.

If Polanyi is right about reduction's limits in science, we need not retreat. We can see each proposed line—life, skills, language, reason, society, morality—as achievements providing more degrees of freedom with new risks and possibilities, as we locate innovation and irreducibility in tacit knowing and emergent being.

## 2. Can a Machine Perform Achievements of Tacit Knowing?

### 2.1 The Structure of Tacit Knowing

Polanyi saw tacit knowledge originating in the body, its general mechanism in tacit knowing's process. He says, "tacit knowledge is comprised of two kinds of awareness, subsidiary awareness and focal awareness."[5] Subsidiary conditions or “clues” come together into gestalt "joint comprehension." We move from tacit clues to focal meaning. This *from-to* structure is intentional and holistic.

*Ida B. Wells (2020): Helen Marshall & University of Chicago Library. Created by artist Helen Marshall of the People's Picture, commissioned by the Women's Suffrage Centennial Commission, and produced by Christina Korp, Purpose Entertainment.*

Consider the portrait of Ida B. Wells created from photos of women and pamphlets important to women's suffrage and civil rights. There is joint comprehension of many individual clues. To see Wells, we attend from individual photos, but when we examine the photos individually, their joint meaning dissolves.

[5] Michael Polanyi, "The Logic of Tacit Inference," *Philosophy* 41, no.155 (1996): 7.

We see individual people, which are the joint comprehension of colors and shapes. Attending further to colors and shapes, those people dissolve into pixels.

For Polanyi, perceptual and bodily skills—seeing in three dimensions, riding a bicycle, playing piano—have the same tacit knowing structure as higher cognitive processes like recognizing Wells's portrait or making scientific discoveries.[6] The *from-to* intentional structure provides a gestalt where particular meaning of clues individually are insufficient to their joint meaning. This creates an irreducibility Polanyi describes as the unspecifiability of tacit clues. Turning to specify clues we were *attending from*, we gain only indirect, incomplete knowledge of clues as they functioned to give focal knowledge. By adopting an alienated perspective we no longer "dwelling in" clues but perceiving them as focal objects themselves.

Irreducibility comes from *irreversibility* in tacit integration: looking back at clues to see them focally, you inevitably attend from different tacit clues. The uspecifiability creates *inexhaustibility* in attempts to make tacit knowledge explicit. Understanding the focal whole and examining clues, we inevitably miss some clues and always have room for further investigation and new meaningful connections.

Moving from clues to joint meaning we cross a “logical gap.” In one sense, this adds to unspecifiability of clues. In another sense, joint comprehensions provide the means to specify clues appropriately. If it is Ida B. Wells, those are *her* lips, *her* eyes. Once we have joint comprehension, it becomes tacit background through which we understand other things, including the clues themselves: the "to" becomes a "through," when particulars coalesce into concepts or when the joint comprehension of an organism’s parts allows us to understand a comprehensive entity, or a representation of one.

Joint comprehension sinks into the tacit background and one sees particulars in terms of meaning they provides. Levels of joint comprehensions build into hierarchies, one level supporting another as it influences the level below, and rules for each level can be discovered. In language, for instance, "The first level…is the production of a voice; the second, the utterance of words; the third, the joining of words into sentences" and "the voice you produce is shaped into words by a vocabulary; a given vocabulary is shaped into sentences in accordance with grammar" etc.[7]

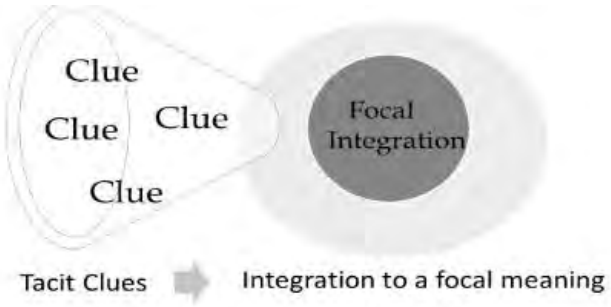
---

[6] Ibid., 4, 8.
[7] Ibid., 16.

**2.2 Modeling Body-Knowing and Learning with Neural Networks**

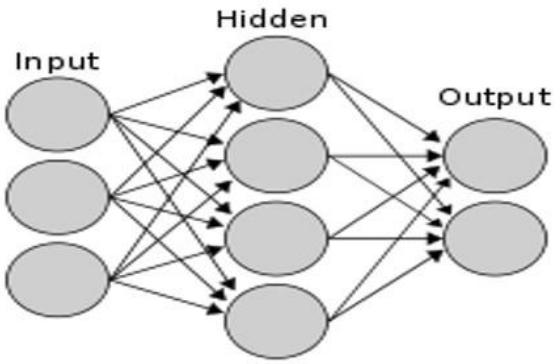
The AI machines most promising for simulating human language and behavior use connectionist or "neural net" architectures. They train on "big data" and "learn" appropriate responses. Connectionist systems excel at pattern recognition, and are used for speech recognition, giving us virtual assistants like Siri and Alexa. They sift through data for regularities to give us the next most likely word in a string of words, setting the basis for ChatGPT and Claude's response to questions. Though Polanyi underestimated early neural network theory,[8] he would now notice strong similarities between tacit knowing and connectionist architectures, and how these networks mimic irreducibility.

**2.2.1 Structural Similarities**: Tacit knowing's structure strongly parallels connectionist systems. Polanyi discusses how clues integrate to provide a focal joint meaning. Similarly, connectionist systems have "input units" connecting into individual nodes, integrating in complex processes producing resultant "output" nodes.



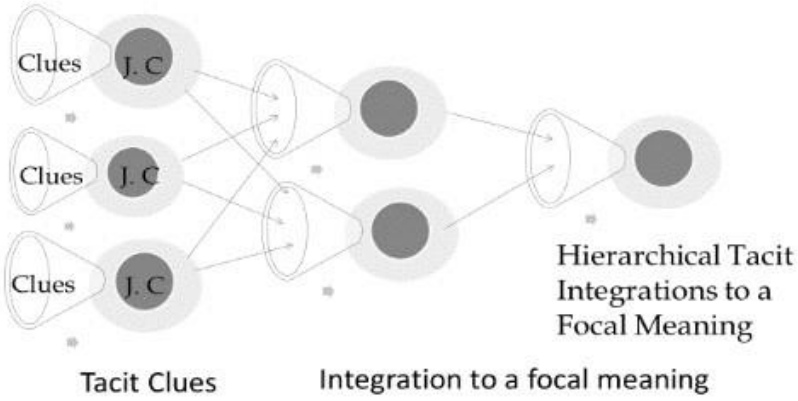

Parallel processing charting inputs to outputs model the difference between clues (inputs) in isolation and their joint meaning (output). Connectionists modeling language discuss "distributed representations," where inputs are sub-symbolic proto-representations and where intermediate vectors or outputs are symbolic representations proper.[9] Different experiences involving the sign "coffee" (inputs) don't individually amount to symbol COFFEE (output). Inputs in isolation are just sounds (like pixels), but outputs gains different, holistic significance (like Wells's portrait).

[8] Michael Polanyi, *Personal Knowledge: Towards a Post-Critical Philosophy* (Chicago: The University of Chicago Press, 1962), 340.
[9] Paul Smolensky, "Connectionism, Constituency and the Language of Thought," in *Connectionism: Debates on Psychological Explanation*, Vol 2, ed. Cynthia & Graham MacDonald (Oxford: Blackwell, 1995), 164–98.




*Structure of connectionist processing.*

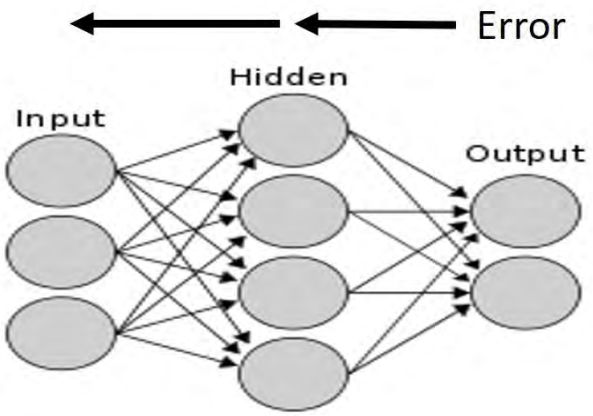
**2.2.2 Hierarchical and Training Similarities**: Connectionist networks, like martial artists, are programmed to perform tasks with coaching; one task builds upon others. Programmers can deliberately modify weights between connections, comparable to instructors providing correction. Networks can also be provided with repeated examples to adjust the weightings and the outcomes, just as practitioners of a physical skill train bodies through repetition, engraining new associations and activation patterns.




*Layers of tacit integration to joint meanings.*

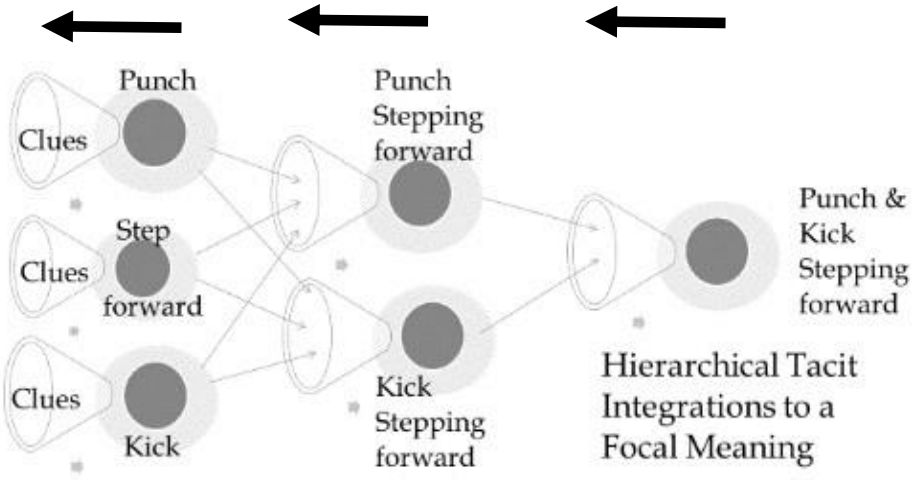
Networks can be modified via backpropagation algorithms that feeds information back into earlier input units, automatically making corrections toward given ends. Weights or connection strengths change, altering activation patterns toward desired goals without constant trainer guidance. This catches how the body itself works toward getting motion right, gaining experience and improving execution not just by repeating but by zeroing in on right moves over time "spontaneously." For instance, after months of attempting to properly follow instructions, the Brown Belt may finally "realize" how to put hip power into a punch correctly.

Polanyi discusses learning in terms of "intuition" and "technical invention."[10] When we don't know the end goal, imagination seeks possible solutions. "Intuition" is spontaneous and creative integration of subliminal clues to joint comprehension calling forth a possible solution. "Technical invention," in contrast is backward engineering, where we already know the end goal and discover how to organize clues or parts to function toward that outcome.




*Backpropagation: feedback loops in neural network learning.*

In somatic processes and body skills, the layered hierarchization of clues to joint foci can be modeled in connectionist networks, as one node (say the performance of a punch with a step) together with other nodes (stepping forward with a kick) feeds forward into another node (stepping forward with a kick and transitioning to a punch).




*Layers in learning a complex skill, with information feeding back.*

[10] Michael Polanyi & Harry Prosch, *Meaning* (Chicago: The University of Chicago Press, 1975), 96–8.

But much of this process is not directed by conscious adjustments. All along the way, we have the process of intuition and invention as the body self-corrects to find its bearings and perform the tasks with balance and power. The feeding back of information to adjust movements and correct errors towards reaching goals is similar to the process of backpropagation that adjust weightings to produce desired results in connectionist computing.

In reaching new skill levels, the body itself works between intuition and invention, moving intentionally toward solutions. The intentionality in the body's discovery, and freedom to diverge from goals, may be modeled in dynamical systems approaches to connectionist networks. Networks can develop their own goals and learn how to learn, independent from the guidance of a programmer or coach. The body's intentionality can be mimicked with "attractors" increasing probability of outcomes as systems develop. A dynamical system interacting with environment can achieve stabilization reinforcing particular patterns or move down gradients of potentiality to new patterns. The body develops its own attractors and repellors, and when working with trainers in wider systems, its intentional trajectory is affected, putting it on paths to new stable actions and reactions.

**2.3 Connectionist Irreducibility**

Connectionist networks may model somatic tacit knowing, but since they also mirror tacit knowing's irreducibility, they do not show human creativity and choice to be fully reducible. This can be seen best in the language-building efforts that were the drivers of AI progress. Jeffrey Elman presents a connectionist "words-as-cues dynamical model" inspired by the notion that "words do not have meaning but rather are cues to meaning,"[11] which seems very friendly to Polanyi's tacit integration. We see irreducibility in non-linear gestalt change from inputs to outputs. Distributed representations aren't yet symbols. We don't get COFFEE by mere aggregation of "coffee" iterations.

Just as there is an unspecifiability in tacit integration, there is an intractibility in effective neural networks, which use "hidden layers" to process data (See figure#). We see inexplicability in hidden layers' tangle between input and output units. These intermediate nodes preprocess information to help obtain solutions, but their functions aren't directly accessible to

[11] Jeffrey Elman, "Systematicity in the Lexicon: On Having Your Cake and Eating It Too," in *The Architecture of Cognition*, ed. John Symons and Paulo Calvo (Cambridge: MIT Press, 2014), 137–8.

programmers. There is intractability as clue or sign integration passes through hidden layers with outputs too complex to decompose.

So in connectionist learning we see another dimension of irreducibly tacit connected to irreversible holistic meaning shift and its intractability—a mismatch between system functioning and efforts to make that functioning explicit, coordinating with unspecifiability and inexhaustibility of clues when one is no longer "dwelling in" clues but attempting to "turn around" on them. Our descriptions work with top-down concepts and explicit rules that are incomplete, needing more *post hoc* refinements as exceptions crop up.[12] The artist tacitly learns the right way to move with help of explicit rules but isn't strictly following rules when artfully performing.[13]

Rules provide higher-level description and don't adequately describe actual lower-level processes involved in network performance (the tacit processes). This unspecifiability and inexhaustibility of the lower in terms of higher is an effect of what Paul Smolensky called semantic "dimension shift."[14] Concepts used to understand emergent activity don't quite fit the processing level, so emergent higher-level description is insufficient to account for the lower when we turn back to look at it.

## 3. Does a Machine's Ability to *Perform* Achievements of Tacit Knowing Mean that *that Knowledge* is Now Fully Explicable?

### 3.1 Collins on Making the Tacit Explicit

Collins rightly makes the most significant cut between tacit and explicit knowledge at social and linguistic levels. But for him, Collective Tacit Knowledge is the only sort not in principle (as yet) subject to fully explicit exposition. Relational Tacit Knowledge (RTK) occurs when experts relay knowledge tacitly to apprentices. Somatic Tacit Knowledge (STK) appears in skillful performance, but both are ultimately fully explicable for Collins, with only human limitations masking that explicability.

---

[12] See the section "Raven's Matrices: Rule-Following without Rules," in Charles Lowney, Simon Levy, William Meroney and Ross Gayler, "Connecting Twenty-First Century Connectionism and Wittgenstein," *Philosophia* 48, no.2 (2020): 643–71.
[13] See Charles W. Lowney, "From Science to Morality: A Polanyian Perspective on the Letter and the Spirit of the Law," *Tradition and Discovery* 36, no.1 (Fall 2009): 42–54.
[14] Paul Smolensky, "On the Proper Treatment of Connectionism," *Behavioural and Brain Sciences* 11 (1988): 11.

Collins contrasts tacit knowledge with explicit knowledge from elaborating or transforming processes into interpretable "strings"—material things engaging in causal relations.[15] For Collins, human bodily skills reduce to physical processes and are thus explicable as material, causal strings.

*Elaboration* is Collins' first sense of explicability. It basically just means we need more details about the process, i.e., we need a longer string. This sort of elaboration makes explicit RTK[16] The details we have, however, may need to be transformed into a different string that provides a better affordance for us to focally grasp; this string *transformation* is Collins' second sense of explication, e.g. if we translate a foreign text into a known language.[17] A third sense of explication comes *if machines can perform human tasks*, such as riding a bicycle, because it transforms the biological process into and explicitly interpretable form. Through this transformation we can analyze machines into explicit parts and state clearly how parts function together to produce desired results. Making a machine is like embedding the process in formal language, and the actual working of the machine acting as syntactical proof. It a robot can mimic human behavior, then Collins argues, there is an explicit in strings of material causes at work.[18]

AI machines like ChatGPT mimic some human cognitive abilities, and so these would also be reducible for Collins, because those skills, like somatic tasks are encoded as interpretable strings. But such neural networks *can't* do what humans can when they employ because CTK abilities can't (yet) be reduced to strings. In *Artifictional Intelligence*, Collins provides conditions that would have to be met for AI to achieve human tacit abilities, such as demonstrating "ubiquitous" common sense by passing a rigorous Turing Test.[19] This, for Collins, would show human tacit knowledge to be explicable and as fully reducible as he believes STK or RTK to be. But (as yet) Collins believes CTK demonstrates a special irreducible ability, while neural networks and human bodies are no more inexplicable than "cats and dogs—or trees and sieves for that matter."[20]

---

[15] Collins, *Tacit and Explicit Knowledge*, 9–10.
[16] Ibid., 81; See here for types of explication.
[17] Ibid., 81.
[18] See Charles W. Lowney, "Ineffable, Tacit, Explicable, Explicit: Qualifying Tacit Knowledge in the Age of 'Intelligent' Machines," *Tradition and Discovery* 38, no.1 (2011–2012): 24.
[19] Collins, *Artifictional Intelligence* (Cambridge: Polity Press, 2018), See chapter 10.
[20] Collins, *Tacit and Explicit Knowledge*, 77.

There is indeed CTK involved in the "programming" of the martial artist or the artisan baker; there are social rules and context sensitivity at work guiding when it is appropriate to do what. As Collins says of the martial artist, "one way to win would be to smash the opponent's head with a baseball bat, but that would not be a *right* move, the criteria are collective."[21] My claim is not that productive skills and arts are not nested in higher-order social contexts which act as constraints on lower-order meanings. I agree that a higher social context can be the tacit background that influences individual acts in a manner that is irreducible to lower-level material causes. My claim is that there are similar dynamic "dual control" systems[22] at lower levels as well, and that those lower levels also display forms of the creativity and tacit knowing we see in higher-level discoveries.

### 3.2 Machines, Animals, and Martial Artists: Reducible Complexity?

Collins provides two main reasons why we should see connectionist computing networks and somatic skills as reducible to determinate causal chains: first, training AI machines, animals, or martial artists works like operant conditioning, which he believes betrays a fully deterministic origin.[23] Second, connectionist programs can run on digital computers; there's physical transformation from non-linear connectionist processes into a linear digital one.[24]

**3.2.1 Operant Conditioning and Dynamical Systems:** While many identify operant conditioning with causal determinism, it's not quite that. Connectionist systems display flexibility and something like intentionality that often diverges from the programmer's will. When combined with dynamical systems understanding, we model ability to learn. Dynamical systems develop attractors providing propensity toward certain outcomes in ways allowing degrees of freedom and creativity.

We train animals, bodies, and connectionist systems with reinforcement like operant conditioning, but this increases response probability; it doesn't act like bottom-up physical cause. Operant conditioning implements upper-level constraint, acting more like telic or formal cause than efficient or material cause. It operates the way dynamical systems create attractors when encountering other systems. Or the way a system, for Polanyi, can generate a "telic field" that

[21] Harry Collins, "Analysing Tacit Knowledge: Response to Henry and Lowney," *Tradition and Discovery* 38, no.1 (2011–2012): 41
[22] Polanyi, "The Logic of Tacit Inference," 15–16.
[23] Collins, *Tacit and Explicit Knowledge*, 75.
[24] Ibid., 75.

moves something forward toward a particular configuration, e.g., the way the DNA of a developing embryo operates together with the epigenetic landscape.[25]

Collapsing the conditioning process into a determinant string ignores the creative contribution of bodies and their role in discovery. It ignores variety of possibilities bodies have chosen from and numerous variables needing balance to initiate proper responses. Bodies try many different movements learning to perform techniques correctly, offering alternatives in the process. We have top-down control in martial artists deciding against using bats, but also bottom-up suggestion of new moves and top-down guidance of bodies themselves in moving muscles to avoid being hit by an unexpected strike. Flattening out the process of learning *post hoc* can make it look like a determinant string, but if you use operant conditioning to train a dog to sit and then a rabbit runs by, you will see that the dog still has a difficult decision to make.

Collins seems to recognize the illegitimacy of this flattening when it comes to CTK, but not for STK. For example, he says a computer cannot tell when a word is misspelled on purpose, because this requires irreducible CTK.[26] He acknowledges that we might repair this glitch by adding some lines of code, which would then allow words to be misspeled under certain conditions,[27] but this is the equivalent of a post hoc flattening. The tracing a string to the desired result doesn't belie the fact that CTK was involved. I see a similar context sensitivity and flexibility for the pre-CTK body.

Operant conditioning also has its explanatory limitations. It takes some of the implicit tacit clues that we can explicitly identify from a complex integrative system and works just with those and peripheral associations. This "behaviorist analysis is intelligible only because it imitates, however crudely, the tacit integration which it pretends to replace."[28] It doesn't identify a strictly bottom-up causal mechanism—let alone the mechanism as indwelt by a living being.

**3.2.2 Running on Digital*: Is Complexity Reduced?***: We saw how neural networks, with their complex non-linear processes, model the irreducibility of tacit knowing, but what of the ability for a connectionist system to run on a digital computer? Does that not prove that all this complexity is ultimately, in principle, tractable and reducible?

---

[25] Polanyi, *Knowing and Being*, 219, 232. See also Gregory Bateson on "deutero-learning" [e.g., in *Mind and Nature: A Necessary Unity*, New York: Bantam Books (1979)], which shows even Pavlovian learning to be more context driven and stochastic than are effects determined simply by bottom-up causes.
[26] Collins, *Artifictional Intelligence*, 4.
[27] Ibid., 11.
[28] Polanyi, "The Logic of Tacit Inference," 14.

Well, it could be that the more intentional and creative features of body-learning have not yet been properly modeled by even connectionist systems—and that is likely. A machine that mimics a human ability is a working model that is analogous but not identical to the target. For example, one can model the movement of the earth around the sun by swinging a rock around with a string. Some information will be made explicit, but all the relevant features will not be modeled. Collins claims that AI machines cannot yet pass the Turing Test, but that might be a matter of time and technological innovation. Turing Tests, etc., are ways to judge whether enough of the relevant features are modeled for an AI to be considered to have our sort of thinking ability—but even here we still go only by the clues that we can make explicit, and base the adequacy of the model being tested on those.[29]

The ability to run purportedly irreducible connectionist programs on digital computers makes it appear that connectionist networks, although complex, could be reducible to linear digital computer processes, and thus be fully reducible to material strings. But this could be another "dual control" system, which still harbors irreducibility, as upper levels influence boundary conditions left open at the lower levels. A connectionist system (upper-level control) constrains digital computer features (lower-level control) to display connectionist functions. But even if there is a reduction from neural networks to digital computing, and there is just one *one-level* machine, this, for Polanyi, could still indicate we're dealing with something emergent not fully determined by enabling conditions, precisely because it's a machine and machines themselves are dual control systems, whose lower-level laws underdetermine the higher-level engineering principles.

### 3.3 The Irreducibility of Tacit Knowing and Emergent Dual Control Systems

We needn't answer whether machines can adequately model STK to answer whether body-knowing has irreducibly tacit features. Even if machines successfully model baking artisan bread, riding bicycles or competing in mixed martial arts matches, we can question whether this reduces tacit knowledge involved. Elaborating, transforming, or untangling material strings to make them explicitly interpretable isn't the main issue, since even simple machines, like sieves,

[29] Polanyi rejects the relevance of the Turing Test, while acknowledging a machine might deceive us. See Polanyi, *Personal Knowledge*, 263.

are irreducible according to Polanyi; there's semantic dimension shift indicating the presence of an emergent entity.

In emergent systems, there's integration of clues to focal meaning and parts to whole within facilitating environment. Here we see how tacit knowing structure acts in tandem with emergent being structure, both blending at rudimentary description levels when discussing living systems and machines. This comes across vividly seeing computers—tangible entities—mimicking tacit knowing processes. The meaning change at higher integration levels is displayed in need to use higher-order principles to account for system operation. Hence, machines represent different ontological entities than their parts and have different meanings.

To identify machines requires higher-level description in terms of function; their very existence cannot be recognized by lower-level principles delineated by physics and chemistry. Complete physical description of steam engines couldn't tell you whether they're working or broken; success or failure as machines is not recognized by physical description. The notion of something working or being broken is higher-level determination requiring understanding of engineering principles. Lower-level descriptions solely in terms of physics or chemistry are insufficient.

Any machine or living system exemplifies dual control systems. No physical laws are violated, but possible physical element configurations are constrained by stable higher-order emergent wholes and their principles and laws. Emergent entities control boundary conditions left open by lower levels, as engines constrain steam to move pistons. There can be success (machine works; animal lives) or failure (machine broken; animal dead), but neither can be comprehended solely in lower-level causal or material string terms.

Irreducibility from tacit knowing, linked with emergent being, displays the problem of understanding any lower-level subsidiary process in terms of higher-level concepts, and any higher-level process in terms of the lower. There are many dependency relations forming dual control systems. Polanyi characterizes them generally as a hierarchy moving from lower to higher levels: physical, chemical, biological, psychological, social, then personal. In the emergentist schema, the emergent level is ordinate and in between the subordinate (parts) and the superordinate (environment it is a part of). When we look at activities at our own level, we have subordinate and superordinate systems acting as tacit clues. For example, in the portrait of Ida B. Wells (figures 6.1 and 6.2), both the colored pixels (subordinate) and our knowledge of human

faces and civil rights heroines (superordinate) are in play. We understand and make particulars explicit in terms of their higher context (it is a picture of Wells), which is not strictly causal, or in terms of lower-level clues (colored pixels making lines and shapes), which won't be enough on their own to tell us who or what we are talking about.

The irreducibility Collins sees in CTK exists at many levels, but it is harder for us to recognize. We see variability and multiple possibilities in social performances, so creativity, intentionality and freedom come to forefront. STK irreducibility is harder to see because it's easier to think we've made something fully explicit when several layers below us on th hierarchy of being —closer to physical causality—and its various potentialities have been flattened out.

**3.4 Limits of Scientific Explanation**

Scientific explanation is Collin's fourth type of explication, and can take several forms: showing how particulars fit into higher-order theory, or how material parts function together according to lower-level laws. But just as there's inevitable mismatch making tacit knowledge explicit from top down, there's also mismatch from bottom up: Looking down from above with emergent higher-level conceptual descriptions, we only approximate tacit causal processes and indirectly circumscribe them. Using lower-level descriptions, we miss higher-level meaning and inadequately describe emergent properties of real activity, the proper way to understand it in higher-order context.

For reasons both of the irreversibility, unspecifiability and inexhaustibility encountered in a top-down approach, and of the underdetermination and semantic dimensional shift of a bottom-up approach, one might say that knowing how cannot in principle be completely captured by knowing that, and an emergent being or even a complex neural network can't be reduced.

Emergent systems, such as animal bodies and machines—though they can die or be broken—are both irreducible, and the epistemic processes of perception, cognition, and skillful activities follow that same structure. So, the possibility that a machine—some sort of AI computer—could mimic human activities and human cognitive capacities would mean different things for Polanyi than for Collins. For Collins, it would mean that we have found transformations of those strings that would make what was always explicable explicit and reducible. For Polanyi, it could mean that the machine has risen to a yet higher emergent status. But can a machine really "know" anything?

## 4. Dwelling in the System: Evolution, Animals and Intelligent Machines

Collins puts animal and human bodily skills on par with deterministic mechanical processes. "[C]ats and dogs and sieves and trees cannot be said to 'know' any explicit knowledge, they shouldn't be said to know any tacit knowledge either... they just transform strings"; likewise for body skills we develop and neural network skills.[30]

There's some truth here; we usually reserve "knowing" for humans who can use concepts, but the line isn't sharp. Martin Davies softens the line when he recognizes that tacit knowing need not involve conceptualization; and so, he uses "cognizing" for *tacit* "knowing."[31] Polanyi, too, softens the line. He emphasizes importance of *knowers* who *attend from* clues and *attend to* focal integration and who *dwell in* clues to recognize emergent meanings.

Davies introduces "cognize" also to avoid the pseudo-problem of how something tacit can be knowledge. Someone attends from clues to joint comprehension. This indicates some consciousness must emerge before tacit knowing reaches higher potential, but clearly tacit knowing process, occurring in bodily skills, predates human knowers.

Polanyi says, "by acquiring a skill, whether muscular or intellectual, we achieve an understanding which we cannot put into words and which is continuous with the inarticulate faculties of animals."[32] This continuity features unformalizable and unspecifiable aspects of tacit knowing that confound strict causal determinism and support emergence of entities with increasing degrees of freedom.

### 4.1 Proto-Structures

For Collins, if we make machines to perform tasks usually requiring apprenticeship or tacit skill, that performance doesn't display tacit knowing in strictest sense: it's "mimeomorphic" and mechanical rather than "polimorphic"[33] and doesn't truly display what is irreducibly tacit. For Polanyi, tacit structure pre-exists linguistic and social knowledge. It is an intentional structure that forms the basis of simple organic activities such as those involved in perception. So one can say of even rudimentary bodily processes that some pre-conscious form of tacit knowing

[30] Collins, *Tacit and Explicit Knowledge*, 78.
[31] He follows and explains Noam Chomsky's similar use of "cognize." In Martin Davies, "Connectionism, Modularity, and Tacit knowing," *British Journal of Philosophy of Science* 40, no.4 (1989): 551.
[32] Polanyi, *Personal Knowledge*, 90; Collins, *Tacit and Explicit Knowledge*, 76.
[33] Collins, *Tacit and Explicit Knowledge*, 55.

or cognizing is taking place, and there are hierarchies of these structures that build on one another in the course of evolution. Animal hunting is not an example of tacit knowing, for Collins,[34] but the skills animals use in hunting are the basis of the skills that we use as we develop crafts and arts. Bodily tacit processes, perception, and the skills of animals, have proto-meanings (one might say) that linguistic and social meanings are built on. As Charles Taylor and Hubert Dreyfus claim, following Merleau-Ponty, motor intentionality is the structural basis for representational intentionality.[35]

An advantage of Polanyi's approach is that it shows a knowing structure that manifests differently at different levels. There is a similarity in structure, though differences in types and orders of meaning. There is indeed the disjunction Collins sees between social and somatic, between humans with language and animals without our language, but these operate like layered subsidiary-focal achievements and the layered sets of dual control systems supporting them. Rather than one big dividing line between explicit and irreducibly tacit located at CTK, we see emergence of tacit structures through evolution, reformed and repurposed to meet new challenges. For example, the way feathers used for controlling temperature, were used to facilitate flight.[36]

The converse of this multiple re-purposing is multiple realizability, where the same property or type of thing can be realized in multiple ways, for instance, the ability to fly can manifest one way with birds, and in other ways with bees or squirrels.

Polanyi sees emergent intentionality and degrees of freedom even in most basic life forms and simplest machines. The tacit process, in rudimentary form, is already at work. Seeing the link between tacit and emergent may have implications for understanding evolutionary process. How evolution retools tacit intentional structures for different goals as species develop (like a proto-*intuition*), how it uses different means to achieve similar effects (like a proto-*invention*) and how a dynamical attractor works like a telic field (like proto-*intention*) might all be part of a properly understood evolutionary picture.

---

[34] Collins, *Tacit and Explicit Knowledge*, 78.
[35] Charles Taylor and Hubert Dreyfus, *Retrieving Realism* (Cambridge: Harvard University Press, 2015), 50.
[36] As William Bechtel notes, "evolution often works most effectively by taking components that were previously employed for one purpose and using them for other purposes. This kind of evolution occurs at the expense of decomposability, since it depends on building up additional connections within the system to build a more integrated system." William Bechtel, "Perspectives on Mental Models," *Behaviorism* 16, no.2 (Fall, 1988): 137–48.

### 4.2 Telic Fields

The notion of an intentional component to tacit knowing goes together with Polanyi's notion that there is some sort of force drawing evolution forward toward more complex and free unities. For Polanyi, living systems respond to "telic fields"[37] helping them move forward in ontogenetic and evolutionary history toward goals. These attractors need not be intelligent beings' work, nor forces at history's end, but can be emergent environments coaxing entities along paths toward richer meanings or ways of being.

The idea of tacit knowing and emergence may go down as far as the simplest forms of life, which show a unity and a meaningful response to their environment that is geared toward self-sustenance. Polanyi would side with Thomas Nagel in saying that, left only to chance variation and ability to survive, "the materialist Neo-Darwinian conception of nature is almost certainly false."[38] But if there is a telic field that unlocks or establishes various potentialities that an entity possesses in particular environments, that itself can be emergent. The material conditions or environment can coordinate in a way that sets up attractors and increase the probability of the expression of certain traits. So, we have something like Polanyi's "maturation,"[39] and the notion of a telic, rather than a pre-existent or final teleological, principle at work.[40] For Polanyi, life, and then consciousness, are emergent. But they are also drawn forward into existence by telic forces or principles that were folded in from the beginning. In development, "this field of forces would also be the gradient of a potentiality: a gradient arising from the proximity of a possible achievement."[41]

### 4.3 Intrinsic Value of Animals

The existence of tacit intentionality in non-human animals may affect understanding of their intrinsic value. There might be important distinction between living things and machines. Although machines are emergent things, we put meaning into machines—they exist as real, but we recognize them only in relation to our purposes. They are “contrivances” that model

---

[37] Polanyi, *Personal Knowledge*, 403.
[38] Thomas Nagel. *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is Almost Certainly False* (Oxford: Oxford University Press, 2012). Compare also Bateson on the telic as a feature of self-corrective systems including larger ecological systems. Gregory Bateson, *Mind and Nature: A necessary Unit* (Toronto: Bantam Books, 1980), 117–18.
[39] Polanyi, *Personal Knowledge*, 395.
[40] See Richard Gelwick, "Michael Polanyi's Daring Epistemology and the Hunger for Teleology," *Zygon* 40, no.1 (2005): 63–76.
[41] Polanyi, *Personal Knowledge*, 398.

intentionality but don't experience it by "dwelling in" clues. Living organisms, in contrast, are "comprehensive entities" that have a form of intentionality that they indwell; there are differences that make differences *to them*.[42] Without us, they have their own meaning and purpose. Having interests, recognizing things as meaningful, indicates valuing and having value.

Showing the intentional and creative structure of body-knowing thus reinforces Holmes Rolston III's conception of emergent value in natural entities and emergent structures; we can see levels of intentional meaning and value in at least sentient beings and higher zoology, and even in lower zoology and botany.[43] So, there are moral implications here when we tie together indwelt meaning with some sort of intrinsic value, which would need to be spelled out in an emergentist framework.

Dreyfus' exposition of body-knowing and Collins' exposition of CTK both show important emergent jumps in knowing, being and value. We can agree that most animals do not share collective tacit knowledge, but we do not have to defend the line between animals and humans to preserve human irreducibility against material reductionism. Our sort of freedom and context sensitivity may only emerge with self-consciousness within society, but degrees of freedom or proto-freedom can be seen to various degrees in various living organisms/bodies. Polanyi, like Dreyfus, can afford to be more generous to our cousins and ancestors in other species, but could Polanyi be as generous with AIs?

**4.4 The Future of Consciousness**

Connectionist networks aren't ordinary machines. They model inexplicability with "hidden layers" and exhibit semantic "dimension shift." They can function like complex dynamical systems developing "attractors" as they interact with environment and reinforce their own "intentional" trends. If connectionist networks can model STK, they might still be intractable and inexplicable. They might model well body-knowing with emergent degrees of freedom from deterministic physical causality, but are they free to the degree animals might be? And might they one day rise to consciousness, even self-consciousness?

I agree computers cannot yet do what people do using collective tacit knowledge. However, in emergentist pictures—given multiple realizability—there's acknowledgment that

[42] On Polanyi's notion of contrivances and comprehensive entities see Phil Mullins, "Michael Polanyi and Contemporary AI" [Publishing info].
[43] Holmes Rolston III, *Conserving Natural Value* (New York: Columbia University Press, 1994).

consciousness might form from different material subsidiaries, leaving open machine consciousness possibility. But experience matters, as do the affordances of mediums (biological vs. silicon), since some experiences are specific to particular mediums in particular environments. To get conscious or self-conscious AI, we'd have to provide right bodily functions and clues, and they'd have to develop right social interactions and become autonomous to some extent. They'd also need emotional intelligence if we're looking for human-like entities. As it is, AIs only manipulate representations; meaning in and meaning out is still overwhelmingly provided by human interpretation.

AIs mimic tacit knowing, but don't seem close to being systematic structure that can exercise tacit knowing by *dwelling in* their subsidiaries.[44] Animals can properly be said to have their own sense of intentions. But no machine can yet dwell in clues to experience even proto-intentionality. Currently, AI is too superficial a model of tacit processes to catch emergent nuances we look for. It would be grander accomplishment to create machines with right affordances and experiences to have their own sense of intention. While there's likely something it's like to be a bat—as Nagel would say[45]—it's unlikely there's something it's like to be Siri, or Claude, or an even much more complex and well-trained machines, like Ava in the movie *Ex Machina* (2014). Currently, advanced processing machines aren't emergent comprehensive entities that dwell in their parts and clues.

**Conclusion**

Rather than reduce humans to machines, we see an irreducible tacit knowing that emerges at multiple levels, from the simplest organisms to the most complex social structures, and, potentially, to future Artifactual Intelligences that will genuinely dwell in their own meanings.

With the emergence of animals with increasing types and degrees of knowing and freedom, we develop the conditions required to appreciate value and make moral decisions. If we develop sentient, conscious, human-like machines, we won't have to worry about similarities with animals, martial artists or computers reducing all that is meaningful and valuable about humans—or this new AI—to rocks and strings. We will instead extend the family of beings that share these irreducible qualities.

---

[44] Polanyi, "The Logic of Tacit Inference," 14.
[45] Thomas Nagel, "What is it Like to be a Bat?" *Philosophical Review* 83, no.4 (1974): 435–50.