CAUSES, GOALS, and REASONS: Clarifying the meanings of teleology

by Dan Bruiger

Hornby Island, British Columbia
dbruiger [at] telus.net

ABSTRACT:

Teleology is a nebulous and historically troublesome concept that still requires clarification. Some ambiguities in the target question are first noted, and then reformulated in relation to the key notion of *agency*. Genuine teleology is proposed as uniquely a property of agents, in contrast to varieties of apparent teleology, which involve projections by agents. The question of the mathematical description of goal-oriented behavior is explored.

1. Redefining the challenge

First, let me comment on what I believe to be tacit assumptions or subtleties in how the target theme for this essay is presented.[1] Clarifying these issues will help to cut to the chase, which I believe is the question of a mathematical or other formal description of goal-oriented behavior. However, as I interpret it, the question presented is: *what is the relationship between mathematically describable causal laws and apparent or real teleology in nature*? Or, for short: what is the relation between causes and goals? Here are a few points to consider:

a. The phrase "*Wandering* Towards a Goal" already suggests *stochastic* processes at the root of any teleology. This is consonant with the modern scientific paradigm, which would ground all order in the random. The very notion of the random, however, is ultimately only intelligible in contrast to its opposite, intention. It is otherwise metaphysically empty. What is random, chance, or accident but what is not intended? Similarly the notions of determinism and indeterminism have no sense but with regard to an agent's ability to *determine* (either to *ascertain* or to *fix* by some action). Determinacy and indeterminacy are states of knowledge, not of nature.

How can mindless mathematical laws give rise to aims and intentions? The blunt answer is that mathematical laws cannot give rise to anything but other mathematics! (Or to human actions based on their predictions.) Natural laws in general have no "governing power." A "law" is an afterthought, a maximally compressed expression of an empirically discovered pattern [Chaitin 2009]. The empirical *correlation* mathematically expressed is one thing, the *causality* involved in "giving rise" to the correlation is quite another [Bunge 1979, p77].

The question quoted above presumes a physicalist worldview in which all phenomena can be reduced to the terms of modern physics. (Or, perhaps, a worldview in

---

[1] As per the "evaluation criteria." http://fqxi.org/community/essay/rules

which the terms of modern physics are reduced to mathematics [Tegmark 2008].) The question as posed also presumes that mathematical laws *are* mindless, whereas they may be viewed (though not by Platonists) as products of the minds of mathematicians. In that light, the question could with equal rights be put the other way around: “How do aims and intention give rise to mathematical laws?” That formulation would presume a very different worldview, in which human being, perception, and action are primary.

Even granted the physicalist perspective, the tricky part of the question lies in the expression “give rise to.” The question might then read: how does the behavior of “matter” (for example, in the brain), which may be mathematically describable, give rise to “mind”? In principle, such a question might be answerable if *aims* and *intentions* could be described in strictly behavioral terms: for example, behavior of neurons accounting for molar behavior of the organism. For, then we remain in the domain of physical description. Alternatively, we remain in the domain of intentional or mental description if mathematical laws can be described in strictly mental terms. For example (perhaps not so outlandish): an intentional account of the strategies of neurons (or molecules) leading to the organism’s goals. Mixing or crossing categories (‘mental’, ‘physical’), however, gives rise to logical inconsistencies and makes the problem seem insurmountable.

b. *These laws [of dynamics] provide predictions by carrying conditions at one moment of time inexorably into the future.* This is a statement of classic determinism, which I contend is a purely mathematical concept (i.e., a property of equations) and not viably a physical one (a property of nature) [Bruiger 2016]. Determinism applies only to certain systems (such as planetary orbits) that correspond (with sufficient approximation) to mathematical models.

c. *The motion of the most basic particle can be described by the action of forces moment by moment or as the attempt to extremize an action integral…* At least in systems presumed to be deterministic, descriptions by means of action principles are no more than equivalent alternatives to dynamical descriptions. This was noted by many commentators since Maupertuis, who were anxious to preserve the independence of science from religion by denying any metaphysical significance to action principles [Stöltzner 2003]. Action principles do *not* imply teleology [Bunge 1979, p83]. A teleological interpretation is always trivially possible in any deterministic scheme [D’Abro 1939], but this implies no attempt by nature toward economy. No relation should be assumed between the goals of living organisms and the tendencies of non-living systems toward certain mathematically defined states, such as extrema, equilibria, or attractors.

d. *How do goal-oriented systems arise… in a world that we can describe in terms of goal-free mathematical evolution?* A simple answer to this question is that they don’t! That is, goal-oriented systems arise in a world far more complex than that described by “goal-free mathematical evolution,” which treats always of simplified idealizations, such as isolated or reversible systems. Mathematical models are inherently idealized and simplistic. They are defined by equations, which only approximately describe real systems. Apart from using the same word, the “evolution” of a mathematical function (usually with time as independent variable) is a categorically different thing than natural evolution by selection. While an equation *might* be found to describe natural evolution over time, one should not confuse such a description with the process itself of selection. The further question—*How do physical systems that pursue the goal of reproduction arise from an a-biological world?*—is at least mildly rhetorical, since the question has been well studied ever since Darwin.

e. *How are goals (versus accomplishments) linked to 'arrows of time'?* First, 'accomplishment' must be disambiguated. It is appropriately the result of a *goal* but could be mistaken for *effect* (which is rather the result of efficient cause). Deterministic systems alone are reversible (no arrow of time), such that cause and effect are interchangeable. But, deterministic systems are mathematical fictions. Goals and accomplishments are not reversible. The deeper question is therefore what role goals or (final causes) might play in the real world in which efficient causal processes are irreversible. In particular, might there be some kind of teleology at work in the inanimate world of physics and cosmology, despite its traditional proscription, if the cosmos turns out to be self-organizing on large scales? [Bruiger 2014]

f. *What separates systems that are intelligent from those that are not?* This question has been addressed by cybernetics in general and within theoretical biology by Maturana and Varela [1980] in particular, in terms of the concept of *autopoiesis*, which expressly avoids reference to human observers. The short answer to the question is that intelligent systems are autopoietic (that is, self-producing). On the one hand, that understanding severely constrains what can be considered artificial intelligence. On the other, if it should turn out that the cosmos in some sense is autopoietic, then the very concept of *un*intelligent matter, as developed since the scientific revolution, would have to be reconsidered.

2. Causes, goals, and agents

What is the relationship between cause and goal? Aristotle had proposed four kinds of answer to the question of why something occurs: material, formal, efficient, and final. The modern notion of causality corresponds most closely to the third, "efficient" cause. In contrast, the modern notion of goal seeking corresponds most closely to the fourth, "final" cause.

In a broad sense, an efficient cause is a necessary or sufficient condition. (These are *logical* relations, however; determinism corresponds to the ideal case of being both necessary and sufficient.) In a narrower and more informal sense used in classical physics, a cause is an event outside a system that precipitates a chain reaction of further events within the system. There is usually an assumption of physical continuity through this chain (sometimes challenged, as in "action at a distance"). Experiments generally try to isolate a single cause, among many contributing factors; and such a cause generally operates in one direction. This interpretation of causality suits a mechanistic view of the world. In reality, however, causes are multiple and may be circular or mutual, especially in the living world.

A vast literature has accumulated on teleology, testifying to the confusing difficulties of the subject. The term *teleonomy* [Pittendrigh 1958] was coined in part to divorce the apparent goal-seeking behavior of organisms from more suspect cosmic or religions notions of teleology. Mayr [1974] defines a teleonomic process or behavior to owe its goal-directedness to a "program." He thereby hopes to assimilate teleonomy to causality, since he views a program as unfolding through material processes. In contrast, he calls *teleomatic* inanimate processes that seem to converge toward some finality but actually are merely "automatic" consequences of natural laws. I believe a further distinction should be made, between organisms or systems as considered from an observer's point of view, and from the (self-defining) point of view of the creature itself. *Teleology* should then properly be reserved for the goal-seeking of an *agent*, which is a system that *acts* on its own initiative and energy, and in terms defined by itself—that is, following goals that are unequivocally its own. This is in contrast, first of all, to inanimate systems that simply *react* through the transfer of efficient cause and energy. But it stands also in contrast to systems whose goals are attributed by observers who speculate about "function." The obvious paradigm of an agent is the individual organism. But observer speculation has also been applied to species, kin groups, evolution at large, and the entire cosmos. In earlier times, God was considered the agent whose goals resulted in the apparent finality of various aspects of his creation. My use of the term, however, is restricted to natural agents and does not include any association with theology.

What is agency? Historically, *agent* complements the archaic notion of *patient*. This distinction is no longer made in a natural science that *defines* matter to be passive and physical processes to involve only efficient cause, never agents and their goals. An agent is an *initiating* (if not an utterly *first*) cause, in contrast to efficient causes, (which only nominally serve as first causes) passively transmitted through physical systems. Only an agent pursues goals.

Every conscious subject is an agent, though not every agent is a conscious subject. There can be no consciousness without internal representation, but there can be representation without consciousness. Scientists are *embodied* agents, who—like many other organisms—make motivated internal connections that serve to represent events in an external world. This introduces a dualism between sign and signified. To avoid confronting this dualism, the subjective agency behind the representation is normally not included in scientific description itself. While logical connections may be held to have their own intrinsic existence, causal connections must be observed or inferred, and someone must do this observing or inferring. Yet, like logical connections, the observed or inferred causal connections are simply considered (by a consensus of agents!) to objectively exist in the external world. Hence, the *internal* connections representing them are not normally part of the physical picture. But it is just these connections one must consider as the actions of an agent. Such an agent does not have to be conscious, nor does the internal connection have to be consciously made. (On the contrary, agency might well serve as the basis to *explain* consciousness!)

An action in this sense—such as making internal connections—is motivated by values and goals, and is justified with reasons. This stands in contrast to events that are effects of efficient causes, and which are not initiated by agents but merely observed by others. Agency in this sense is a fundamental concept, a basic category that includes more than conscious human beings and their actions. Physical science historically excluded agency, along with final causes, in the mechanist program limited to efficient causation, passively inert matter, and space-time description. By denying its own agency, physics has positioned itself poorly to study agency in the world.

3. Teleological, teleonomic, or teleomatic?

Let us reserve 'teleology' for an agent's *own* purposes or goals. 'Teleonomy' is then the *ascription of agency* to a system.[2] When reading goal seeking into the behavior of a living creature, for example, one is oneself an agent imputing goals to another agent. But this projection occurs also when one ascribes goals to merely inanimate systems. Following Mayr, let us call these *apparently* teleonomic processes 'teleomatic'.

We have our own goals in ascribing goals; and the idea we have of a creature's goals does not necessarily correspond to the goals it is actually pursuing. If additionally the creature has an idea concerning its own goals (or ours), that idea is not necessarily the idea human observers have in mind. On the other hand, if we read goal-seeking into the observed behavior of a *non-living* system—unless it is demonstrably an autopoietic system—we are not dealing with an agent at all, but simply indulging a metaphor or figurative way of speaking.

Even in dealing with real agency, one must distinguish the domain of the observer from the domain of the observed. This already introduces the subjectivity normally and historically excluded from a science that fails to deal with agents. In addition, one must distinguish between real and apparent teleology. Real teleology is exclusively a property of agents. Its domain of discourse consists precisely in the considerations of agents: reasons, goals, purposes, intentions, etc. (the very sort of "mental" thing eschewed by physicalist description). On the one hand, observers may understandably be prone to impute their own reasons and goals to other agents. On the other, they may ascribe apparent teleology to a class of real phenomena (the teleomatic) that merit their own explanations, most likely in conventional causal terms. One should by no means assume that such phenomena involve agency. The relevant domain of discourse instead consists in efficient causes, space-time descriptions, state spaces, etc., as well as derivative concepts such as mathematical attractors and "as if" language that speaks *figuratively* of goals and purposes.

In cases of the teleomatic, in which a physical system seems to converge toward an end-state, that end-state is only known after the observed fact. Mathematical treatment (integration) may give the illusion that the final state is knowable in advance, or exists timelessly alongside the initial state. But this effect owes to the properties of equations, not of natural systems. Thus, extrema, attractor points, biological fitness, evolution, and even the 2^{nd} Law of Thermodynamics, may appear to involve end points but do not involve goals in the sense of something causing the end-state. *Only* agents have goals in that sense and involve that sort of finality. The only agents we know of are individual biological organisms. And the only agents to whose goals we have direct access (without observer speculation) are ourselves.

---

[2] This distinction corresponds roughly to that made by Reese [1994] between *purposive* and *purposeful* behavior. The former means *serving a purpose* (which is not necessarily the system's own purpose, and may be but an appearance only). The latter means *having a purpose*, which is necessarily the system's own.

4. Programs

Mayr [1974] views a *program* as materially existing before the onset of the teleonomic process. He defines a program as "*coded or prearranged information that controls a process (or behavior) leading it toward a given end.*" This is less a description of a kind of behavior or process than a postulate about its origin. Can we prove or discover the existence of a program responsible for a given behavior and can we know its details? Is it in fact material? A computer program, after all, is not strictly material but consists abstractly of a *text*, which then becomes a *state* of the physical computer. The programmer, not the machine, specifies the goals of the program.

The notion of program, even if it corresponds to an objective material entity (such as DNA), does not circumvent the question of whose purpose or intentionality it manifests. Claims about the function of a program are no different than claims about the function of various organs or behaviors, such as kidneys (to filter the blood) or migration (to avoid starvation in winter). Cashing them out boils down to claims by observer agents such as: 'If I had designed this system, this would be the purpose I intended for that function', or 'if I were this creature, this would be my intention'. That is, speculation about the "purpose" or "function" of something amounts to substituting one's own intentionality for the operation of the system or creature.

The idea of teleonomy as program is attractive because, first of all, a program can be expressed formally or mathematically. Secondly, as a *material* thing or process, it can potentially be explained in terms of causal processes. That is, the coded information of the program can guide the causal unfolding (for example) of chemical and cellular processes of development of an organism. And the formation of the program itself can potentially be explained in causal terms, for example though natural selection of mutations. DNA, which consists of identifiable molecules, illustrates the validity of the concept. Thirdly, because it exists objectively, it thus seems to evade any transfer or projection of human intentionality or anthropocentric terms, so objectionable in some teleological thinking. This is not entirely so, however. The very concept of program derives from the programming of computers, which is a human action. (Indeed, a 'computer' was initially a person doing computations.) Even Mayr's formal definition, involving "coded or prearranged information," can be misconstrued to apply to inanimate processes, especially when causal powers are credited to "information" as a basic ontological category of physics. This is an unfortunate byproduct of a view of the universe as a mechanism—even literally a computer, as some would hold, with the laws of physics as its program.

5. Causal laws and teleomatic processes (convergences)

Let us begin to zero in on answers to the target question: *what is the relationship between mathematically describable causal laws and apparent or real teleology in nature*? We have broken down "apparent teleology" into sub-categories: the teleomatic and the teleonomic. What then is the relationship between causal laws and the teleomatic? Simply put, the relationship is trivial. Mathematical laws can be formulated to describe processes that appear to converge on some end. Such laws are often called causal, though they are purely descriptive. In particular, they are not pre-existing or prearranged, have no governing power, and do not foresee or guide physical processes toward an end. This is not to deflate the meaning of causality as a human category or deny its usefulness. It is only to say that, for most inanimate physical processes, the appearance of convergence towards an end is a result of nothing above and beyond the causal processes themselves. There are no genuine goals involved where there are no agents in play.

6. Causal laws and teleonomic processes (programs)

While laws do not govern, a program governs because it is constituted to do so. In other words, while a law is an empirical generalization of observations, a program is a set of commands to achieve an end. This distinction trades on an ancient ambiguity in the meaning of law, as originally referring to the *edicts* of an authority (such as God or the king), then extended to mean *regularities* in nature. A program, as a set of commands, rules by edict. However, it is speculation to propose that some natural structure (such as DNA) serves as a program—to *generate* regularities that have been observed and formulated as empirical laws. The "program" cannot be thought of in the fixed and deterministic way implied by the usual understanding of household computer programs. The organism is not a robot, but self-programs, self-updates, and adapts constantly to the world (and, as it is able, adapts the world to itself). The information contained in DNA represents but a fraction of this interactive process.

Natural selection is a causally describable process that can give rise to biological programs such as the genetic "code," without recourse to any notion of future goals since it operates strictly by rewarding current success. Because of ever-changing environmental conditions, selection does this precisely by *not* committing to any future goal [Mayr 1974].

7. Causal laws and genuine teleology (agents)

Cosmic teleology, in its various forms over the ages, has essentially projected human intentionality onto nature, or onto a god who then exercised divine intentionality to create the natural world. The world was comprehensible *because* it was created. This is literally a very different matter than a world that is simply come upon, pre-existing and perhaps eternal. We have a privileged access to human intentionality, which alone we can know from the inside. Our own consciousness invokes the unique domain of reasons, purposes, goals, intentions, etc.—expressions of our own agency. All other domains in which we claim that such things apply are no more than projections of our own agency: for example, as acts of God, as tendencies within nature, or as goals imputed to creatures. The positive side of this limitation is that we know that "mathematical laws" and "aims and intention" are compatible with each other *insofar as they are alike expressions of human agency.* But no further.

8. Agency in nature

One task is then to clearly distinguish such projections from any real agency that occurs in nature. An agent is an autopoietic system, and the only such systems we reliably know of are individual organisms. Other systems (such as species, ecologies, the biosphere, self-organizing systems, etc.) are only agents if they are truly autopoietic systems—that is, not only self-organizing, self-maintaining, and self-reproducing, but also self-defining. This is a narrow category, taken from the obvious example of living organisms. It is conceivable that other autopoietic systems exist—even that the cosmos at large is such a system. However, in the case of any particular system, it must be demonstrated that it satisfies the definition of agent. Otherwise, one simply indulges in projection of some sort. That is not to deny the pragmatic benefits of such projections. It is mathematically *useful* to treat some systems in terms of action principles, attractors, etc. It is *convenient* in ordinary language to speak of aims and intentions when dealing with creatures and even with inanimate systems. It should be borne in mind, however, that such figures of speech (or mathematical devices) reflect our own strategies and not necessarily the natural world.

9. Conclusion: formal description of goal-oriented behavior

Is the mathematical description of teleological behavior feasible? In this context, the paradox of mathematics is that it transcends physical reality while derived ultimately from it by conscious agents. I take the position that mathematics, like science, is a form of human cognition. As the most generalized and abstract form of cognition, it may be uniquely qualified to describe agency itself.[3]

Yet, mathematical models are essentially simplistic, whereas nature is essentially complex—perhaps indefinitely so. The prospect of mathematically describing complex processes, leading from molecules to “aims” and “intentions”, is sufficiently daunting to warrant a negative appraisal. To be sure, there are projects underway to map the micro-neurology of the brain, just as there are projects to map various genomes. I am of the opinion, however, that no aspect of nature can be *exhaustively* mapped, because any such project already imposes its own limited models and conditions on what is sought and believed to be significant. This will always leave something out of the picture, perhaps crucially. As Vico held, the only things we can know with complete certainty are things we have made, never the ones we have found in nature.

To formalize goal-oriented behavior is effectively to program a robotic agent. But an agent is effectively self-defining; the idea of programming a machine to be self-programming may be paradoxical.[4] And yet, given that *organisms* exist, and *if* organisms are indeed self-programming machines, then perhaps it is feasible. After all, *nature*

---

[3] One heroic attempt to formalize *perception* [Bennett et al. 1989] avoids crossing ontological category barriers by postulating that “objects of perception” are other “observers,” defined in formal notation. In other words, there need be no direct reference to a real external world even in describing perception. I do not believe we need to go to that extreme, so long as perception is counted a form of behavior, motivated by creaturely existence in the natural world.

[4] Arguments against the machine modeling of mind have been proposed on the basis that any fixed system can be transcended only by a human being [Lucas 1961]. An autopoietic system, however, is not fixed; it has not been defined by an external agent, but is self-defining. As such, it potentially has the same capacity for transcending its own state (as given to itself) that the human observer has to transcend the given state of a “machine.” The agent we account for must be able to logically transcend states of itself it has been able to identify. In short, it must be an organism capable of self-reference.

"programmed" organisms to be self-programming! Whether one can discover or reproduce the dynamics of that process is another question: that is, whether *we* can write a program for a self-programming machine. In any case, the actions of an agent are not determined by "mindless mathematical laws," but rather determine them. An agent's goal-oriented actions are not produced either by efficient physical causes outside it nor by someone else's goals. They mean something to the agent itself, in its own terms. The most appropriate description of such a system would be from *the system's own point of view*, or one that allows for it. Formulating such a type of description could be the basis of a future research program.

It appears that *nature's* programming happens through natural selection among incremental changes—from the bottom up, so to speak—in contrast to the top-down rational planning by a pre-existing intelligence. Artificial self-organizing networks of connections (neural nets) exist, of course, but there is no guarantee that we can understand how such programs self-organize or how they solve problems we put to them, any better than we can understand how organisms solve the challenges we perceive that nature puts to them. Even if they are not true autopoietic systems, we can now build/program machines that program themselves; but we have little idea how they work. Such machines (programs) appear to manifest aims and intention, but these reflect human aims and intention and their powers remain within the limits of problems defined by humans. If we cannot understand even this restricted case, can we hope to understand the case of genuinely autopoietic teleology?

The relationship between cause and goal is that causality, like teleology, is a category of thought and description applied by *observers*, whereas *agents* pursue goals. An observer is an agent, of course, and an agent may be an observer. But they are essentially different roles, conventionally involving first and third "person" points of view. Science has traditionally excluded first-person description from its accounts, and with it agency and teleology. To account formally or mathematically for natural goal-seeking behavior, in a logically consistent manner, requires re-instating the fundamental notion of agency.

REFERENCES

d'Abro, A. (1939) *The Rise of the New Physics* Vol 2 Dover, p265

Bennett, Bruce M.; Hoffman, Donald D.; Prakash, Chetan (1989) *Observer Mechanics: a formal theory of perception* Academic Press

Bruiger, Dan (2014) "An Argument for a Second-Order Cosmology" Academia.edu

Bruiger, Dan (2016) *The Found and the Made: science, reason, and the reality of nature* Transaction Publishers

Bunge, Mario (1979 [1959]) *Causality and Modern Science* Dover

Chaitin, Gregory (2009) "Leibniz, Complexity and Incompleteness" *APA Newsletter on Philosophy and Computers* Vol 9, No.1, pp7-10

Lucas, John R. (1961) "Minds, Machines and Gödel" *Philosophy* 36: 112-127

Maturana, H. and Varela, F. (1980) *Autopoiesis and Cognition* Reidel

Mayr, Ernst (1974). "The Multiple Meanings of 'Teleological'" (PDF). Boston Studies in the Philosophy of Science.

Pittendrigh, C. S. (1958) "Adaptation, natural selection, and behavior," in *Behavior and Evolution*, ed. A. Roe and G. G. Simpson, New Haven: Yale University Press, 390–416

Reese, Hayne W. (1994) "Teleology and Teleonomy in Behavior Analysis"
The Behavior Analyst 17, 75-91 No. 1 (Spring)

Stöltzner, Michael (2003) "The Principle of Least Action as the Logical Empiricist's Shibboleth" *Studies in the History and Philosophy of Science Part B* 34 (2):285-318

Tegmark, Max (2008) "The Mathematical Universe" *Foundations of Physics* **38**:101–50.