Explaining Biological Functionality: 
Is Control Theory Enough?

John Collier
Philosophy and Ethics
University of KwaZulu-Natal
collier@ukzn.ac.za

Abstract
It is generally agreed that organisms are Complex Adaptive Systems. Since the rise of Cybernetics in the middle of the last century ideas from information theory and control theory have been applied to the adaptations of biological organisms in order to explain how they work. This does not, however, explain functionality, which is widely but not universally attributed to biological systems. There are two approaches to functionality, one based on etiology (what a trait was selected for), and the other based in autonomy. I argue that the etiological approach, as understood in terms of control theory, suffers from a problem of symmetry, by which function can equally well be placed in the environment as in the organism. Focusing on the autonomy view, I note that it can be understood to some degree in terms of control theory in its version called second order cybernetics. I present an approach to second order cybernetics that seems plausible for organisms with limited computational power, due to Hooker, Penfold and Evans. They hold that this approach gives something like concepts, certainly abstractions from specific situations, a trait required for functionality in its system adaptive form (i.e., control of the system by itself). Using this cue, I argue that biosemiotics provides the methodology to incorporate these quasi concepts into an account of functionality.

Introduction
It is generally agreed that organisms are Complex Adaptive Systems. Since the rise of Cybernetics in the middle of the last century ideas from information theory and control theory have been applied to the adaptations of biological organisms in order to explain how they work. This approach has had considerable success in some areas, and many hoped that the same general principles of Cybernetics could be applied to machines, organisms and minds. The limitations of what is now called First Order Cybernetics became apparent towards the end of the century. The main problem was that functionality was deeply entwined with the autonomy of organisms and minds, as developed by people like Maturana and Varela (autopoiesis) (Maturana and Varela 1972) and one of the originators of complex adaptive systems theory, Robert Rosen (closure to efficient causation) (Rosen 1991). The problem is that First Order Cybernetics is notably mechanical in nature, depending only on inputs, outputs and internal state transitions. The system is capable of adapting its behaviour to its environment (within a limited
range), but it is not capable of controlling itself so as control how it interacts with its environment.

In order to deal with this problem, the relatively obvious solution was to add another layer of control, yielding what is called Second Order Cybernetics, which became popular in the last couple of decades. Originators of the approach are Heinz von Foerster (1974), Gordon Pask (1966), Gregory Bateson and Margaret Mead (1973). Others, like Maturana and Rosen (cited above) took up the ideas. The basic idea is to have a controller within the system that controls the system's responses, giving it some degree of autonomy. The basic idea is shown in Figure 1.

**Figure 1**
(Figure courtesy of Wikipedia Commons)

![Figure 1](Figure.png)

Two feedback loops are used instead of the one found in First Order Cybernetics, where an engineer designs the control parameters for the system. Instead, the control of the behaviour of the system is placed inside the system.\(^1\) The idea is that such a system is autonomous. It turns out that in theory a two level control system can control anything in principle. The devil, however, is in the details. Actually creating a dual level control system for a particular control problem can be exceptionally difficult.

\(^1\) See Heylighen and Joslyn (2001) for a good account of Second Order Cybernetics and its contrast with the first order version.
Unfortunately, the calculations required for biological systems for this design, using methods inherited from first order cybernetics are implausible.

To deal with this problem, C.A. Hooker, together with two control theorists responsible for the control system of the F16 fighter, Ron Evans and Bruce Penfold (Hooker et al 1992), argued for a new paradigm of control that just requires local control, and not full differential equations for the system. This greatly reduces the computational load. The second layer of control in Second Order Cybernetics needs a kind of abstraction of the control problem at the first level, and their approach requires only local abstraction. In that respect at least (their claim is somewhat stronger) the local controller of the behavioural parameters of the system is like an embodiment of a concept. If so, this is quite remarkable, since Second Order Cybernetic systems are found, like the first order version, across machines, organisms and minds. All three, if their adaptivity is organized along Second Order lines, have something like representations. We then might want to ask if there is anything that separates this representational character of the F16 from that of an E. coli or of a human. Intuitively we would say yes, F16s do not represent anything in and of themselves (though they may well represent something to us, or perhaps even to E. coli), so either our intuitions are wrong, or else the Second Order Cybernetic model is incomplete for bodies and minds, and it needs to be augmented. This, then, is the problem for explaining functionality: what more do we need beyond control theory to account for functionality and avoid the symmetry problem. The answer, I will argue, is in autonomy of organisms, which takes a special form, and that the role of the concept like things in autonomous systems of this sort need to be understood in terms of the emerging field of biosemiotics.

Control
Control fundamentally involves the transmission of signals that trigger, cause or imply further activity, typically of a different sort. Cybernetic systems are control systems that use information to direct other forms of flows (including other information flows). This suggests that information directing flows is sufficient to explain the behaviour of organisms. Information theory is useful for understanding control because transmission of information can typically use very small energies to control much larger energy flows. According to Biological Information Theory (BIT), information theory is a useful technology for representing biological systems, but more than that, biological processes involve not only matter and energy flows, but also information flows. Typically the energy and information budgets are to some extent independent, and information theory is more than just instrumentally useful, but has a substantive application within biological systems (Collier 2008). The information flow in biological (and other) systems can be unified with causation (Collier 2011), but the partial decoupling of energy and information budgets in biological systems allows an analysis of biological control internal to the system that can be integrated into control theory. The details are rather technical, and I will not go into them here. One thing to note, though, is that the theory of information flow in distributed systems (Barwise and Seligmann 1997) requires that information tokens, which carry information, are classified into types. The simplest way to do this is with a code, such as the genetic code, but more generally it can be done through physical information systems (Collier 1986), which allow the combination and recombination of basic units, such as macromolecules. In either of these approaches no interpretation is assumed, but the classes must be objective, laying the ground for possible interpretation. Although BIT can give a good account of
information flow in biological systems, and can form the ground for a control theory account, in itself it has problems from the perspective of explaining biological functionality. First of all BIT is purely syntactic; it has no semantics in itself. Consequently BIT is not able to give a complete account of function. Functions have goals, or aims, whereas information flow is neutral. Given the way that information is carried in an organism, it has a meaning (or something similar to a meaning), this means a semantics is needed. The information in a hormone, for example, is about some state of the body. The information in nerves are about some state of perception, though not about in the same way as thoughts are about. There is no nontechnical term for this, but the biosemiotics community has called ‘interpretant’, a term from C.S. Peirce. The interpretant is what makes sense of the information. Another problem is that organisms’ actions make sense in a way that those of rocks and waves do not. This making sense is the basis for thinking they are meaningful. They have a point to them

This suggests that we might be able to invoke the principles of control theory to add what we need to get to functionality. In the case of First Order Cybernetics, given a distributed network, the mutual information among nodes defines the properties of the relations in the network. Furthermore, feedback and feedforward allow for control of the information flow through the nodes. This approach allows us to analyze the properties of control in specific network organizations. It can explain how a system behaves in terms of the relation of information and flows, but not why. In order to explain function we need to postulate some sense of something being done for the sake of something, in this case, the organism. In particular, we need a way to understand how an organism can do something for its own sake. This requires some degree of autonomy.

Second Order Cybernetics was introduced specifically to try to take account of autonomy by placing control into a higher order system within the organism that controls its behavioural parameters. Although I will argue that Second Order Cybernetics is not sufficient to explain autonomy, let alone functionality, autonomy requires Second Order Cybernetics. Autonomy requires self-regulation, which requires that the steerer is also steered within the system. Standard approaches to Second Order Cybernetics based in the principles of First Order Cybernetics require that the system has a model of all possible behaviours it can engage in. Specific behaviours are chosen to satisfy constraints on the second order controller, whatever they may be. Everything is already there. To call this choosing is a bit weak. This, though, as mentioned above, is computationally implausible for organisms and our limited minds.

I think the problem can be resolved by moving to a new form of Second Order Cybernetics. The major innovation of Hooker, Penfold and Evans (1992) was to limit the requirements of control to local variations in behaviour, so that the second order controller can check the behaviour at a given time, and see if it is design parameters. If it is not, it adjusts the behaviour of the system locally to reduce the error. This is far more computationally tractable, as it focuses not on all possible system behaviours, choosing the one that reduces the error in behaviour, but on revising immediate responses to reduce the error. As mentioned above, Hooker, Penfold and Evans suggested that the second order controller functions like a concept that must be satisfied by the behaviour of the system. I think there is something right about this, but there is a serious problem. Because the control is local, there is no reason to think that the various control processes in the system are integrated in a way to produce coherent system behaviour on the large scale. Although we have an account of functioning, we do
not get an account of function from this model. This can be made apparent by compar-
ing the control of an F16 (for which the paradigm was effectively applied) and that of
an E. coli bacterium. F16s don’t represent anything to themselves in any sense and the
F16 system itself serves a function within a larger system, so its self-controlling capac-
ities gain their function derivatively. Furthermore, it could be said that the controller
of the F16 uses the airplane to achieve its functions (deadly flight). This symmetry is
fatal to our usual understanding of biological function. E coli cells don’t function for
anything but themselves, unlike F16s. There is also no symmetry: the functions of E
coli function for the cell (and perhaps its lineage, but then these are really functions of
the lineage), but the cell does not function for its parts. We need to account for these
two properties in explaining biological function. I believe this symmetry problem has
been overlooked in many accounts of biological functionality.

Function
There are two competing accounts of biological functionality. The one most popular in
the English speaking world is the so called etiological account, due originally to Larry
Wright (1973). On this account, the function of T is to F if and only if T exists because
it Fs. A biological trait T, on this account, is functional if and only if T is an adaptation
due to its Fing (T was selected because it Fs). For example, the function of the heart is
to pump blood because it was selected for due to its pumping blood. There have been
objections to the etiological view on the grounds that we do not need to know the eti-
ology of a trait to know its function. Furthermore, it seems to put the cart before the
horse: a trait is selected because it already performs a function. Proponents of the etio-
logical approach, however, are willing to swallow these consequences, and simply
deny that a trait is functional until it is selected, and that we really can’t know the
function of a trait unless we know its actual selection history. It seems to me, however,
that there is a fatal flaw with this account due to a symmetry problem similar to the
problem in the F16 case. The organism exists because it supports the heart (if it did
not, the heart would not exist). So the function of the organism that has a heart is to
support hearts. This is ridiculous, of course, but I think it is all the worse for the etio-
logical account. The organism exists only if it supports hearts, so the function of the
organism with a heart is to support hearts. The analogy to the F16 case seems to me to
be exact.

The other view of function is more popular in continental Europe and South Amer-
ica. This is the autonomy account, favoured by Maturana and Varela (1972) and Rob-
ert Rosen (1991), but which arguably goes back to Kant. It is basically an organiza-
tional account. The autonomy of an organism is that organization that maintains the
organism's viability (Collier 2004b). Something is functional for an organism if it con-
tributes to its autonomy. Note that biological functionality, on this account, is likely to
be selected for. This allows some reconciliation with the etiological account, though it
is functionality in general, not specific functions that is likely to be selected. Given
enough modularity, we can speak of specific functions of traits, but the approach also
allows for integrated and holistic functionality. The symmetry problem is avoided,
since functionality is primarily an organizational property of the organism (though
there is a residual worry that symmetry might sneak back in because of the environ-
mental role in organism viability – this will need to be dealt with in a full account of
functionality).
One point that is worth noting: Both Maturana and Rosen require complete closure (autopoiesis and closure to efficient causation, respectively). Autopoiesis is the process by which things can be self-maintaining. In Maturan's case (he invented the term), the maintenance must be entirely inside the organism. Closure to efficient causation occurs when their closed causal loops among modules. This happens when one thing controls another, and it controlled by a possible succession of other modules. The closure condition can be relaxed, giving an account more compatible with open systems accounts of complex adaptive systems, by allowing for environmental interaction that forms loops, allowing feedback through the environment, and offloading some control into the environment, further reducing computational load (Collier 2004a, 2006). For a good example of this, see Paul Churchland's crab model (Churchland 1986), in which functioning is determined in part through environmental interactions, allowing the crab's vision to interact with what it is observing and its motor functions so as to grab an object. Hooker, Penfold and Evans (1992) cite this case approvingly as an example of local control. If we assume that real crabs are so organized, and the object is a tasty morsel, we can say that the autonomy of the crab is organizationally extended into its environment.

**Biosemiotics to the rescue?**

Using the Hooker, Penfold and Evans idea that their second level controllers are analogous to concepts, or signs for something else, we can take this seriously and try to use the theory of signs and their interpretation, or semiotics, in order to analyze the functioning of biological control systems. Fortunately, there is an emerging field call biosemiotics that is developing the required technology. If we can understand the role of signals in general, and how interpretation takes place in each of machines, organisms and minds, if it does at all, then we may be able to distinguish the sort of control peculiar to machines, organisms and minds respectively. Taking it as obvious that current machines like the F16 or IBM's Big Blue do not interpret, we can say that semiotics is irrelevant to the functioning of such things, and any functionality that they have is not their own, but is derivative (which leaves such cases open to the symmetry problem). Minds, on the other hand, have explicit representations that we don’t expect to find at the biological level (our gut may have more nerves than our brain, but it at best produces feelings, not explicit representations like the brain does). I am going to ignore strictly mental functions here, and concentrate on biological function. The branch of semiotics that applies to biology has roots in the early part of the Twentieth Century in the Baltic region, and has matured recently into a respectable science, with its own books series and journal. A number of issues are still controversial, though, and I will be taking distinct sides on some of them. What is biosemiotics, and how can it enhance control theory to give an adequate explanation of biological function?

Biosemiotics is first of all an empirical hypothesis that stands on its own, irrespective of its relations to other theories and methods. Specific methods and a core theory specify what it is to be a biosemiotic system. It is an empirical question what real systems (if any) fit this general model. One of the deep problems in current biosemiotics is what are the minimum set of tools required for it to be useful.

Saussure is responsible for the branch of semiotics called semiology. On this approach a sign represents something, standing for it within a system of signs. The most important case is a word that stands for an idea in someone's mind. Applied to the biological case a gene might stand for a given trait. However a trait might equally stand
for a given gene. On the two-valued (dyadic) semiological approach it is inevitable that signs and their objects can be inverted. Ideas, for example, could equally well be signs of words. It is purely arbitrary to declare one the sign and the other the object. The symmetry problem is acute. The same symmetry problem arises for so-called biological codes, of which the genetic code is the most famous example. While genes can be taken as signs for traits, it is equally plausible to take traits as signs for genes. For example, blue eyes indicate a certain genetic makeup. We need to block this symmetry.

Traits are often thought to gain their function by being selected by the environment, and the genes underlying the traits are often said to represent the environment. However it is equally true that the environment contains signs for traits, and therefore genes. Is it therefore the function of the environment to select traits and their generating genes, as I suggested above? I think not. It seems that an explanation of function grounded in dyadic relations will inevitably have problems.

Peirce developed a triadic account of signs in which the triads are not reducible to combinations of dyadic relations. There is the physical or material form of the sign (a quality, dubbed and icon), its object, and its interpretant, which gives the sign its meaning, or in the biological case, I will argue, its function. According to Peirce a sign must be interpreted to be connected to its object. This is done through interpretation, the function of the interpretant role of a triadic sign. Basically, it is what gives the sign sense. The whole can be the material form for a further sign, and so on until we reach some final interpretant. For example, we may have a particular experience that we take to resemble other experiences that we commonly call red, and infer that the experience is of something red. In the case of traits, we have a trait that in the context of a particular organism contributes to its (likely) survival. The trait is thus ultimately a sign of survival (and not of specific aspects of the environment, though a biologist might take the trait to be a sign of a particular environment).

The requirement of the interpretant for a sign makes the inversion of sign and object impossible: interpretants require a certain degree of organization, which environments do not imply. Certainly there are highly organized environments, but these are typically constructed by organisms, and are thus part of their biology, often social. Language is an especially strong example of this.

It will be useful to give some examples. Suppose we have smell A (icon) of something dangerous (interpretant), then it is incumbent to avoid (object). For good biological reasons, this sort of interpretation has very short chains. Smell is a primary guide as to whether something is edible or dangerous to eat. Chains related to survival are typically longer, and can form complex networks. Figure 2 shows the Peircean sign involved, as an irreducible triad. The sign itself (the whole triad) can be an icon for a further sign (or signs), which have different objects and further interpretation.

Suppose we have smell B (icon) that indicates food (interpretant) that can be eaten (object). This itself is a sign (icon) that falls under survival (interpretant) indicating it should be
accepted (object). Typically there will be longer chains both for the interpretant and often the icon. This is shown in Figure 3. The smell B is a sign of food, and the interpretation is to eat it. This sign has the object “accept” with the interpretation of survival. This final interpretation, “survival”, is as far as we can go. It is the ultimate interpretant. The function of Smell B is to indicate food, which is responded to by the inclination to eat. This second sign has the object to accept (and eat), which has the interpretation of survival. Survival is the basis of organizational function on the autonomy account. If the account is correct, then all signs ultimately lead to an interpretation of survival or its opposite, death (see Figure 3).

Autonomy, then, involves not just a dual level form of control, but integration into an organization that increases the likelihood of survival of the thing that has it. The overall organization is what promotes survival, and individual traits are functional inasmuch as they contribute to this. The “meaning” in functionality is ultimately survival in biology, though it may be something else for minds and artefacts. The second-order cybernetic “concepts” of Hooker, Evans and Penfold are functional just inasmuch as they promote survival. They represent just inasmuch as they are ultimately signs of survival. The account rules out fighter jets from being functional independently of their use by us, and rules out their having representations, except derivatively. The problem of bidirectionality of functionality between environment and traits is also resolved. This is because the ultimate interpretant (survival or death, as the case may be), is internal to the organism, following from the technology of Peircean biosemiotics. Furthermore, biological traits are not functional independently, but only inasmuch as they contribute to autonomy.

Summary
In summary, BIT does not account for function or semantics, but does account for information flow, networks, control, higher order control, and the dynamics of biological form. The operations of BIT are all included under the theory of information flow. The theory of information flow requires types classifying tokens (Barwise and Seligmann 1997, for details see their book or Collier 2011). This is best achieved with a code (or, more generally, a physical information system, Collier 1986). The classifications required for distributed information flow imply nothing more than that certain tokens are filtered by something that puts them into groups that are system indistinguishable. Nothing is implied about interpretation, meaning, or anything of the sort.

In some cases functional systems make an internal distinction that have something very like semantic value (e.g., signs for something else), implying semiotic values. On
the autonomy account of function, all functionality implies these quasismismatic values, and hence semiotic values, the most basic of which are signs for survival and signs of death. The problem of symmetry is avoided by making all interpretation internal to the organism, or at least to its environmentally closed processes.

References

Bateson, Gregory Stewart Brand and Margaret Mead. 1973. Interview with Gregory Bateson and Margaret Mead, CoEvolution Quarterly, 32-44.


