Behaving Well

Ranulph Glanville
CybernEthics Research, 52 Lawrence Road, Southsea, Hants PO5 1NY, UK; and
The Bartlett, University College London, 22 Gordon Street, London WC1H 0QB.

email ranulph@glanville.co.uk.

Abstract

One way of thinking about second order cybernetics is that it's more a way of looking at the world than a subject. Second order cybernetics is concerned with circular causality and interaction (it is also known as the cybernetics of cybernetics). For von Foerster, father of second order cybernetics, this was best expressed through what he called his "ethical imperative:"

Act always so as to increase the number of choices.

It can be seen, thus, that ethical considerations are at the heart of second order cybernetics. In this paper I explore the ethical dimension of second order cybernetics to consider what its ethical implications are, by considering four key cybernetic understandings and their ethical implications, which include, for instance, autonomy, responsibility, sharing (and interaction); and the obligation to act with generosity towards both each other and the environment.

Implicitly, consideration is given to how these understandings cast light on how humans might best (inter)act—with animals and machines.

Introduction

Students of second order cybernetics have long made assertions concerning its ethical implications. At an earlier conference in this series, I spoke about attitudes to causality and blame (Glanville 1995). I have also written about generosity and the need to listen (Glanville 1991, 2001b). Heinz von Foerster, who invented the approach that is second order cybernetics, was also explicit about this. He published an aesthetic and an ethical imperative:

act always so as to increase the number of choices.

(which has an accompanying aesthetic imperative:

if you desire to see, learn how to act.)

The recent book of von Foerster's selected papers in cybernetics and cognition includes the explicit "Ethics and Second Order Cybernetics" and the less obvious

"The Responsibilities of Competence" (von Foerster 1992, 1972). Von Foerster distinguishes ethics from morals thus: ethics come from within, whereas morals are imposed from inside.

Other scholars in the field—Maturana, Varela, von Glasersfeld, Scott and Pask amongst them—have made statements and argued positions that explore more or less explicitly ethical considerations and implications of their discoveries.

It is therefore reasonable to consider some of the ethical implications of the approach that is second order cybernetics, in spite of Wittgenstein's familiar warning that ethics is an area for action, not argument! This paper will essay this. To make sure there is an original contribution I shall develop my argument primarily within the framework of my own work. And to make sure that I am talking of ethical considerations, rather than moral, I examine what arises out of certain characteristic second order cybernetic models, rather than what can be applied to them.

Autonomy

Autonomous systems have been characterized as organisationally closed but informationally open.[1] This means their organisation is sustained from within the system itself, that the organisation organises itself without reference to any external factors. The prototypical cybernetic organisationally closed system is the autopoietic system, which Maturana (1981) has described thus:

> *"[Autopoietic systems are]… a class of mechanistic systems in which each member of the class is a dynamic system defined as a unity by relations that constitute it as a network of processes of production of components which: (a) recursively participate through their interactions in the generation and realization of the network of processes of production of components which produce them; and (b) constitute this network of processes of production of components as a unity in the space in which they (the components) exist by realizing its boundaries."*

The organisation of a system can be said to generate its boundary (generally through interaction with an environment) and to be that which determines the changes in behaviour (i.e., signal, i.e. information) that are observed in it. There may be no actual connection—indeed, there may never be any possible demonstrable connection—between the organisation and the observed behaviours. Thus, the account which attributes to the organisation a causal effect on the observed behaviours locates the cause in the realm of the observer's account of his/her observations, rather than in the action of the organisation on the signal. The device through this may be studied *in extremis* is the Black Box (see later).

---

[1] They may also be energetically open.

Nevertheless, we talk of the observations we make of a system's behaviour as if they were the systems' behaviour, and the structure we deduce as causing observed changes in behaviour as if it were actually causing this; whereas whatever cause we observe lies in our observations and our reflections on these observations, where we locate our explanations.[2]

Pursuing this way of talking, we can say that autonomous systems are organised so that they not only maintain themselves but also so that they maintain themselves in the face of signals that they process and which are expressed as behaviour. The (observed) change in behaviour is caused by the (internal) organisation of the system. Thus, the response of the autonomous system is entirely performed within the system itself.

Under these circumstances, it is difficult to argue that the behaviour of the system is not the responsibility of the system itself. The system's organisation is so arranged that it sustains itself, and it processes signals accordingly: and that is the source of its autonomy and its responsibility. We are thus responsible for our behaviour. As an extreme example, and to rephrase Sartre (as reported in Chödrön 1997): while we may not be able to control when and how death will present itself to us, but we can control how we face it.

The argument, however, should not be allowed to rest here.

There are systems that observe others while those others observe them. These systems are closed. One obvious example is the conversation (to which we will return), where system A observes system B and system B observes system A, thus maintaining a circularity (I use the term observe in the most general sense). The progenitor of Conversation Theory, Gordon Pask, described a conversation as organisationally closed and informationally open, a phrase also used to describe autopoietic systems.

Thus, there are systems in which the observer is included in system. Second order cyberneticians would argue that most (*in extremis*, all) systems are of this sort, and that those which seem not to be are essentially trivialisations of those that are. Indeed, the very description of the autonomous system given above indicates that it is such a system. It has been emphasised that the behaviour is not the behaviour of the system, but the behaviour observed to be of the system: and thus the observer is incorporated.

Take this to the extreme. Let the system A observe itself. This is without doubt a formulation of an autonomous system. If there was any room to doubt the autonomous system's responsibility for its actions, there is now none: the observations are now clearly of and in the system. Expanding this to our earlier descriptions, we see that, while a system may be deemed to be autonomous by an observer, there is another autonomy possible: the autonomy where the observer is

---

[2] Science works, I claim, at the level of description (of some supposed reality), developing descriptions of descriptions that act as explanations. See Glanville 2001a.

within the system. In the example we began with, the system we judged autonomous can be taken to have the character of a self observing system, but our interactions with it as observers asserts that there is another autonomous system which exists when there is organisational closure between the observing and the observed systems. Furthermore, the observed system is, itself, in this paradigm, (in the role of) the observing system of the observing system (which become its observed system).[3]

Such systems are called, but the author, Objects: self-observing systems which generate time through the alternate assumption of the roles of self-observing and self-observed (Glanville 1975).

For all these systems, autonomous systems, we can say:

Autonomy = Responsibility

As I have argued at an earlier conference, autonomy denies us the possibility of "Chasing the Blame." The Buck Stops Here (Glanville, in this series, 1995).

Distinction

How can we congise/(re)cognise/recognise (observe) a system, an object, a thought? George Spencer Brown developed a logic of distinctions that admits to formal logics the notion that it is we who distinguish what we distinguish: that it is not distinguished *a priori*. In the "Laws of Form" Spencer Brown (1969) commands us to "Draw a Distinction!" and then proceed to examine how an entire (and tremendously powerful) system of logic can be developed originating in this simple command.

Drawing a distinction cleaves a space: it creates (in Spencer Brown's terminology[4]) a mark and a value. The mark is the mark of the distinction that cleaves the space, and the value is the value of the space cloven (on one side). It matters not which side the space is: one clever aspect of Spencer Brown's calculus is that, unlike more traditional logics, the spaces on each side of the distinction (which, in the propositional calculus, would be denoted p and not-p) do not have to add up to totality.

The distinction, cleaving space, thus makes the space of each (both) of its sides, as I have argued (Glanville 1990). (Spencer Brown's analysis concentrates on just one side.)

---

[3] There is a critical problem that relates to this way of looking: the difference of view from within and without a system. This relates to how we understand distinctions (discussed in the next section). I shall not attempt to deal with this problem here. Although central for any account of the observer and his/her relationship to the observed, it is not central to the arguments in this paper concerning how to behave well.

[4] I have long argued there are limitations that Spencer Brown is reluctant to admit in his calculus. In this paper, however, the aim is not to contest such points—which do not affect what is being argued here—and so, for the moment, I ignore them.

Furthermore, we can ask what would be the point of drawing any distinction if this action did not distinguish at least two items. To distinguish one is to make no difference: rather as Gregory Bateson claimed about madness, it's a social matter. There is no self, for instance, without another (how else could you know there was a self), just as there is no madness without the sane to compare it to. To distinguish oneself (in the formulation used above, to observe oneself, which is to bring oneself into being in a universe of observation) is to do nothing. If I am all, I am not.

Distinguishing, thus, always involves a triad: the distinguishing, and both distinguisheds, which are brought into being through the act of distinction drawing.

At this most fundamental level of the coming into being, the beingness of each side of the distinction results from the initial act of distinction drawing. Since every possibility depends on this coming into being, and the quality of coming into being is the same and has, for each side, the same origin, it follows that what is possible on one side is, in principle, possible on the other side.

Thus, for all distinctions we can say:

Distinction = Mutuality

I have stated this more fully as the "Principle of Mutual Reciprocity", which goes:

> *That which is possible on one side of a distinction is possible, at least in principle, on the other side.*

Thus, whatever I might (come to) attribute to me I might attribute to you. Whatever I attribute to you I might attribute to me. And (when we have developed a means to communicate), whatever you attribute to you, you or I might attribute to me, as whatever you attribute to me I might attribute to me, also.

Might does not mean will, it does not mean must. It only means it must be possible. Not every possibility must be taken up.

Conversation

Autonomous systems understand what they understand in their own ways. Their understanding is theirs. This is a consequence of both their organisational closure and their responsibility for their own understandings (and actions). Communication is a consequence of distinction, where drawing a distinction brings into being both the distinction itself, together with 2 spaces, not 1 (that is, both sides of the distinction). It is also how we share and the prototypical form of interaction.

The form of communication that permits communication to take place without the in principle need for the understanding of one participant to be imposed on the understanding of another is the conversation, studied as a general means of communication by Gordon Pask.

In a conversation, with the minimum of 2 participants (call them A and B), the understanding of A is somehow expressed (by noises, movements or whatever) to a second participant B, in a (conversational) space that lies between them. B, sensing this, makes his/her own understanding. There is, thus, no connection at all between the two understandings. B then expresses his understanding in the conversational space so that A can build another understanding, which can be compared by A with his/her earlier understanding. If there is a big mismatch, an error correction strategy may guide the next expression, attempting to allow B to correct his understanding so that A's understanding of B's understanding (of A's understanding) is closer to A's (initial) understanding: an error correction loop.

There is much more to a conversation than this, but this constitutes a minimum specification. There is the question of shortcuts (language); of the number of levels necessary for a conversation to take place on (Pask maintained 2, I maintain 3); of the willingness not to insist on immediate error correction so the conversation can flow and self-correct.

This is an organisationally closed system, which (depending on your position) may be seen as informationally open: so a conversation is autonomous and, even were it not for the feedback mechanism, an essentially second-order cybernetic item.[5]

It also depends on a spirit of generosity. If we are not generous, if one participant tries to impose his/her understanding on the other rather than accepting that each understands differently; or if one participant does not allow for discrepancy and is not willing to allow a little of the oil of mismatch to flow, there will be no communication.

The (technical) conversation of Pask's Conversation Theory is merely a cleanly expressed version of the everyday conversation that we enter into all the time, that guides me in writing this account and you in reading it.

Furthermore, the most familiar experiences of conversation, that we get lost in it, become as one with the others, and end up where we had no idea or intention of coding (which is the source of creativity in conversation) could never happen were a conversation not as it is.

Thus, for communication, for the sharing of being (for that is what we intend when we use this word), we can say:

Generosity = Necessity

This assertion is not unlike von Foerster's Ethical Imperative:

---

[5] The alternative to conversation is coding. I have argued the primacy of conversation over coding—in contrast to the usual position. Coding destroys necessary difference, which is at the base of the arguments put forward in this paper, and the crudeness of which can be seen in the training of people who are obliged to use it: e.g. the armed forces. See Glanville 1996).

Act always to increase options

It also relates to the aim of those who practice the field I mainly teach: architecture. An obituary for the architect Denys Lasdun quoted him thus "Our job is to give the client not what he wanted, but what he never knew he wanted till he saw it." Increase the options, see the best possibility, act in the assumption of the best. It is better to be disappointed acting well than confirmed acting badly. To give you what you want, ask a builder.

How the strange noises we make actually work remains, thankfully, a mystery, as the old master magician reminds us (von Foerster and Poerksen 2001).

Black Box

The Black Box is the supreme and extreme expression that reminds us that we deal with descriptions of observed behaviours (which we describe to make explanations, often in the form of causal mechanisms) rather than the mechanisms of a "Mind Independent Reality," Out There.

The Black Box is a phantasm inserted by an observer when (s)he observes a change in a behaviour (a signal), at the point of that change. The reason for its invention and insertion is to allow the observer to account for the observed change.[6]

The observer, having inserted this phantasm observes the changing behaviour as two behaviours modified by the (invented) Black Box. The observer, having inserted this phantasm, can now attribute to it the change between these behaviours. And having constructed some explanation for the observed change in behaviour, the observer can now change the "input signal" to see if the resulting "output signal" is still covered by that explanation (i.e., is predicted), or whether the explanation itself needs to be changed.

Notice that the observer never gets inside the Black Box (indeed, (s)he can't: it's a phantasm, and it's Black (meaning that what's inside, if anything, cannot be seen—by definition), and that both the explanation (based on descriptions of observations), and Wittgenstein's stricture, that because something has always happened does not mean it will continue to happen, are built in.

The explanation of the change in behaviour is developed by the observer, but it is not his/hers alone, for it comes about through the interaction of observer and Black Box. Each acts on the other, and the description is not of the Black Box but of the observer's observations of what (s)he takes to be the results of this interaction, as exhibited in observed behaviours.

---

[6] I have never been successful in my efforts to trace the Black Box idea to its reported origin in the work of James Clerk Maxwell. The concepts I introduce here derive largely from my own interpretations, particularly in Glanville (1982).

This is why everything in a system involving a Black Box depends on interaction, and the so-called Whitening of the Black Box (the creation of a viable explanation) which is taken to actually be—correspond one to one with—the internal workings of the Black Box, which is surely a fantasy greater than the phantasm that is the Black Box! Whitening is not opening the (unopenable) Black Box, but making a viable explanation in the space between the Box and the observer. Those who suggest that the Black Box can actually be "Whitened," that we can know what happens inside it, see the "real" mechanism, forget both its origin and its purpose. Furthermore, it is reasonable to argue that the observer is, him/herself, a Black Box to the Black Box!

Thus, the explanation built of the descriptions of the apparent observations of the behaviour of the Black Box by the observer, is not the observer's alone. That is, the explanation may be the observer's, but the observed behaviour on which it is based is neither the observer's nor the Black Box's, but exists in the space that is shared between them (even though the Black Box is a phantasm). And the explanation is not of what is in the Black Box, of which the observer remains totally ignorant, but is of the observed behaviours which are shared and are not of the Black Box alone any more than they reveal what is "actually" in the Black Box, of which the observer remains entirely ignorant.

The remarkable power of the Black Box, in this understanding, is that it allows us to create explanations and to behave accordingly when there is nothing but a phantasm, when we cannot know what (if anything) is there: we are always present. Thus, it is the prime depiction of how we can operate as it seems we believe we do, in a universe where we assume is no Mind Independent Reality. It asserts that we live in ignorance.

Thus, for Black Boxes we can say:

Observing = Sharing

When we have sharing we can no longer make absolutist assertions. Our assertions come not just from one of us but from some form of sharing, and they derive not from opening the Black Box but from us building our descriptions and explanations based on these observations of shared behaviours. We remain ignorant of what, if anything, goes on in the Black Box and thus are always in a state of blissful ignorance.

And this tells us we neither have access to that Mind Independent Reality, nor do we need it to be able to make an image of a universe and live in and with that.

Coda

There are two further qualities that may be derived from the above and which are ethical in dimension.

The first is trust. Elsewhere, I have argued at length for trust as an essential quality of participants in second order cybernetic systems (Glanville 2001c). I will not repeat that argument here, restricting myself to bringing to the attention of the reader the relationship between generosity and trust. Without generosity, there can be no trust.

But, almost reciprocally, trust is what makes generosity work. We can sum this by asserting

Generosity = Trust

The second quality is ignorance. The discussion of the Black Box, above, is based in handling ignorance. Ignorance is our reason for invoking the Black Box. But ignorance appears, implicitly. elsewhere in this argument. For instance, trust is a quality that is needed to overcome ignorance—or, rather, the fear we feel that is engendered by ignorance. So we can say, for instance, that

The Black Box = Ignorance

Conclusions

I do not seek to tell you how to behave. I seek only to show the hidden implications in these central cybernetic insights. Some may lead to logical problems in themselves: if we see and make the world in our own images, then allowing that each such image is valid means that I cannot deny the image that insists it is the only one. Indeed, my image, that we make our own images, is an overriding one that some will have images that deny. The way out of this is acceptance.

But that discussion is for another day.

Meanwhile, here are those hidden implications that constitute the meat of this paper, again:

Autonomy = Responsibility
Distinction = Mutuality
Generosity = Necessity
Black Box = Ignorance
Observing = Sharing
Generosity = Trust

References

Chödrön, Pema (1997) When Things Fall Apart, London and Boston, Shambala

Foerster, H von (1972) Responsibilities of Competence, Journal of Cybernetics vol. 2 no 2

Foerster, H von (1992) Ethics and Second Order Cybernetics, Cybernetics and Human Knowing vol. 1 no 1

Foerster, H von and Poerksen, B (2001) Understanding Systems, New York, Klewer Academic/Plenum Press and Heidelberg, Carl-Auer-Systeme

Glanville, R (1975) A Cybernetic Development of Theories of Epistemology and Observation, with reference to Space and Time, as seen in Architecture (Ph D Thesis, unpublished) Brunel University, 1975, also known as The Object of Objects, the Point of Points,—or Something about Things

Glanville, R (1982) Inside Every White Box there are two Black Boxes trying to get out Behavioural Science vol. 12, no 1

Glanville, R (1990) The Self and the Other: the Purpose of Distinction, in Trappl, R Cybernetics and Systems '90 the Proceedings of the European Meeting on Cybernetics and Systems Research, Singapore, World Scientific

Glanville, R (1991) Generosity: an Ethos: an Endpiece, in Glanville, R and de Zeeuw, G (eds.) Mutual Uses of Cybernetics and Science (2 vols.) Amsterdam, Thesis Publishers

Glanville, R (1995) Chasing the Blame in Lasker, G (ed.) Research on Progress—Advances in Interdisciplinary Studies on Systems Research and Cybernetics, vol. 11, IIASSRC, Windsor, Ontario

Glanville, R (1996) Communication without Coding: Cybernetics, Meaning and Language (How Language, becoming a System, Betrays itself), Invited paper in Modern Language Notes, vol. 111 no 3 April (ed. Wellbery, D)

Glanville, R (2001a) An Observing Science, invited paper for special issue of Foundations of Science vol. 6 nos 1–3

Glanville, R (2001b) Listen! (The Listen Inn) in de Zeeuw, G, Vahl, M and Mennuti, E (eds) Problems of Participation and Connection, Lincoln, LRC

Glanville, R (2001c) The Man in the Train: Complexity, UnManageability, Conversation and Trust in Würthrich, H, Winter, W and Philipp, A (eds.) Grenzen Ökonomischen Denkens, Wiesbaden, Gabler

Maturana, H (1981) The Organisation of the Living: A Theory of the Living Organisation, Cybernetics Forum vol. X nos 2 and 3

Spencer Brown, G (1969) Laws of Form, London, George Allen and Unwin

Published in:

Cognitive, Emotive and Ethical Aspects of Decision Making in Humans and in AI

Volume II

Edited
by

IVA SMIT
E&E Consultants, Inc.
The Netherlands

WENDELL WALLACH
WW Associates
USA

GEORGE E. LASKER
University of Windsor
Canada

Published by

The International Institute for Advanced Studies in Systems Research and Cybernetics

11