Dirk Baecker

The Intelligence of Ignorance in Self-Referential Systems

in: Robert Trappl (ed.), Cybernetics and Systems '94. Proceedings of the Twelfth European Meeting on Cybernetics and Systems Research, Vienna, Austria, 5-8 April 1994, vol. II, Singapore: World Scientific, 1994, S. 1555-1562

ABSTRACT: Self-referential systems theory does not provide for a concept of intelligence. There is even a certain resistance to intelligence that seems to block any explicit exchange of concepts with artificial systems theory. This paper describes the intelligence service in self-referential systems as the self-referential and, hence, paradoxical switching from the self-reference of these systems to other-reference. How this might work is shown by means of G. Spencer Brown's calculus of indications and Heinz von Foerster's notion of double closure.

1. Inclusion

One of the peculiar aspects of general and sociological systems research is that there seems to be no need to talk about intelligence. At the very moment of the introduction of self-reference into systems analysis, the search for intelligence was suspended. To continue to speak of intelligence, however, is the hall-mark of any systems analysis that avoids introducing self-reference. Biological and sociological systems theory (e.g., Maturana/Varela 1980, Luhmann 1990) seems to resist intelligence as insistently as artificial systems theory (e.g., Newell/Simon 1976) shuns self-reference. The former appears to say: life, thinking, and communication rely on unfoldment of self-reference, that is, paradox (see also Barel 1989), whereas the latter keeps repeating: there is no way to run any program as long as it risks hitting paradox. There is, however, an occurrence of intelligence in sociological systems theory when Niklas Luhmann (1984, p. 158) defines "intelligence" as an indication that it is impossible to observe how a self-referential system consulting itself chooses one and not another solution for a problem. Similarly, various kinds of paradox in artificial computer science keep popping up whenever it has to admit a "kludge" somewhere that will only be displaced, but not removed, by debugging–and even sometimes happens to be taken as the design itself (see Sharkey/Pfeifer 1984).

I am not sure whether this state of affairs already explains why there is the communication barrier between biological and sociological systems theory on the one hand and artificial systems theory on the other, we all experience. But I think that this lack of communication is placing a burden on the development of cognitive sciences that we should be aware of.

As a sociologist, I am not in a position to show how self-reference might be introduced into computer science (see, however, Weston/von Foerster 1973, Löfgren 1979). Instead, I will concentrate on the question of how we could possibly introduce a concept of intelligence into self-referential systems theory. My attempt will be loyal to self-reference insofar as it emphasizes the service furnished to self-referential systems by means of self-referential intelligence. Thus, I adopt the closure-type description of self-referential systems (see Varela 1984) and will not provide for any possibility to feed intelligence into systems as we feed information into a computer. I take intelligence not to be an input to be fed into a system nor an output to be taken out of it. But it comes very close to be taken as such an input, or output. I will try to show that the intelligence service featured by self-referential systems consists of switching the self-reference of such a system over to other-reference, which is, of course, nothing but self-reference in disguise. This is the paradox intelligence is called upon to take care of.

There are, of course, some ideas concerning intelligence which might prove helpful in showing us a way to develop a self-referential concept of intelligence. Jean Piaget (1947), for example, takes intelligence to have something to do with the assimilation of an organism to its environment, and describes this as accommodation, that is, as adaptation of the organism to its adaptation to the environment. W. Ross Ahsby (1981, pp. 295-306) defines intelligence as the "power of appropriate selection", and Gordon Pask (1970) adds that it is not the sticking to a goal but the ability to set and change goals that proves to be intelligence. Terry Winograd and Fernando Flores (1986) underscore that intelligence appears to consist in the avoidance of breakdown as well as in the handling of breakdown as soon as it takes place.

We take "appropriate selection" to be a very handy definition of intelligence that covers the setting of goals as well as the handling of breakdown, and try to introduce this definition into the philosophical notion of intelligence that attaches importance to the reflective capabilities of intelligence: Intelligence appears to consist in a refraction of, and refraining from, any immediate access to whatever reality (see Gadamer 1967).

There is some kind of hesitation involved in appropriate selection that we are only able to understand when we take reality to be not the reality but a reality watched by an observer who almost always checks on the reality he or she takes to be the reality. Why does he hesitate? What does he know about himself when he starts to think intelligently about it? I assume it is the appropriateness of any selection that gets checked by intelligent double-checking. But how does an observer start to think about the appropriateness of a selection? Well, he tries to think about anything he might have excluded when choosing this selection rather than another. He is thinking about opportunity costs, an economist would say. So what is intelligence all about? It is an attempt to include not only what was excluded but the very fact of exclusion, and even to ask what might be excluded by including the excluded. And that seems to be a very down-to-earth definition of the problem of intelligence. It has the advantage of shedding light on the eternal fascination that any intellect exhibits toward anything resembling exclusion, on the one hand, without forgetting about its love story with simple life as such, on the other.

One might, then, distinguish between an intelligence function that handles the exclusion problem in general, and an intelligence service that is specifically engaged in selecting excluded issues to be included. An intelligence service chooses specific conditions of how to handle the intelligence function. The intelligence function works on the problem of exclusion, whereas the intelligence service works on the problem of inclusion. They are like two sides of the same coin, nevertheless, since there is no working on exclusion without including it, and no working on inclusion without considering it as exclusion in the first place. The difference between the two sides lies in the selection specificity of the service, a specificity that may itself be handled by the function which becomes, thereby, a service.

2. Distinction

The problem of intelligence resides in the problem of how to include the excluded, and how to watch out for the effects of exclusion propelled by the inclusion of the excluded. It is comparatively easy to describe an intelligence service as long as the selection specificity is answered for on a higher level. And it is comparatively easy to see what happens when this higher level, for its answering for selection specificity, relies on the intelligence service itself, trapped by its very name, if you like. Just take a look at the MfS of the GDR to get an impression of the tangled short-circuiting of selection involved. For any intelligence an intelligence service provides for is an intelligence only as judged by others. You can't rely on the service itself to judge the function.

In what follows I propose a concept that might be able to handle the more difficult problem of an intelligence service and an intelligence function that arises when they are operating on the same level, introducing level distinctions which are themselves taken care of by the service and the function. The question, then, will be <u>how to reenter the distinction between service and function into the service, or, more properly, into the system that makes use of both service and function</u>. You may read one of the late John Le Carré's spy novels which describe intelligence services on the brink of readopting an intelligence function, if you want to get an idea of the difficulties involved. Or take a look at the Israeli Mossad contacting the PLO on behalf of peace talks–yet relying for those contacts on governmental order. Our concern here will be how to view and model a distributed intelligence. As it happens, that is an idea that is pursued by AI as well.

The mechanism proposed hereafter relies on the <u>mechanism of re-entry</u> developed by G. Spencer Brown (1977) in his calculus of indications, combined with Heinz von Foerster's (1981, pp. 304-306) <u>concept of double closure</u>.

We take intelligence to be a function and a service inside an observing system that reproduces itself autopoietically by having recourse to its own operations of cognition which are distinct from an environment. This system may be a mental system or a social system, that is, operating by means of consciousness, or communication (see Luhmann 1990). Until further notice, we will not assume this mechanism of intelligence to operate in anorganic or organic systems as well. Social and mental systems rely on <u>double closure</u> insofar as they do not only reproduce operations by means of operations of the same type – that is, thinking is only reproduced by thinking, communication only by communication – but moreover organize the reproduction of their operations by means of the organization of the reproduction of their operations. Double closure demands nothing less than the <u>possibility to use operands as operators, and operators as operands</u>. Heterarchy (McCulloch 1989, pp. 40-45), or strange loops in tangled hierarchy (Hofstadter 1979), that is, circularity, appears to be the logical minimum to be required for the implementation of such a switching from operands taken as operators, to operators taken as operands, and vice versa.

But the main question seems to be how to understand those operators, and operands. How are we to model operations that are circular in their operation in that they take themselves as operands as well as operators? It is precisely this question that is answered by Spencer Brown's <u>calculus of indications</u>. This calculus takes any operation to be an indication that relies on a distinction that distinguishes the indicated side of the distinction from everything else. Any distinction thus has two sides, a marked state and an unmarked state, and both sides taken together constitute <u>the form of the distinction</u>. There are only three ways to handle a distinction: (a) one may call it again, and thus confirm the distinction; (b) one may cross from the marked state of the distinction to the unmarked state, und thus cancel the distinction; (c) and one may re-enter the distinction into the indicated state, and thus be able to observe (i.e., distinguish) the form of the distinction, its two-sidedness. The first two possibilities are called the "law of calling", and the "law of crossing", respectively (Spencer Brown 1977, pp. 1s.). They are the only "laws of form" which, based on the definition of the distinction as "perfect continence" (p. 1), or "a form of closure" (p. 77), enable Spencer Brown to get his calculus started.

Revealingly, the re-entry is not a law at all but the rather improbable event of the distinction of the distinction as such. Spencer Brown has to upgrade his calculus to equations of second degree, that is, to infinite expressions, in order to be able to show how an re-entry may come about, taking, as it were, a part of the expression for the whole expression (p. 58). But it is this improbable event that proves a distinction to be a self-referential operation–an operation that takes itself as the precondition of itself (see Kauffman 1987 and Luhmann 1993). Yet, it is a law that as soon as you reenter the distinction into the marked state you observe its form, and, thus, feed yourself back to the instance when you drew the distinction in the first place, now seeing what you did not see then: (a) that you drew the distinction; (b) that you distinguished yourself by drawing this distinction; (c) and that you did not see either the distinction or yourself, the drawer of the distinction (see Glanville 1979).

"We see now that the first distinction, the mark, and the observer are not only interchangeable, but, in the form, identical" (Spencer Brown 1977, p. 76). In reentering the distinction into the marked state you reveal the distinction to be the blind spot of its own operation, and you start to observe this blind spot (that is, the distinction itself and yourself, the observer) by using a further distinction that you now know to have a blind spot itself even if you see that only by using a third distinction, etc. "Thus the world, when ever it appears as a physical universe, must always seem to us, its representatives, to be playing a kind of hide-and-seek with itself. What is revealed will be concealed, but what is concealed will again be revealed" (p. 106).

In upgrading his calculus from equations of first degree to equations of second degree, Spencer Brown made only explicit what implicitly (see the notion of "unwritten cross", p. 7) enabled him to start the calculus in the first place. But note that the distinction, as it is used on the first level, is not exactly the same distinction that is reentered into the distinguished space on the second level. On the first level the distinction works as a "cross", as in "Draw a distinction." On this level, it is the injunction that is, if obeyed, the distinction itself. Referring to the wording invented by J. L. Austin (1962), we may call the distinction on the first level a performative utterance. On the second level, the distinction is not a "cross" but a "marker": it is not to be taken "literal", for it does not draw the boundary, but represents it. You observe a "cross", but a "marker" you just watch – perhaps as if you listened to a constative utterance.

Now, if we take intelligence to be the handling of the problem of inclusion, i.e., the choosing of an appropriate selection that knows its dynamics of exclusion, then we may rely on the notion of the form of distinction in order to be able to show how it is possible to observe both aspects of any operation, the aspect of inclusion and the aspect of exclusion. Intelligence relies on the observation of the form of distinction, that is, on the observation of the identical of the different: It is one distinction, but it has two sides (see Luhmann 1993). Any intelligence is thus a calculation of forms of distinction that takes account of marked states as well as of unmarked states–the distinction itself, as long as it is used, always being part of the unmarked state itself (the blind spot).

The enigma of intelligence, as viewed from this angle, consists in the fact that it not only computes marked states, inferring one state of knowledge from another state of knowledge, but also takes unmarked states into account. Intelligence consists in inferring states of knowledge from states of ignorance. Peirce, perhaps, would have spoken of "abduction".

One of the few instances of a calculus that makes use of unmarked states, apart from Spencer Brown's, is Gotthard Günther's kenogrammatic. He introduces his notion of a "proemial relationship" in order to show how a distinction may become something that is distinguished, and how something that is distinguished may become a distinction (Günther 1979, p. 226), that is, how an operand turns out to be working as an operator, and vice versa. This proemial relationship, that is a relationship between a symmetrical exchange relation and a nonsymmetrical ordered relation, shows how any operation of intelligence, taking recourse to unmarked states, or to "the other side" of a distinction, might work, namely by combining two distinctions, one in a position proemial but orthogonal to the other, and deliberately exchanging one side of a distinction for another, or reversing and re-reversing the order of the distinctions. With respect to rhetoric and psychoanalysis, both accustomed with the intricacies of intelligence, one might speak of metonymical operations in the case of symmetrical exchange, and metaphorical operations in the case of the nonsymmetrical ordering (see Burke 1969, pp. 503 ss.; Lacan 1966, pp. 273 ss.).

3. Distribution

There is not one intelligent operation. It takes at least two distinctions to observe each other. And it takes an infinite number of distinctions to reenter one of them into the realm of the distinguished in order to be able to observe its form. That puts a limit to any attempt at implementing intelligence on a local and isolated level, as, for instance, in a machine. Instead, intelligence relies on distribution and recursion.

We may even go a step further and maintain that intelligence, coming down, as it were, to computations of unmarked states, is the consideration as such of the environment of a system. Now this sounds more trivial than it is, taking into account that we are dealing with operationally closed systems. We have a couple of fairly good descriptions of the self-referential operations of these systems, reproducing themselves inside a medium of perturbations by connecting self-referential operation to self-referential operation. We know how to describe their evolutionary drift by means of variation and selective retention (see Campbell 1969). But we do not know very much about their observational capabilities. We take any observation to be a distinction that distinguishes a marked state from an unmarked state. But that leads only to a description of system formation, because we can see now how the connection of marked states (an arrangement of distinctions, in the wording of Spencer Brown, p. 4) makes up for a system.

But how do we account for intelligence, that is for the inclusion of the unmarked states as well? The endeavor is paradoxical, but that should not stop us. I assume the basic intelligent operation (that is, remember, never just one) to be an inference of a marked state from an unmarked state, or a position of knowledge from a position of ignorance. This takes a system, mental or social, that is itself but a bunch of bits of knowledge, to displace itself into a position of ignorance (remember the philosopher who demanded to know your ignorance?), and to assume its environment, the unmarked state, to be in a position of knowledge. And that is but a different wording for the switching from self-reference to other-reference. It takes intelligence to do that. It takes a system to prefer knowledge over ignorance. And it takes intelligence to reverse that preference, to know the arbitrariness of this preference, and to assume itself to be in a position of ignorance. Of course, that intelligence is and remains the intelligence of a system, the reversal of the preference confirms the preference, a sudden flash of a symmetry between system and environment confirms the asymmetry.

For AI, this statement of a reentering of symmetry into the asymmetry between system and environment is nothing new. It assumes its expert systems to be "part of an intelligent environment" (see Schwartz 1989, p. 181), which is clearly a paradoxical notion because only systems can be intelligent, but which reveals the main mechanism of artificial intelligence which consists in smuggling the admission of ignorance into social systems that thus lend their intelligence to the expert systems boasting to be intelligent. It is like Gorbatchev declaring perestroika to be the solution to the problem of socialism, leading socialism to accept itself as the problem. AI offers solutions to problems, these solutions being, ultimately, the systems themselves.

Any introduction of intelligence into a system relies on the recursive operations of the system itself. Rather than an introduction, it is a production of intelligence by the system attributed to the environment of the system. And that means, it is a reentering of the distinction that constitutes the system into the system which is now able to regard the form of its distinction, and to observe the exclusiveness of its inclusiveness.

The research work this article is based upon is sponsored by the Deutsche Forschungsgemeinschaft. This paper is a shorter version of my article "Über Verteilung und Funktion der Intelligenz im System" in: Werner Rammert (Hrsg.), Soziologie und Künstliche Intelligenz: Produkte und Probleme einer Hochtechnologie, Frankfurt am Main: Campus, 1995, S. 161-186 (wiederabgedruckt in Dirk Baecker, Wozu Systeme? Berlin: Kulturverlag Kadmos, 2002, S. 41-66).

Bibliography

Ashby, W. Ross (1961): Mechanisms of Intelligence: Ross Ashby's Writings on Cybernetics. Seaside, Cal.: Intersystems 1981.

Austin, J. L. (1962): How to Do Things with Words. London: Oxford UP.

Barel, Yves (1989): Le paradoxe et le système, essai sur le fantastique social. Nouvelle édition augmentée. Grenoble: PUG.

Burke, Kenneth (1969): A Grammar of Motives. Reprint. Berkeley: Univ. of California Pr.

Campbell, Donald T. (1969): Variation and Selective Retention in Socio-Cultural Evolution. In: General Systems 14, pp. 69-85.

Foerster, Heinz von (1981): Observing Systems. Seaside, Cal.: Intersystems 1981.

Gadamer, Hans-Georg (1967): Philosophische Bemerkungen zum Problem der Intelligenz. In: Kleine Schriften I: Philosophie Hermeneutik. Tübingen: Mohr, pp. 220-230.

Glanville, Ranulph (1979): Beyond the Boundaries. In: R.F. Ericson, ed., Improving the Human Condition: Quality and Stability in Social Systems, Washington, D.C., pp. 70-74.

Günther, Gotthard (1979): Cognition and Volition: A Contribution to a Cybernetic Theory of Subjectivity. In: Beiträge zur Grundlegung einer operationsfähigen Dialektik. Vol. 2. Hamburg: Meiner, pp. 203-240.

Hofstadter, Douglas R. (1979): Gödel, Escher, Bach: An Eternal Golden Braid, New York: Basic Books.

Kauffman, Louis H. (1987): Self-Reference and Recursive Forms. In: Journal of Social and Biological Structure 10, pp. 53-72.

Lacan, Jacques (1966): Écrits. Paris: Seuil.

Löfgren, Lars (1979): Unfoldment of Self-Reference in Logic and Computer Science. In: F. V. Jensen, B. H. Mayoh, K. K. Møller, eds.: Proceedings from the 5th Scandinavian Logic Symposium. Aalborg, pp. 205-229.

Luhmann, Niklas (1984): Soziale Systeme: Grundriß einer allgemeinen Theorie. Frankfurt am Main: Suhrkamp.

Luhmann, Niklas (1990): Essays on Self-Reference. New York: Columbia UP.

Luhmann, Niklas (1993): Die Paradoxie der Form. In: Dirk Baecker, ed.: Kalkül der Form. Frankfurt am Main: Suhrkamp, pp. 197-212.

Maturana, Humberto R., and Francisco J. Varela (1980): Autopoiesis and Cognition: The Realization of the Living. Dordrecht: Reidel.

McCulloch, Warren S. (1989): Embodiments of Mind. Second printing. Cambridge, Mass.: MIT Pr.

Newell, Allen, and Herbert A. Simon (1976): Computer Science as Empirical Inquiry: Symbols and Search. In: Communications of the ACM 19, pp. 113-126.

Pask, Gordon S. (1970): The Cybernetics of Behaviour and Cognition Extending the Meaning of "Goal". In: Cybernetica 13, pp. 139-159 and 240-250.

Piaget, Jean (1947): La psychologie de l'intelligence. Paris: Armand Colin.

Schwartz, Ronald David (1989): Artificial Intelligence as a Sociological Phenomenon. In: Canadian Journal of Sociology 14, pp. 179-202.

Sharkey, Noel E., and Rolf Pfeifer (1984): Uncomfortable Bedfellows: Cognitive Pychology and AI. In: Masoud Yazdani, Ajit Narayanan, eds.: Artificial Intelligence: Human Effects. Chichester: Ellis Horwood, pp. 163-172.

Spencer Brown, G. (1977): Laws of Form. Second american edition. New York: Julian.

Weston, Paul E., and Heinz von Foerster (1973): Artificial Intelligence and Machines that Understand. In: Annual Review of Physical Chemistry 24, pp. 353-378.

Varela, Francisco J. (1984): Two Principles for Self-Organization. In: Hans Ulrich and Gilbert J.B. Probst, eds.: Self-Organization and Management in Social Systems: Insights, Promises, Doubts, and Questions. Berlin: Springer, pp. 25-32.

Winograd, Terry, and Fernando Flores (1986): Understanding Computers and Cognition: A New Foundation of Design. Norwood, NJ 1986.