Certifying Life

P.A. Hancock

University of Minnesota

Preamble

Systems have three possible states: stable, transient, and failed. When a system is stable no certification is necessary. If a system is in transition no certification is possible. If a system has failed no certification is relevant. I argue, therefore, that certification is a palliative and an anodyne to societal concerns over the potential destruction that advanced systems can wreak. I further submit that the manifest 'need' for certification is part of an occidental view that nature must be tamed, constrained, and contolled. It is unlikely that our cultural myopia will be excised by the present polemic. But, within the fullness of time, mutual co-evolution and validation by nature itself will fulfill the argument for me.

Introduction

When my father died, I was some 40,000 ft. above Iceland. I will never be able to reconcile myself to the fact that I could not see him one more time before he died. So when I saw him in Cheltenham hospital's 'Chapel of Repose' much of what I felt was anger and frustration, diffusely directed.

I realize now, some years later, that part of that frustration had to do with life itself. As I stood in front of his body I could not help but feel that he was only asleep. After all, he had not changed substantively since I had last seen him. But physical appearance belied what we all know and all must eventually face ourselves; what made my father *my father* had gone.

As I started to write this diatribe on certification, I realized that some doctor had been asked to certify that my father was dead. Indeed, it became obvious that many agencies required this certified evidence to remove him from the 'lists of the living.' As you might imagine, and I hope you do not experience, the bureaucracy of dying is as obscene as the event is disturbing. (I exonerate the process only on the grounds that the individuals involved proved both sensitive and caring in a job where repetition must eventually dull the sensibilities.) In the Cheltenham office of the Department of Health and Social Services, I pondered on the comparison between my father's death and the demise of any system in society, biological or technical.

Human Factors Certification of Advanced Aviation Technologies
Edited by J. A. Wise, V. D. Hopkin, and D. J. Garland

Certifying Failure

My immediate response was, why bother? Nothing in the process of certification was going to bring him back and therefore, for me as an individual, the process was redundant. However, while my father was important to me as an individual, he was, as an entity, important to society in another manner. It was in this latter sense that he was being liquidated. As a result, one critical question on a societal *and* individual level, for both humans and machine systems, was why they failed. Whether it's a post-mortem or an accident investigation the process is the same; a post hoc analysis of what went wrong.

The fundamental assumption is that knowledge of what went wrong last time will help us to avoid the 'same' sequence of events leading to failure next time. With respect to my father such reasoning is vacuous. There can be no 'next' time. With respect to a theory of technical systems operation such reasoning is also becoming progressively more naive. That is, events of failure are becoming more idiosyncratic and less deterministic. Despite our continual attempts to extract or even impose pattern on failures, we are faced with the certainty that as nature explores the combinatorial explosion of interactive states that complex systems can achieve, no two failures will be exactly the same and are liable to become progressively divorced in type. Hence our search for pattern will devolve to ever higher meta-levels of description until we provide the single parameter, unified field theory of failure, namely; "It Broke!" The alternative to this is articulated in a recent chapter (Hancock & Chignell, 1993) and is directly dependent upon the resolution of an empirical proposition concerning the demise of natural ecological systems. This implies there is a power law relationship between the frequency of failure occurrence and the magnitude of that occurence. Should the power law, founded in the application of non-linear dynamics, apply to technical systems, it would provide strong evidence that technical systems are as "natural" as ecosystems and are subject to the same constraints. In particular, it would imply that catastrophic failures of systems are the magnification of events that propagate through systems and predominantly result in minor, frequently unnoticed, perturbations. Where and how intentionality might supress or deflect such propagations remains a moot issue until a veridical power-law relationship has been demonstrated. Had my father been murdered, as a technical system might be sabotaged, we use such knowledge to apportion blame. However, that is an outflow of certification of failure, a prime reason for it.

We *have* to believe in regularity, since we humans at a fundamental level *invent* it. Therefore, we have to, in the occidental world at least, subscribe to the notion that the future is, at least partly, predictable from the past and therefore controllable. To subscribe to a radically differing version of this belief is to risk being labeled, almost literally, 'insane.' Indeed, as Schrodinger (1944) observed about the self-consistency of these rules:

> It is well-nigh unthinkable that the laws and regularities thus discovered should happen to apply immediately to the behavior of systems which do not exhibit the structure on which those laws and regularities are based.

However, at the heart of the schizophrenia of this position is our manifest dissonance between a view of time as a linear dimension in which unique progression obviates exact repeatability versus time as cyclic phenomenon in which repetition and recurrence dominate (Toulmin & Goodfield, 1965). Our present zeitgeist is to believe that the future must be like the past in some way, but cannot be the past exactly. And the 'laws' that have structured the past should have a consistent influence in the future. This latter assumption is a belief[1], not an empirically supportable statement, as is the predicate of regularity and consistency in a more subtle way.

I also felt reasonably confident in asserting that the post-mortem cause of my father's death was about as accurate as the most cursory of accident investigations. In reality doctors deal with death in much the same way they deal with illness. They do not have the time for exhaustive diagnosis of particular problems, hence they frequently treat symptoms or provide palliative agents of widespread capability that will cover the source of the problem without ever necessarily identifying exactly what the problem is. Cause of death is even less liable to receive in detailed examination since the problem rarely proliferates. If death is likely or if there is some specific reason (e.g., homicide), they do what failure analysts do anyway: they pass the problem on to a specialist (e.g., a forensic pathologist). What is frequently not acknowledged is that because we do not fully understand the phenomenon of life, we cannot always specify why life is extinguished. With respect to complex systems, as they grow less determinate in their actions (indeed as many such systems already are), their 'cause of death' may become equally difficult to specify. Right now many professional medical personnel will acknowledge that some individuals die because 'they no longer wanted to go on living.' Can we expect an analog of this in our machine systems?

In summary, certifying death is fundamentally irrelevant. At a bureaucratic level, there are many boxes to be checked and some superficial reasons why we need a paper that records demise. But as with the munchkin doctor in 'The Wizard of Oz,' repeatedly asserting the absence of life is hardly an answer to the future of life. (I do not comment on the social function of leave-taking and grieving, but note that such processes occur when we let go of our possessions as well as our loved ones.) Certifying failure states in complex systems is similarly redundant. Post-mortems identify a concatenation of circumstances which connote progressively longer chains of interactive failures, where *a priori* prediction of such failures has not, and some would suggest, cannot be anticipated. The search for pattern in such failures will inevitably turn up some commonalties, since humans can turn up commonalties in the most diverse array of electro-magnetism. However, prevention based on post-mortem is inevitably a losing battle.

Certifying Stability

If certifying failure should prove irrelevant, shouldn't we certify systems for stable states of performance? That is, shouldn't we be able to assure ourselves that withinside the design

---

[1] The foundation of these beliefs has been most eruditely articulated by Sheldon Glashow in the New York Times (October 22, 1989) which stated: "We believe the world is knowable, that there are simple rules governing the behavior of matter and the evolution of the universe. We affirm that there are eternal, objective, extrahistorical, socially neutral, external and universal truths and that the assemblage of these truths is what we call physical science. Natural laws can be discovered that are universal, invoriable, inviolate, genderless, and verifiable. They may be found by men or by women or by miced collaborations of any obscene proporations. Any intelligent alien anywhere would have come upon the same logical system as we have to explain the structure of protons and the nature of supernovae. This statement I cannot prove, this statement I cannot justify. This is my faith."

operational envelope, the system reliably does all that we say it should? In part this depends upon what we mean by complex systems. Let us consider the nature of machines and consider indeterminancy in machines in the same manner we consider the potential for intelligence for machines. Some four decades ago, Scriven (1953) could be fairly unequivocal. He asserted that:

> "Machines are definite: anything which was indefinite or infinite we should not count as a machine."

Today we cannot be as certain. As a result, Scriven's (1953) subsequent argument about the incompleteness of Godel's theorem is not without problem. However, the process of reasoning is instructive.

> Godel's theorem must apply to cybernetical machines, because it is of the essence of being a machine, that it should be a concrete instantiation of a formal system. It follows that given any machine which is consistent and capable of doing simple arithmetic, there is a formula unprovable-in-the-system – but which we can see to be true. It follows that no machine can be a complete or adequate model of the mind, that minds are essentially different from machines.
>
> We understand by a cybernetical machine an apparatus which performs a set of operations according to a definite set of rules. Normally what it is to do in each eventuality; and we feed in the initial "information" on which the machine is to perform its calculations. When we consider the possibility that the mind might be a cybernetical mechanism we have such a model in view; we suppose that the brain is composed of complicated neural circuits, and that information fed in by senses is "processed" and acted upon or stored for future use if it is such a mechanism, then given the way in which it is programed – the way in which it is "wired up" – and the information which has been fed into it, the response – the "output" – is determmined, and could, granted sufficient time, be calculated. Our idea of a machine is just this, that its behavior is completely determined by the way it is made and the incoming "stimuli": there is no possibility of its acting on its own; given a certain form of consturction and a certain input of informaiton, then it must act in a certain specific way.

In arguing the mind cannot be like a machine, Scriven is limited in a number of ways. First, there is no rationale for suggesting that a mind can explore all possible states of a statement space. That is, as we cannot know everything it may well be the things we don't know that contain anomalies intrinsic to Godel's contention. Second, the argument about seeing what is true, but is improvable in the system can rapidly become a teutology in which we ask *how* the seeing or realization is done. Thus the theoretical difference between mind and machine may be obviated by practical exigency. For the purpose of the present argument, we cannot then state all possible conditions within an operational envelope with certainty. What certification devolves to in this case is an assessment of probability. As a consequence, the heart of certification would seem to represent a customer warranty. For small individual objects, this interaction may be appropriate, since the vendor and the customer are divorced in some spatio-temporal fashion. However, the complex systems about which the present discourse revolves are not the creation of one individual nor are they bought by one individual. In essence, society is at one and the same time, both vendor and client. Even within this global perspective, it is frequently the agency that operates a system that regulates and certifies a system.

We would like to think that if all individual parts of a system were certified then the overall system would be 'safe.' This is bottom-up, wishful thinking. It is the *sine qua non* of design, that objects and systems are created for stability of action and hence should be 'certifiable' with the design space. Yet, here it is the combinatorial explosion of potential interactions, as much as nature's own test and evaluation of those interactions which defeats the hoped for assertion. I should note here that combinatorial explosion of interaction alone does not connote instability as represented by the transient states of operation. This is examined below.

More critically, what are we designing such systems for? It is the frequent observation of the more experienced members of the design community that you never get the opportunity to create complex systems from the ground-up. Almost always they are evolutionary in that new elements are added to older system until the working environment is a palimpsest of overlaid versions. If this is the case, we will never be able to completely specify the parameters of a system that is itself 'underspecified.' More to the point, as we build systems that are beginning to cost in the billions of dollars (e.g., national airspace system, Intelligent Vehicle Highway Systems [IVHS]), we will want them to deal not only with existing conditions but also with future anticipated demands.

Hence, future complex systems must be generative and creative in exploring potential operational 'spaces' in order to be cost effective (Hancock, 1993). In consequence, such systems perforce will be underspecified, for not to do so would be to defeat their evolutionary purpose. Systems that are intentionally underspecified cannot be certified for all phases of operation. Thus we arrive at an impasse. That is, the very systems that we seek to certify should, by design, defy certification.

However one seeks to justify certification for stable states of system operation, one will devolve to this paradox. The paradox is that certification is a guarantee of future operation and implies a predictive determinism about that future state. If such deterministic foreknowledge could be achieved, the operation involved would be completely automatic and by definition not a complex system of the sort relevant here. However, as the future is conceived of as either partly deterministic or totally indeterminant, we want systems to adapt to unforeseen conditions and to explore 'strange new worlds' in order to justify their cost. Under neither circumstance is certification necessary or indeed feasible.

In his argument concerning the potentiality of machine intelligence, Turing (1950) examined the same issued from an inverted position and countered the argument that machines could not be intelligent because of the informality (or indeterminacy) of behavior. He indicated that:

> *It is not possible to produce a set of rules purporting to describe what a man should do in every set of circumstances. One might for instance have a rule that one is to stop when one sees a red traffic light, and to go if one sees a green one, but what if by some fault both appear together? One may perhaps decide that it is safest to stop. But some further difficulty may well arise from this decision later. To attempt to provide rules of conduct to cover every eventuality, even those arising from traffic lights, appears to be impossible.*

Given both the paradox of certification and the improbability of comprehensive future prediction, certification around stability appears a vacuous endeavor indeed.

Certifying Transition

If we do not need to or cannot certify failure and are excused from certifying stability, surely we have to explore certification in the intervening realm where systems fluctuate between stability and failure – the regions of transition. This appears most relevant, since it is during the process of incipient failure and recovery from potential failure that represents the most critical active phase of operation. The problem again is one of predictability.

Certification is an assurance of determined causality. That is, we undertake to state that if a sequence of conditions prevail and a sequence of processes are in operation, a series of outcomes are guaranteed. However, when we step into transitional states, we enter regions that by definition provide increasing uncertainty.

I noted above that the societal investment in large-scale complex technical systems implies that they should be generative and explorative. I shall extend this description to imply that such systems should also be 'skillful.' I use skillful in a specific context here. The context is one that has been used in examining adaptive systems (Holland, 1991). It has been posited that adaptive systems are so structured in response to their initially experienced environmental contingencies. That being that adaptive systems, of which life is the pre-eminent example, grew at 'the edge of chaos.' The latter condition is one where the phase plane of operation devolves from a stable condition toward a chaotic one. (A random regime does not allow sufficient consistency to allow responsive systems to develop, a system in energetic stasis cannot develop adaptive strategies.) It is at the edge of chaos that adaptation develops. 'Skill' in this context is the ability to explore the edge of chaos and the advantages intrinsic to residence in that region without fallback to immaleability or transition into chaos itself.

Systems in transition reside in the region between stability and chaos (not to be directly equated with complete failure). Hence, certification of skillful systems in transition is to suggest that we can 'predict' the response of an adaptive system whose primary function is to cope with unanticipated conditions. The imperilment of such a procedure is now surely laid bare. We cannot certify a system in such conditions, since to do so would be to constrain the very stages of response of a system that we want to be open and unconstrained in order to recover to a state of operational stability.

Certifying What?

I have presented a polemic which has used an analogy with life. Life is a successful adaptive complex system that is predicated upon the environment but is, we believe, not totally constrained by it in terms of its response. Within some bounds we can engage in a certification of life, but why would we? I have suggested a parallel between the failure of a system and death. By extension, the parallel holds for health (stable states) and disease and trauma (transition states), although I have not articulated these latter conditions in as much detail.

I have suggested that certification of stable and transient system states is a relatively futile exercise, since I posit that the very systems we are focusing on are ones which imply open, explorative, and non-deterministic functions. Certification of failure is a time honored societal endeavor to provide information on how to obviate failure in successive systems. In deterministic systems with high frequency of occurrence in the same fundamental state (e.g., DC-10's), this can be a useful function. For one-off large scale systems of progressive indeterminacy such certification serves a more social role in apportioning blame or accountability. I submit that the latter function is a societal palliative for the fears that such indeterminacy brings. I further submit that this is an occidental pre-occupation and one that stems from the notion of controlling and taming nature. As I have previously indicated (Hancock, 1991), the Titanic is the leitmotif of this 'world view.' I take all other aspects of certification to be 'lowest common denominator' insurance.

Hope For The Future

In reviewing the above, it might appear to be a rationale for doing nothing with respect to the design, test, and evaluation of systems and to fatalistically accept the uncertain outcome that nature 'chooses' to provide. I reject this fatalism wholeheartedly. What is objected to is an attitude of mind that proposes that we can 'know' all the states of complex systems we have already created and are creating by the moment.

Therefore:

i) I advocate a great exercise of humility, aspecially with respect to an understanding of the influence and effect of the technology we create.

ii) I advocate a societal change in attitude from the legalistic 'blame' we seemed destined to fix, to a recognition of societal responsibility for the things we collectively build.

iii) I advocate a recognition of the explorative and adaptive nature of ourselves and by extension the manufacturanda we create to extend ourselves.

iv) I advocate the need for the immediate integration of those whose innovative work is enlightening complex adaptive system operation with those who design, test, and evaluate such technical assemblies.

v) Finally, I advocate a strong thrust of research in the area of 'skillful' systems who possess an acknowledged degree of skill in recovering to stability.

In sum, I advocate the replacement of the procedures of certification with the exploration of training 'skill' in complex human-machine systems. I am not foolish enough to believe such recommendations are liable to actually enact change. I adhere more strongly to these statements even following the meeting and the interaction which occured. I also take as a cunard the notion of certification as process, since there is then no fundamental difference between certification and design, test, and evaluation. I take such an arguement to be without meaning.

Acknowledgements

I am happy to acknowledge the support and contribution of so many people who have made possible this and other of my work. First and foremost, I would like to thank the organizers, John Wise, David Hopkin, and Dan Garland (and express my regret that Dan was unable to attend the meeting he had done so much to promote). My thanks extend beyond the immediate organizers to Paul Stager and Marv Smith. They have been kind enough to invite me to a series of interesting and insightful meetings and my benefit and therefore debt is immeasurable. I am happy to acknowledge support from a number of agencies. In particular the Minnesota Department of Transportation(Mn/DOT) and the Center for Transport Studies continues to be a major supporter of our Laboratory. The issues of adaptive complex systems is supported by the Naval Air Warfare Center where Jeff Morrison is our enlightened contract monitor. I am especially grateful for the support of NASA through Sandy Hart who was the first to support my research efforts. Of course, the views expressed here are my own and do not represent the viewpoint of any of the agencies named.

References

Hancock, P.A. (1991). The aims of human factors and their application to issues in automation and air traffic control. In: J.A. Wise and V.D. Hopkin (Eds.). *Automation and Systems Issues in Air Traffic Control*, NATO. New York: Springer

Hancock, P.A. (1993). On the future of hybrid human-machine systems. In: J.A. Wise, V.D. Hopkin., and P.Stager (Eds.).*Verification and validation of complex systems*. Martinus Nijhoff: The Netherlands.

Hancock, P.A., & Chignell, M.H. (1993). On human factors. In: J. Flach, P.A. Hancock, J.K. Caird., and K. Vicente (Eds.). *The ecology of human-machine systems*. New Jersey: Erlbaum.

Holland. J.H. (1991). *Adaptation in natural and artificial systems*. Cambridge: MIT Press (University of Michigan Press, 1975)

Schrodinger, E. (1944). *What is life*? Cambridge: Cambridge University Press.

Scriven, M. (1953). The mechanical concept of mind. *Mind*, 62,

Toulmin, S., & Goodfield, J. (1965). *The discovery of time*. Harper & Row: New York.

Turing, A.M. (1950). Computing machinery and intelligence. *Mind*, 59, 433-460.