Institute of Measurement and Control
Measurement+Control
The Journal of the Institute of Measurement and Control

On the concepts of self-repairing systems

Colin Bell[1*], Richard McWilliam[2], Alan Purvis[2] and Ashutosh Tiwari[1]

1. Cranfield University, Cranfield, Bedfordshire, MK43 0AL, UK
2. Durham University, South Road, Durham, DH1 3LE, UK

*c.a.bell@cranfield.ac.uk (corresponding Author)

Abstract

Systems fail. Period. No matter how much planning and fault analysis is performed it is impossible to create a perfectly reliable machine. The existing approach to improving reliability invariably involves advances in fault prediction and detection to include specific mechanisms to overcome a particular failure or mitigate its effect. Whilst this has gone a long way to increasing the operational life of a machine, the overall complexity of systems has improved sharply and it is becoming more and more difficult to predict and account for all possible failure modes. What is discussed here is a possible shift in approach from specific repair strategies to autonomous self-repair. Rather than focusing on mitigating or reducing the probability of failure, the focus is instead of what can be done to correct for a failure that will invariably occur at some point during operation. By taking this approach, it is not just expected failure that can be designed for – unexpected failure modes are also inherently compensated for, extending the potential life of a system and reducing the need for through-life servicing.

Introduction

It is impossible to discuss the concepts of self-healing and self-repair without having some notion about what their meanings. There are currently no universally-accepted definitions of these terms, but instead intuitive notions about the concepts involved. It is not the purpose of this article to suggest a new taxonomy, but instead to look at what the overall aims are of this emerging field and perhaps reflect on what is achievable now. To make these issues more awkward, there are currently many terms for the similar ideas, and conversely many distinctly different ideas that are referred to by the same name. Furthermore, different fields of research such as electronics or mechanical design can have vastly different interpretations and objectives. A good example of this is modular or physical [1-3] redundancy in electronics – these concepts could perhaps be thought of as inefficient if the same principles are applied to a purely mechanical system that contains more material or elements than are strictly necessary for an optimized design.

(Authors' pre-publication version, 2013)

In layman's terms, perhaps what we are looking in self-repair are systems that are able to *maintain some degree of functionality after a failure has occurred*. This might be a controversial interpretation however as it can be argued that certain self-preservation or preemptive actions such as prognostics, or mitigation through fault tolerance are an intrinsic element of self-healing; and hence we should not focus solely on what happens *after* the event of a failure.

The above definition is similar to the general or biological definition of 'resilience', which is commonly interpreted as the *ability to recover from adversity* [4]. Hence fault-tolerant approaches might better fall under this general umbrella of 'resilience' rather than self-healing.

Fundamentally one crucial distinction is the difference between a reactive or proactive system. In fault-tolerance, where the system is able to absorb a finite number of fault events without explicit repair or reconfiguration, it is assumed that failure can to a certain extent be prevented. For this purpose of this discussion however we will assume that failure can and does occur.

Achieving Self-Repair

To achieve a self-repairing system, it is clear that the system must have an element of self-awareness. Amor-Segan et al., [5] state that the ultimate aim is to develop a system with "the ability to autonomously predict or detect and diagnose failure conditions, confirm any given diagnosis, and perform appropriate corrective intervention(s)". Following this logic, Figure 1 offers a proposed approach that can theoretically be applied to any system. By breaking the process down into a number of finite steps we can better assess the current progress toward achieving an idealized self-repairing system.

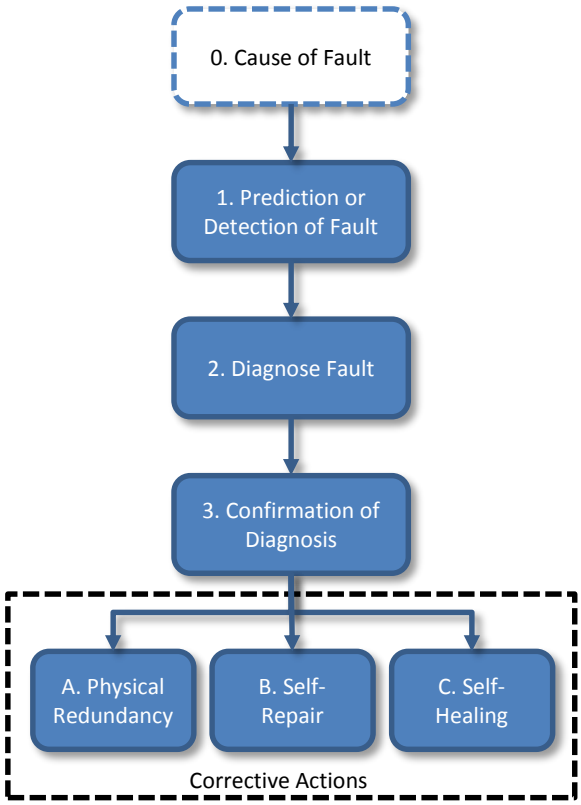

Figure 1. Proposed approach to self-correcting systems

Perhaps the first point that can be drawn from this proposed process is that the underlying cause of fault is not considered crucial. There is a whole research sector dedicated to function-based failure analysis [6-8], and whilst there will invariably be some degree of cross-over between the disciplines, it is better here to focus instead on what happens after failure has occurred.

Detection and Diagnosis

Any critical fault will almost invariably lead to a fundamental change in the behavior of the system. This could perhaps be most easily interpreted as a deviation from the prescribed behavior, utilizing either internal or external telemetric data. One of the difficulties with this in complex system is in defining 'expected behavior', however this problem is not insurmountable and a great deal of progress has been made in this research area [9].

Conversely the diagnosis of a fault is perhaps a more difficult proposition. This is partly due the difficulty in validating large, complex system models because of the vast number of possible system states [5]. Furthermore, there is the issue of confidence in diagnosis, i.e. how much certainty must be present to initiate repair? Because of this, an additional step is proposed in which the diagnosis must be confirmed, to avoid undesirable events such as 'good' components being unnecessarily removed or routed around. Several methods are currently available for this:

- Model-based: Abductive reasoning: compare observation with predicted observation: I expect 'X' but get 'Y', therefore I must correct 'Y' to get it to match.
- Bayesian belief networks: probabilistic graphical model (a type of statistical model) that represents a set of random variables and their conditional dependencies: If 'X' and 'Y' happen, it's likely a failure with 'Z'
- Case-based reasoning methods: anecdotal evidence, if 'X' happens, do 'Y'. – Accounts for expected failure only

Currently, there has been some progress in these areas in electronics with BIST (built in self-testing). Silicon electronic devices are susceptible to a variety of upset events, including transitory events (e.g., random single upset events caused by radiation) and permanent fault conditions that can be triggered by a vast variety of events. Rather than eliminating the underlying cause, BIST has been developed for computer DRAM, where special structures are included in the memory chips that are activated when attached to production test machine. This enables rapid and reliable allocation of redundant memory cells to replace faulty cells which are commonly found in high density memory. Perhaps what we now seek is a shift from this external detection to in-system detection and correction, such as self-contained BIST logic that can operate independently of the (expensive) production test machine during the operational life time of the memory chip. Data error detection and repair is particularly pervasive in electronic systems; it protects critical memory areas such as on chip cache which cannot tolerate transient upset errors.

Corrective Actions

Perhaps of most interest in self-repairing systems is the corrective action itself. If it were possible to fully automate this process then there are huge potential savings in MRO (maintenance, repair, and operations) costs. The precise methodology employed will almost invariably have to be application specific, however a number of possible approaches are available:

- Physical Redundancy: An alternative load path or system is available should the primary fail
    - Currently this is the easiest approach to include and is already implemented on mission-critical systems
    - At a very basic level this can simply be a complete facsimile of the primary system (modular redundancy) that can take over if failure occurs
    - Its relative efficiency can perhaps be measured by how much of the primary system has to be physically replicated to provide the backup

- Self-Repair : The system, as a whole, has the ability to partially or fully fix a given fault to continue operation
    - This is the approach that is perhaps most achievable in the immediate future
    - One approach is to extend the concept of redundancy to the use of degenerate modules that have the ability to perform the same function or yield the same output even if they are structurally different [10].
    - Using this approach rather than having individual back-ups for each module, a single spare module can be reconfigured to provide cover for any defective module
    - Alternatively this concept of self-repair through self-reconfiguration does not necessarily require additional materials, instead performance can be sacrificed to ensure continued functionality utilizing only the currently available resources

- Self-Healing: The system is able to physically bring itself back to its initial state of operation after a fault has occurred
    - True self-healing systems are currently prohibitively expensive and infeasible for all but the most basic of systems or limited to exotic materials
    - An idealized example couple be the ability to automatically re-straighten a mechanical element (through a chemical process) after it has been bent, or physically fix thermal damage in an electronic component
    - An alternative approach would be to have entirely adaptable systems, such as ‘smart dust’ [11], where there is a finer level of granularity and near infinite possibilities for reconfiguration

To better emphasize the distinction between the corrective actions, Table 1 shows a simple example of the repair of a car-tire puncture and how this compares to biological approaches.

Table 1. Biological inspiration for Self-healing

Corrective Process	Car tire mechanism example	General approach	Biological parallel in broken-skin
Redundancy	Run-flat tire – Stiffened tire wall that is able to temporarily carry load in the event pneumatic pressure is lost	If primary load/electronic/signal path fails, an alternative is used instead	Areas of skin continuously worn develop calluses to provide additional protection against skin breakage
Self-Repair	Tire-weld or similar – substance is used within tire to automatically seal puncture	System is repaired using some peripheral materials automatically	A scab is formed over the cut to prevent further damage and enable continued operation
Self-Healing	Low transition temperature rubber tire that is able to automatically melt to seal a puncture	System is healed/repaired at a molecular level with little or no evidence that repair has taken place	Epithelialization collagen synthesis, contraction, and remodeling occur to produce a near-perfect restoration of the skin

One salient point that becomes apparent when looking at biological parallels is that each of the corrective actions does not necessarily have to occur in isolation. Indeed in the broken skin example it is common for self-repair and self-healing to occur sequentially to produce a coupled-whole system. Indeed this is normally preceded by assisted repair in which the wound is externally bandaged.

Current Progress

The electronics domain is perhaps leading the way with regards to self-repairing systems. An evolution from external testing to self-contained testing is already underway with the next proposed approach BISR (built-in self-repair). EDC (Electronic Data Capture) methods offer BISR functionality via special hardware structures. A limitation here is that permanent faults cannot generally be handled by EDC, and system failure will result. Permanent faults can be protected by introducing system redundancy such as TMR (Triple Modular Redundancy), which was first proposed over 50 years ago [12], however one must assume that the voting logic itself is trustworthy or else can also be replicated.

A less popular approach is that of fine grained fault tolerance employing interconnect interleaving and quadded logic [13], which requires additional logic and signal routing hardware but which is able to 'absorb' certain permanent fault events without loss of functionality. The basic principles at work here are the fine granularity of the underlying transistor and interconnecting structure which offers many possibilities for reconfiguration and fault tolerance. Beyond this, there is significant interest in new bio-inspired approaches that use cellular based architectures. Inspired by the early observations of Von Neumann on the intrinsic fault tolerant properties of biological systems [4], this offers the possibility of electronics systems whose operation is governed by localized interactions between electronic 'cells' i.e., circuits not requiring global coordination, and hence BIST and BISR can be executed at the cellular level [14, 15].

Conclusion

Looking at the overall concept of product reliability, if viewed from the perspective of the user, a system with an integral resilience-mechanism would appear to be more 'reliable' – it is able to maintain operation for a longer period of time than would otherwise have been possible. However from a design approach, systems with additional procedures built-in are invariably more complex and hence the primary system becomes intrinsically less 'reliable', even though it is able to bring itself back to a normal operating condition. Getting the balance right between this intrinsic reliability and apparent reliability is important to ensure self-healing technologies are accepted by the end user.

Despite vast improvements in system modeling and prediction, most machines still fail in the face of unexpected damage [16] and one of the long-standing challenges of creating a reliable system is achieving robust performance under uncertainty [17]. Self-repairing techniques inherently must be designed to compensate for a wide-variety of failure modes, thus overcoming some of the problems associated with uncertainty. Although specific solutions have not been suggested, proposed methodologies for developing self-repairing strategies should not focus on a finite number of underlying causes. Instead the focus should be on how these causes manifest, how they can be detected and ultimately how they can be corrected autonomously.

References

[1] S. Habinc. "Functional triple modular redundancy (FTMR) VHDL design methodology for redundancy in combinatorial and sequential logic design and assessment report", http://www.gaisler.com/doc/fpgan 003n01-0-2.pdf, 2002.

[2] J. Davies, T. Steffen, R. Dixon, R. Goodall, A. Zolotas, and J. Pearson. "Modelling of high redundancy actuation utilising multiple moving coil actuators", in International Federation of Automatic Control (IFAC) World Congress, Seoul, Korea, 2008, pp. 3228–3233.

[3] J. Davies, H. Tsunashima, R. Goodall, R. Dixon, and T. Steffen. "Fault Detection in High Redundancy Actuation Using an Interacting Multiple-Model Approach", in IFAC Symp. on Fault Detection, Supervision and Safety of Technical Processes (SafeProcess), Barcelona, Spain, 2009, pp. 1228–1233.

[4] J. Von Neumann. "Probabilistic logics and the synthesis of reliable organisms from unreliable components", Automata studies, vol. 34, 1956. pp. 43–98.

[5] M.L. Amor-Segan, R. McMurran, G. Dhadyalla, & R.P. Jones. "Towards the Self-Healing Vehicle", Automotive Electronics, 2007 3rd Institution of Engineering and Technology Conference, IET, pp. 1-7, June 2007

[6] I.Y. Tumer, and R.B. Stone. "Mapping Function to Failure During High-Risk Component Development," Research in Engineering Design, 14(1): 25-33. 2003

[7] S.G. Arunajadai, R.B. Stone and I.Y. Tumer. "A Framework For Creating a Function-Based Design Tool for Failure Mode Identification", Proceedings of the 2002 ASME Design Engineering Technical Conference, Design Theory and Methodology Conference, Montreal, Canada. 2002

[8] I.Y. Tumer, R. Stone, R.A. Roberts, and A.F. Brown. "A Function-Based Exploration of JPL's Problem/Failure Reporting Database," Proceedings of the 2003 ASME International Mechanical Engineering Congress and Expo, IMECE2003-42769, Washington, D.C. 2003

[9] M.L. Visinsky, J.R. Cavallaro, & I.D. Walker. Robotic fault detection and fault tolerance: A survey. Reliability Engineering & System Safety, 46(2), 1994. 139-158.
[10] G.M. Edelman, & J.A. Gally. Degeneracy and complexity in biological systems. Proceedings of the National Academy of Sciences, 98(24), 2001. 13763-13768.
[11] J.M. Kahn, R.H. Katz, & K.S. Pister. "Next century challenges: mobile networking for 'Smart Dust'". In Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking. August 1999. pp. 271-278.
[12] R.E. Lyons, & W. Vanderkulk. "The use of triple-modular redundancy to improve computer reliability". IBM Journal of Research and Development, 6(2), 1952. pp200-209.
[13] P. A. Jensen, "Quadded NOR Logic", IEEE Transactions on Reliability, vol. R-12, no. 3, September 1963. Pp.22-31
[14] J. David, R. McWilliam, A. Purvis, "Designing convergent cellular automata", Biosystems, vol. 96, no. 1, 2008. pp. 80–85.
[15] A. M. Tyrrell and A. J. Greensted, "Evolving dependability," J. Emerg. Technol. Comput. Syst., vol. 3, no. 2, July 2007.
[16] J. Bongard, V. Zykov, & H. Lipson. "Resilient machines through continuous self-modeling (sic.)" Science, 314(5802), pp.1118-1121, 2006
[17] S. Thrun, W. Burgard, and D. Fox, "Probabilistic Robotics". MIT Press, Cambridge, MA, USA. 2005.