The Architecture of Robust, Evolvable Networks

Organization, Layering, Protocols, Optimization, and Control

by John C. Doyle

Biological systems are robust and evolvable in the face of even large changes in environment and system components, yet can simultaneously be extremely fragile to small perturbations. Such universally robust yet fragile (RYF) complexity is found wherever we look. The amazing evolution of microbes into humans (robustness of lineages on long timescales) is punctuated by mass extinctions (extreme fragility). Diabetes, obesity, cancer, and autoimmune diseases are side-effects of biological control and compensatory mechanisms so robust as to normally go unnoticed. RYF complexity is not confined to biology. The complexity of technology is exploding around us, but in ways that remain largely hidden. Modern institutions and technologies facilitate robustness and accelerate evolution, but enable catastrophes on a scale unimaginable without them (from network and market crashes to war, epidemics, and global warming). Understanding RYF means understanding architecture—the most universal, high-level, persistent elements of organization—and protocols. Protocols define how diverse modules interact, and architecture defines how sets of protocols are organized.

Insights into the architectural and organizational principles of networked systems can be drawn from three converging research themes. (1) With molecular biology's description of components and growing attention to systems biology, the organizational principles of biological networks are becoming increasingly apparent. Biologists are articulating richly detailed explanations of biological complexity, robustness, and evolvability that point to universal principles. (2) Advanced technology's complexity is now approaching biology's. While the components differ, there is striking convergence at the network level of architecture and the role of layering, protocols, and feedback control in structuring complex multiscale modularity. New theories of the Internet and related networking technologies have led to test and deployment of new protocols for high performance networking. (3) A new mathematical framework for the study of complex networks suggests that this apparent network-level evolutionary convergence within/between biology/technology is not accidental, but follows necessarily from the universal system requirements to be efficient, adaptive, evolvable, and robust to perturbations in their environment and component parts.

"robust yet fragile"

Additional Details and Selected References

The Internet is an obvious and familiar example of how a protocol-based architecture facilitates evolution and robustness. The hourglass protocol "stack" has a thin, hidden "waist" of universally shared feedback control (TCP/IP) between the visible upper (application software) and lower (hardware) layers. This allows "plug-and-play" between modules that obey shared protocols; any set of applications that "talks" TCP can run transparently and robustly on any set of hardware that "talks" IP, accelerating the evolution of TCP/IP-based networks. Diverse applications use additional protocols to exchange objects such as web files, images, and email. In the hardware layer, signals are also exchanged via shared protocols and interfaces. Indeed, recent theory proves that with suitable versions of TCP/IP and network provisioning, an otherwise arbitrary (e.g. large and heterogeneous) network has guaranteed global dynamics that robustly optimizes aggregate application utility. (Similar results hold for flocking and oscillator synchronization.) Other theory shows that overall system robustness obeys conservation laws, suggesting RYF behavior can be managed but not completely eliminated.

Within a layer these protocols can be visualized as "bowties" with large fan-ins and -outs of energy, materials, and information via a thin "knot" of universal protocols specifying carriers, building blocks, or interfaces. Other engineered examples of bowties include networks which connect energy sources to consumers via carriers and standard socket-plug interfaces, sellers to buyers via money, fibers into diverse clothing via sewing, and raw materials to assemblies via standardized building blocks for advanced manufacturing. In these and the biologic examples below, the currencies, carriers, intermediates, precursors, plugs, packets, conserved residues, and interfaces in the bowtie "knot" are highly constrained by protocols. Yet this shared universality allows robust adaptation and evolution to extreme diversity at the inputs (energy, sellers, fibers, materials, nutrients, receptors) and outputs (consumers, buyers, clothing, assemblies, products, effectors), as long as they have appropriate (and typically hidden) layers of feedback control.

The robustness and evolvability of the cell exemplify reliance on layers and protocols analogous to those in TCP/IP and other engineered architectures, but with wondrously different hardware components and more extensive and sophisticated use of feedback control. Perhaps the ultimate bowtie involves the universal, central protocols that transcribe and translate DNA to proteins via conserved codes, polymerases, and RNA intermediates. As in IP, feedback control of polymerases yields a functional network with regulated levels of macromolecules. A simple cell in a constant environment needs little more, but real cells add robustness to changing demands with a TCP-like feedback layer of allostery and post-translation modifications acting on faster time scales. These TCP/IP-like layers (the waist of a cell's hourglass) of control protocols constitute much of biochemistry texts.

In the microbial biosphere, genes that "talk" transcription and translation protocols can move by horizontal gene transfer (HGT), thus accelerating evolution in a kind of a "bacterial internet." Furthermore, proteins newly acquired by HGT or through duplication and divergence or domain rearrangements can work effectively because they can exploit additional shared protocols. For example, a bacterial cell's application layer consists of biochemical fluxes with various additional bowtie protocols such as core metabolism (nutrients to biosynthesis and energy supply via conserved precursors and carrier-based group transfers) and signal transduction (receptors to downstream effectors via conserved residue pairs). Thus selection acting at the protocol level could evolve and preserve this variety of shared and ultimately conserved architecture, essentially "evolving evolvability".

All life and advanced technologies rely on protocol-based architectures. The evolvability of microbes and IP-based networks illustrate how dramatic, novel, dynamic changes on all scales of time and space can also be coherent, responsive, functional and adaptive, despite implementations that are largely decentralized and asynchronous. New genes and pathways, laptops and applications, even whole networks, can plug-and-play, as long as they obey protocols. Biologists can even swap gene sequences over the Internet in a kind of synthetic HGT. A related aspect of sophisticated architectures is that actuator signals (e.g. from intermediate fluxes and concentrations in core metabolism to cardiopulmonary rates and renal concentrating effects) have extremely high variability in order to keep other critical signals (e.g. from key metabolic products to blood oxygenation and pH, to core body temperature,) tightly regulated despite fluctuating supplies and demands. Such natural physiological variability can be a source of confusion if homeostasis is misinterpreted as implying that all signals, including actuators, are held nearly constant.

"plug-and-play"

Typical behavior is fine-tuned with this elaborate control and thus appears boringly robust despite large internal and external perturbations and fluctuations. As a result, complexity and fragility are largely hidden, often revealed only by rare catastrophic failures. Since components come and go, control systems that reallocate network resources easily confer robustness to outright failures, whereas violations of protocols by small malicious or even random rewiring can be catastrophic. So programmed cell (or component) "death" is a common strategy to prevent local failures from cascading system-wide. The greatest fragility stemming from a reliance on protocols is that standardized interfaces and building blocks can be easily hijacked. So that which enables HGT, the web and email also aids viruses and other parasites. Large structured rearrangements can be tolerated, while small random or targeted changes that subtly violate protocols can be disastrous. Thus another source of high variability is in the external behavior of RYF systems, where typical behavior is boringly robust but rare events can be catastrophically large.

Chaos, fractals, random graphs, criticality, and related ideas from statistical physics inspire a popular and completely different view of complexity, where behaviors that are typically unpredictable and fragile "emerge" from simple and usually random interconnections among homogeneous components. The attraction is understandable because a variety of features, continually recycled in series of "new sciences," greatly simplify modeling and analysis. For example, random internal rewiring has little effect. Tuning, structure, and component heterogeneity is minimal, as is system functionality, robustness, and environmental uncertainty. Perhaps most appealing is that "modularity" reduces to superficial self-similar patterns in ensemble averages, avoiding the complexity of protocols and architecture. Unfortunately, highly evolved and architecture-based RYF complexity is utterly bewildering when viewed from this perspective, which has led to widespread and remarkably persistent errors and confusion about complex systems of all kinds. There are striking parallels with irreducible complexity and intelligent design, which has even more successfully popularized persistent errors and confusion, though more narrowly targeting biology and evolution.

Particularly persistent errors surround the origin and nature of power laws. The statistical invariance properties of Gaussians and exponentials for low variability phenomena are well-known enough that their presence is little cause for surprise. But for high variability phenomena, power laws have even richer statistical invariants, and could rightly be called "more normal than Normal." Thus their ubiquity in the abundant high variability of architecture-based RYF systems should be no surprise, and certainly no evidence on their own for any specific mechanism beyond the presence of such high variability.

One thing all can agree on is that the search for a "deep simplicity and unity" underlying natural and technological complexity remains a common goal. Fortunately, our growing need for robust, evolvable technological networks means the tools for engineering architectures and protocols are becoming more unified, scalable and accessible, and happen to include the more popular complexity views as extreme special cases within a much richer structured, organized, and dynamical perspective. These developments are bringing much-needed rigor and relevance to the study of complexity generally, including biology, but will not eliminate the need for attention to domain-specific structure and detail. Quite the contrary: both architectures and theories to study them are most successful when they facilitate rather than ignore the inclusion of domain-specific details and expertise. L E E

John Doyle is the John G Braun Professor of Control and Dynamical Systems, Electrical Engineering, and Bioengineering.

Read more at: http://www.cds.caltech.edu/~doyle

References

[1] H. El-Samad, H. Kurata, J.C. Doyle, C.A. Gross, and M. Khammash, "Surviving Heat Shock: Control Strategies for Robustness and Performance," *PNAS* **102**, 8 (2005).

[2] Doyle et al, (2005), "The 'Robust Yet Fragile' Nature of the Internet," *PNAS* **102**, 41, (2005).

[3] M. A. Moritz, M. E. Morais, L. A. Summerell, J. M. Carlson, and J. Doyle, "Wildfires, Complexity, and Highly Optimized Tolerance," *PNAS*, **102**, 50 (2005).

[4] H. El-Samad, A. Papachristodoulou, S. Prajna, J. Doyle, and M. Khammash, "Advanced Methods and Algorithms for Biological Networks Analysis," Proceedings of the IEEE, **94**, 4, 832-853 (2006).

[5] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, "Layering As Optimization Decomposition," Proceedings of the IEEE, **95**, 1 (2007).