Chapter 1:
System level risk

The world's urban population now exceeds the world's rural population. The growth of towns and cities concentrates risk and the interdependencies of risk.


龙元建设

Understanding and coping with the increasing risk of system-level accidents

Herman B. "Dutch" Leonard, Arnold M. Howitt[1] – The world has seen a number of recent events in which major systems came to a standstill, not from one cause alone but from the interaction of a combination of causes. System-level accidents occur when anomalies or errors in different parts of an interconnected system negatively reinforce one another, spiraling up out of control until they eventually drive the system outside of its sustainable boundaries, resulting in system "collapse". Systems with multiple components that are tightly linked to one another are prone to such events. Increasingly, our industrial, commercial and social systems are coming to have the characteristics that predict system-level accidents – in some cases, driven by consistent economic forces that cause tighter interconnections to form within existing systems – and there seems to be a rising frequency of such events. Since we inhabit an increasingly tightly interconnected global collection of such systems, finding ways to reduce and to manage systemic risk is an important priority.

A particularly heavy snow storm in South China in 2008…

For nearly a month in January/February of 2008, an unusual weather pattern in South China produced a series of snow and ice-storms that were distinctly out of the ordinary. The snow, combined with icing of power lines (and associated power outages), snarled traffic from railroads to highways and disrupted factories, schools, government services and households. To be sure, the weather was significantly worse than normal. Normally, if snow falls at all in this region, its appearance is brief and its effects transitory. This time, the snow came and stayed and came again.

…coinciding with a national celebration led to multiple system failures.

Even allowing for the fact that the snow was worse than is normally encountered in this region – this event was widely described as the worst series of snow and ice events of the last 50 years – the impacts of this event seemed out of proportion to the event itself. Everything seemed to go wrong at the same time. The snow arrived on the eve of the spring festival, when millions of Chinese citizens, who have gone to work some distance from their homes, seek to return to their places of origin to celebrate the new year. Half a million people wound up crowded into the Guangzhou rail station, with no trains to take them to their destinations, no food, little water, and what might at best be referred to politely as primitive sanitary conditions. Unwilling to leave, they were also virtually impossible to sustain in situ. Troops and national political figures hurried to the scene to enforce order and plead for calm. Throughout the region, breakdowns that were different in detail but similar in overall pattern were taking place. Power outages were widespread and persistent – partly because icing caused breaks in power lines, but also partly because transport obstacles prevented repair efforts from moving rapidly to affected areas. With transport at a halt, coal supplies ran short at power stations; in the resulting power outages, coal became difficult to transport (on the newly converted electric train grid) until diesel-fuelled engines could be mobilised and assembled in the area (and, meanwhile, diesel fuel also became scarce).[2]

[1] The authors are the faculty Co-Directors of the Program on Crisis Leadership, a joint programme of the Ash Center for Democratic Governance and Innovation and the Taubman Center for State and Local Government at the John F. Kennedy School of Government at Harvard University, and faculty Co-Chairs of the Kennedy School's Leadership in Crises executive programme. We are indebted to Douglas C. Ahlers for inspiration, guidance, and technical knowledge about systems and system dynamics and for suggesting a number of important insights and to Arrietta Chakos, David Giles, and Jason Qian for research support, comments, and suggestions. This research was generously supported by Swiss Re. Any remaining errors, alas, were always and still are our own.

[2] Much of this description is drawn from an extended Program on Crisis Leadership case study by Jason Qian, currently in draft.

Why were the effects of the snowstorm – unusually severe, to be sure, but hardly cataclysmic – so widespread and profound? What turned a modestly novel moderate-sized event into a major, embarrassing, out-of-proportion, "landscape-scale" disaster?

The high interconnectivity of a number of systems…

Simply put, *one thing led to another.* Fundamentally, the challenge of the South China snowstorm was not that the storm itself or its immediate effects were so severe – it was, instead, that the consequences pyramided on one another and reinforced one another, spiraling up to create, collectively, a collapse of the wider *system* of tightly interconnected utilities (power, water, transport, communications), economic activity (manufacturing, power generation, food distribution), social activity (using transport systems to journey home for the holidays) and governmental activity (transporting additional supplies and assistance to affected areas). Under normal conditions, these different parts of the larger system operate in reasonable harmony with one another, and the system exhibits some level of resilience, absorbing small shocks and perturbations without creating major discontinuities in services or activities.[3] By contrast, when the repeated snow and icing conditions simultaneously affected a wide range of interconnected components of the utility and economic system, the resulting consequences fed on one another and pyramided – eventually pushing each other collectively past the capacity of the system to absorb the shifts out of its normal operating ranges. In short, the extended snow, ice, and cold had caused a system-level failure or "collapse" – the failure was no longer just of the individual system elements, but also of the network of components – that is, it was a failure of the system as a whole.

…can trigger sudden and unexpected failures, to which societies are becoming more susceptible…

Events like the failure of interconnected elements of systems leading to wider system failure events – as in the example of the South China snowstorm – seem to be becoming more common, and we believe this is in fact a pronounced and important trend. We believe that there are strong forces in the normal and common ways that economic, financial, natural and man-made physical systems co-evolve in the modern world that make the prevalence of the conditions that lead to system-level failures more prevalent. To put it another way, the risk of system collapse is secularly increasing, and system collapse as a form of large-scale social hazard should be understood as an increasingly likely and profoundly important phenomenon that needs to be addressed through changes in system designs, through policy, and – since it constitutes a risk – arguably through the development of new approaches to defining and insuring the resulting risks. If this is correct, the implications are potentially both deep and wide.

…necessitating greater understanding of such systemic threats.

We will begin by defining more carefully the nature of system failures and the conditions that facilitate them (and, indeed, in many cases make them inevitable). We will then examine some examples of past system collapses and some possible system collapses in the future, using this as a lens through which to identify some of the forces that seem to be systematically increasing the prevalence of this phenomenon.[4] We conclude with some preliminary thoughts on how we might seek to manage this rising form of significant social hazard.

[3] Such systems can be described as having *stability* or as being "Lyapunov stable".

[4] The phenomenon we are describing is what Charles Perrow labelled a "system accident". We will use the terms "system-level failure", "system collapse", and "system accident" more or less interchangeably.

The nature of system accidents

Systemic challenge is the product of subsystem interaction…

Large-scale disasters often involve circumstances where two or more phenomena interact negatively, reinforcing one another's negative consequences. A rainstorm might bring down power lines that halt pumps needed to avoid flooding; the resulting flood may increase damage to the electrical grid, accentuating damage to the pumping process and making it more difficult and more time-consuming to reset the elements of the system so that they are again operating within their normal ranges (and in harmony with one another). Generally speaking, we can define a "system challenge" as a situation in which a series of "subsystems" interact in ways that are harmful to one another, pushing each other away from their normal operating conditions. This can be caused by an external shock to one or more of the systems or by an "error" – a breakdown of one or more components of a subsystem that takes the subsystem outside the range in which it is supposed to operate – in one subsystem or in two or more subsystems at the same time. A "system accident" or "collapse" occurs when a challenge rises to the level of pushing one or more elements of the overall system beyond their design limits, resulting in breakdown.[5]

…which will, eventually, inevitably lead to system collapse.

The phenomenon of system breakdown has been widely studied both in practice and in the abstract, particularly in the field of "system dynamics", which is devoted to understanding the nature of flows and interactions among elements of systems with different design features, structures, control and feedback mechanisms, and degrees of component reliability. A particularly compelling and accessible description of these issues is provided in Charles Perrow's *Normal Accidents.*[6] Perrow outlines the general characteristics of systems that can be shown mathematically to have a non-zero probability of experiencing a system collapse – and which, therefore, if they are operated for long enough, will *inevitably* produce a system accident (hence the use of the term "normal", indicating that [eventual] system collapse is the ordinary state of affairs in such a system). In addition to presenting these results in an abstract and general form, Perrow provides a wide array of examples of systems that exhibit the characteristics that make eventual collapse statistically unavoidable. Perrow's work builds on (his own and others') earlier theoretical research in system dynamics, and is also the source and inspiration of much of the technical work that has followed since.[7]

Systems are designed to work within a specific space…

What are the conditions that create the statistical certainty of system collapse in the long run? To describe these conditions we need first to specify a bit more precisely the ways in which systems are designed to operate. The "state" of a system or subsystem can be characterised by the status of its relevant features at any given moment (for example, in a car engine, its state variables include its revolutions per minute, temperature, power output, fuel consumption rate, lubrication status, and so on – a set of characteristics that are interrelated to one another and which jointly define all that is relevant about the system's operations). Correspondingly, the system's "state space" consists of the set of all possible combinations of the state variables. When the system is operating in some regions of the state space, it will exhibit high performance; when it occupies other regions, its performance will be lower; in still other regions it may cause damage to its components, so that it cannot easily be restored to the high-performance region.

[5] Breakdown or failure can be due to either a low fault-tolerance (sudden failure when a breaking point is reached) or by a degradation of system performance past a lower-bound threshold of acceptable performance.

[6] Charles Perrow, *Normal Accidents: Living with High-Risk Technologies* (Princeton, NJ: Princeton University Press, 1999). The authors refer to Perrow's pioneering work on normal accident theory throughout this paper.

[7] System Dynamics was originally developed by Jay W. Forrester. See: Jay Forrester, *Industrial Dynamics* (Cambridge, MA: MIT Press, 1961); and Jay Forrester, *Urban Dynamics* (Cambridge, MA: MIT Press, 1969).

…and will break down outside of their design limits.

Systems and subsystems are generally designed to operate within a "normal operating range" – that is, the variables describing the state of the system are supposed to vary (often in ways that are correlated with one another) within a limited domain of values. Performance may be better when the system is in some parts of its normal range than others, and the control processes are designed to keep it in the part of its normal range that is optimal for producing its intended outputs; but it is designed to operate successfully (and not to break down) so long as it stays within the normal boundaries. Indeed, systems are generally designed to continue to operate somewhat outside their normal ranges without resulting in catastrophic failure; while performance (understood as the cost-effective production of the outcomes that the system was designed to generate) may decline, and perhaps decline markedly, most systems can operate without immediate and complete destruction within a (wider) range of state variables outside the normal operating limits. We might refer to this broader range as constituting the "design limits" of the system. When systems are intentionally and intelligently designed, their engineers generally try to insure that they will not fail dramatically under any combination of state variables within their design limits, and they are generally (and professionally) conservative in making those calculations, so that the point at which significant damage to the system will occur actually lies somewhat beyond the design limits. If, however, the state variables of the system are pushed sufficiently far outside the design domain – past its "tolerance limits" – by definition the result will be breakdown and destruction.

Figure 1 illustrates the concepts of the optimal performance zone, the normal operating range, the design limits and the actual limits of the tolerance of the system (the boundary beyond which significant damage to the system will occur) in a notional system with two state variables (and thus a two-dimensional control space). As Figure 1 illustrates, in general the tolerance limits will lie outside the design limits – but if designers make errors in calculating how the system will behave, they may incorrectly estimate that the system is safe in areas where in fact it is in peril. Thus, when designers are not sufficiently cautious, there may be areas where the actual tolerance limits lie inside the calculated design limits, creating a zone of "unexpected failures". Beyond the actual tolerance limits of the system, it will self-destruct.

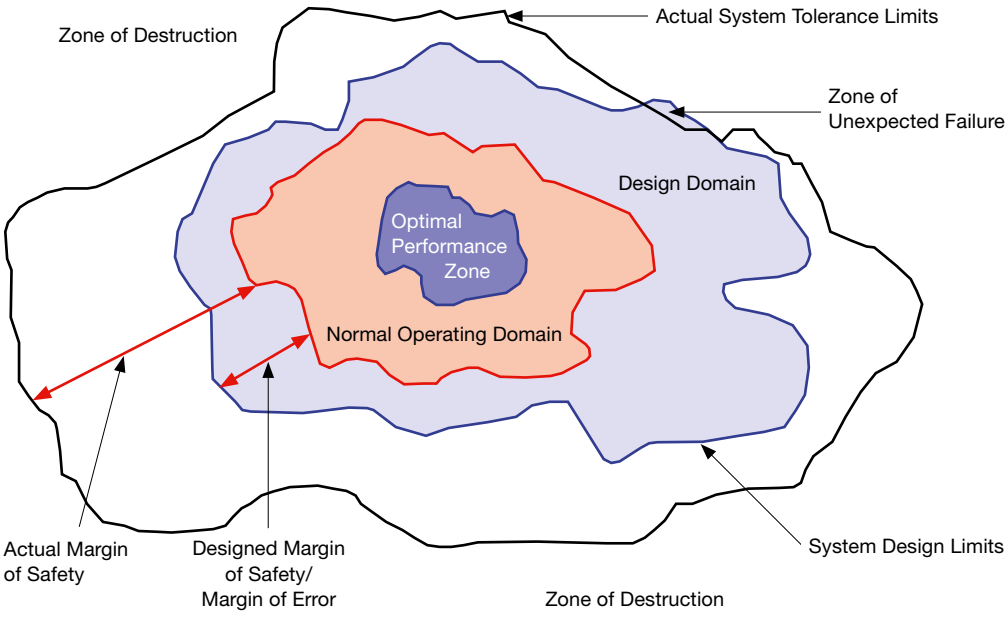

Figure 1: Notional illustration of relationship between performance, normal operations and design limits

Systems can be characterised in a number of ways…

Characterising systems in this way helps us to define better and to understand the nature of system accidents. As Perrow and others have detailed in an extensive literature providing mathematical proofs of the implications of the underlying characteristics, the key feature of a system subject to system collapse is a combination of complexity and interaction. Perrow details the requisite features of a system that will eventually experience a system accident:

1) **Multiple subsystems:** the system must consist of at least three subsystems;
2) **Complex interactions:** the subsystems must be interconnected to one another with multiple feedbacks that form loops (and thus allow for effects to cumulate);[8]
3) **Nonlinear interactions:** the outputs from one subsystem that are inputs to one or more other subsystems must be nonlinear – that is, a one-unit variation in the internal conditions within the subsystem that drive the output does not always result in the same amount of variation in the output;[9] and
4) **Tight coupling:** the outputs of one module or subsystem flow directly and immediately into the next module as inputs (with no "buffer" to mediate the "shock" of an ouput change before it becomes an input change to the next module); and
5) **Imperfect reliability:** subsystem components are imperfect, so that eventually an error will occur in more than one module at the same time.

Figure 2 shows a pictorial illustration of a system that meets these criteria (so long as some of the interactions or feedback flows are nonlinear).

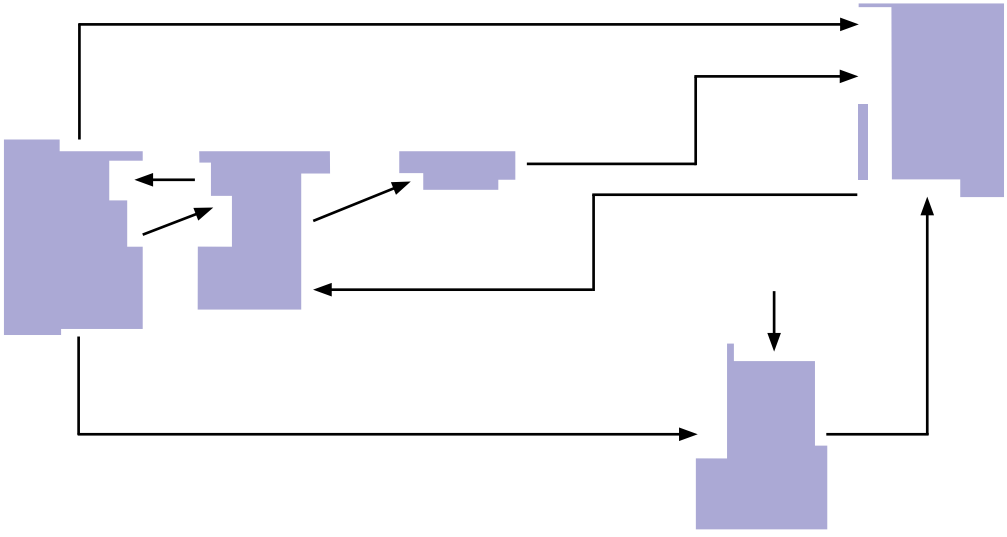

Figure 2:
Diagram of a system showing multiple complex tight couplings among subsystems

[8] Perrow describes the complex interactions that will turn out to be problematic as "those of unfamiliar sequences, or unplanned and unexpected sequences, (that are) either not visible or not immediately comprehensible". While Perrow does not specifically speak to this in *Normal Accidents,* complex systems derive much of the force of their complexity from their scale as a result of the square law of computation – that the number of interactions (or edges) in a system is the square of the objects (or nodes) in that system. The square law of computation was first articulated in G. M. Weinberg, *An Introduction to General Systems Thinking* (New York: Wiley-Interscience, 1975).

[9] Technically, this means that if there is a set of internal variables within a subsystem, x_1, x_2, x_3, … x_n, that drives an output of the module, y, where the output $y = f(x_1, x_2, x_3, \ldots x_n)$ is a function of the subsystem's internal variables, then y does not vary along a straight line when at least one of the values x_1, x_2, x_3, … x_n are changed, but instead varies along a curve. This implies that for some input variable x_i, the amount by which y changes when x_i varies by one unit changes as the value of x_i changes – sometimes, the variation of the x_is makes more of a difference than it does at other times, and this variability introduces a form of instability into the system.

…which provide indicators as to their vulnerability.

What Perrow and others have shown is that systems with these characteristics are vulnerable to – and, indeed, will inevitably generate (given a long enough operating period) – a system accident in which the system is pushed first outside of its ordinary operating limits and then beyond its collective design limits to a point where it will fail simultaneously and catastrophically as a system.

Complex interactions amplify errors in one subsystem…

While the proof that such an outcome is an inevitability is difficult and highly technical, the general idea behind system collapse and the way in which these characteristics will tend to generate system collapse is not difficult to discern. The first two conditions (multiplicity of subsystems and of interconnections or feedbacks between subsystems) jointly create a set of complex interactions that are often unknowable in advance and thus are unexpected when they occur. This is then accentuated by the nonlinearity condition. Once an error (condition 5) – or an external shock – enters the system, its consequences can flow in complex ways among modules, accumulating in feedback loops and encountering and perhaps being reinforced and amplified by other consequences generated by the same initial force, and thus growing and pushing the system farther from the center of its normal operating range.[10]

…a risk largely negated by more simple system constructions…

With simple interconnections, stable subsystems and one source of error, such a system might be manageable over a long period. It would be the easier to manage the more the flows of outputs from one module were only loosely coupled to the input to the next module – as, for example, when the output from one module flows into an inventory (a holding tank, of sorts) and the flow out of that tank is managed, so that a change on its input side is not instantly transmitted as a change in the output side that drives the next module. This might make it possible to manage it in a way that keeps the system as a whole from exhibiting system collapse. However, when the subsystems themselves have internal instabilities, when there are multiple sources of error flowing into the system simultaneously, when the interlinkages are complex – and, especially, when there are no buffer zones between modules to mediate the fluctuations in output signals from one before they become input signals to the next – the prospects for a chain reaction that pushes the system beyond its breaking point rises sharply.

…with clearer design limits and correction measures…

Very simple systems can be successfully operated within their design limits – and captured and either halted or corrected when they threaten to break through their design limits. By contrast, systems with a requisite level of complexity cannot, in principle, be managed always to stay permanently and reliably within their design limits. This is why system accidents are "normal", and it is a fundamental characteristic of systems with the requisite form and level of complexity and interactivity.

[10] Both positive and negative feedback loops can exist, with positive feedback loops tending to amplify and negative feedback loops tending to create binding constraints over time (and often over very short periods). See Lars Skyttner, *General Systems Theory: Ideas and Applications* (Singapore: World Scientific Press, 2001). Both positive and negative feedback loops can be disruptive, as amplification of an output can be destructive as input to another subsystem, just as constraint can restrict output flows that are needed to sustain a downstream tightly-coupled subsystem.

…that avoid the inevitability of collapse present in more complex systems.

System accidents pose real threats on a socially important scale. When we build systems of sufficient complexity, system accidents become inevitable. When the systems affected are large and important, the level of hazard and damage is correspondingly large. Yet, system collapses have another feature that multiplies their impacts – by their nature of having torn asunder the value-creating fabric for which the system was designed and built in the first place, *they are difficult to reconstitute.* Systems are a complex and often delicate ecology – that is part of what makes them vulnerable to system-level accidents, but it is also part of what guarantees that recovery from hazards of this kind may be long and difficult.

Examples of historical system accidents and the potential for future accidents

There are many systems – ranging from modest to large to very large in scale – that exhibit the characteristics of vulnerability to system accidents. We have seen several system collapses on very different scales that demonstrate the phenomenon and emphasise the wide range of scales on which it can operate. In addition to the South China snowstorm example described above, we will briefly outline three examples as an illustration.[11]

Three Mile Island demonstrated key vulnerabilites to multiple systems failure…

Example 1: Three Mile Island. A widely-examined instance in which a system accident actually took place, detailed at length in Perrow's work, is the collapse of the nuclear-powered electricity generating system at Three Mile Island that occurred in March of 1979. Nuclear power plants have a number of subsystems (heat/steam generation, cooling, electric power generation from steam, control systems, safety and monitoring systems, and so on). There are many complex and nonlinear interactions among these "modules", resulting in feedback loops that interconnect the overall system in ways that make power plants an archetypal example of the principles of vulnerability to system-level accidents. Perrow and others have traced how a series of breakdowns (and errors by the operators) spun up from being a minor maintenance issue to being an actual and near-complete meltdown of the nuclear core at the heart of plant. Safety systems built into the plant were designed to capture and render harmless significant departures of the system from its ordinary operations, but in the event proved unable to rectify the rising tide of interactions that were making the system as a whole increasingly unstable and increasingly out of control. Indeed, the safety systems themselves contributed to the problems – or, at least, the operators thought they had. This resulted in the operators shutting off parts of the safety systems (including the emergency backup cooling system, at a time when the core was already in the process of overheating) – which further exacerbated the spiralling consequences of the event.

[11] In addition to the examples given here, elements of the Hurricane Katrina catastrophe can also be understood through this lens. For example, Leavitt and Kiefer offer an argument for how the levee failures in the immediate aftermath of the storm constitute a system accident in: William M. Leavitt and John J. Kiefer, "Infrastructure Interdependency and the Creation of a Normal Disaster: The Case of Hurricane Katrina and the City of New Orleans", *Public Works Management and Policy,* 10 (2006): 306–314.

…while the 2003 US power grid collapse involved multiple power providers…

Example 2: The Northeasten power grid in August 2003. In the case of Three Mile Island, the system involved was a single power plant (and, actually, only half of it, as the second reactor at Three Mile Island was unaffected by the system collapse of its sister reactor). In August of 2003, the northeastern United States and southern Canada experienced a power grid collapse that plunged 50 million people into darkness – and, while in most places power was restored within 12 hours, in some parts of the affected area power was off for more than three days. The outage resulted from a series of anomalous flows of power that apparently originated in the electric grid in the Ohio River valley. Instead of being damped out (by automatic system monitoring, control and regulation devices, and vigilant operators), a combination of what apparently were operator errors together with unprecedented reversals of power flow through the grid resulted in an a spiral of consequences that literally shattered the grid – and which destroyed a sufficiently large number of components that reassembling the system into working order was very time-consuming. How long this took is still being debated; signs of difficulty appeared several hours before the collapse, but nearly all of the large-scale anomalies developed within the system in the course of the final 20 minutes before the collapse – and most of those actually arose within the last few seconds before the system blew itself up.

…with climate change as one of the most complex cases of systems interaction.

Example 3: Global climate change. The global ecological and climate system clearly exhibits the characteristics that dispose a system to the potential of system-level accidents (and some observers believe that we are in the early stages of what will become a catastrophic system collapse). Ecological systems by their nature involve multiple subsystems that are intimately interconnected with many complex linkages. Similarly, the climate system involves distinguishable components that interact with one another. Importantly, the two disparate systems of ecology and climate are themselves tightly bound together (and increasingly bound to the human industrial, agricultural, and commercial systems). Some climate changes can be self-reinforcing (as when warming melts snow and ice, reducing the earth's reflectivity and thus causing greater absorption of heat from the sun), but landscape-scale ecological effects can also affect climate (as when changes in average ambient temperature change the biomass carrying capacity of a forest region, reducing CO_2 absorption). While it is arguable whether the climate and ecological systems are currently in the process of generating a system collapse, we are certainly witnessing changes in those systems that are very large on an historical scale, and the intrinsic characteristics of these systems make large-scale accidents a plausible and therefore significant risk.

Complex systems interactions in very different scenarios have some common qualities…

These examples illustrate how prevalent system accidents already are – and they also indicate that other systems in our midst might be similarly vulnerable. For example, the South China snowstorm suggests that our general transport and food production and distribution systems might be vulnerable. Without doubt, they consist of multiple subsystems that are tightly bound to one another through multiple and often non-linear linkages. On an even larger scale, many of those who note that the global ecology and climate systems have these features view the ongoing evolution of global climate change as a (gradual-onset, but rapidly evolving) system collapse already in progress – and thus feel great urgency about finding ways to prevent or mitigate it.

…but some retain greater resilience than others.

Of course, not all complex systems exhibit these conditions, or accumulate them. Some systems have strong equilibrating mechanisms that push operations back towards the normal operating range (or towards the high performance zone) when they start to depart from those conditions. These systems exhibit stability – and they are therefore intrinsically more manageable, and are not the systems we need most to be concerned about. Our point here is not that all systems are intrinsically vulnerable, but that *some* are, and that there are systematic forces that tend to create (and to increase the prevalence of) vulnerable systems – which is the issue to which we now turn.

Forces that generate system vulnerabilities

Many system interactions evolved rather than being designed…

Where do these vulnerable systems – and system vulnerabilities – come from? Surely not from having been intentionally designed into any important systems that we utilise – in fact, to the contrary, most carefully-designed systems have been deliberately constructed so as to avoid the kinds of spiralling negative interactions that produce system accidents. *However, many of the systems we use (and inhabit) were not* ***designed*** *– they* ***"formed."*** In other words, many of the systems that are likely to have the greatest effects on our future circumstances are not the product of a conscious, intentional, intelligent design process (which might have avoided building in the characteristics that generate system vulnerability, and presumably would have done so to the extent that those features were understood and avoidable). Instead, systems – the interactions and interlocking of destinies among different component subsystems – formed as a result of natural or human forces that brought or pushed them together, causing them to interact.

…as can be seen in ecological systems…

Ecological systems provide a good example of systems that "form", driven by natural forces that end to create tight linkages. It is frequently to the evolutionary advantage of species, for example, to diversify their food sources – thus linking them to a larger collection of species, and increasing the extent to which effects in one part of the system will be transmitted (often rapidly) to others. The process of evolution leads life forms to "find" (through adaptation) new advantages, and while this will sometimes mean moving into a niche that is uncontested and unconnected to others, often it will mean forming additional links in an already strongly interconnected system. The intrinsic redundancy of such a system may contribute at the same time to its general resilience (in the face of small perturbations) and to its susceptibility (in the face of a larger disruption) to a cascading accident.

...which can be described as "self organising" systems that develop to a critical state...

Technically, the process of system "formation" from components is referred to as the creation of "self-organising" systems – distinguishing them from systems that were intentionally constructed (and, hopefully, designed intelligently). Such systems often exhibit a phenomenon called "self-organised criticality".[12] It is helpful to describe this concept with reference to a simple physical example, and it is often illustrated by reference to the development of a pile of sand on a table. Consider a table with a funnel above it into which sand is poured at a steady rate, so that a pile of sand begins to form on the table top. Suppose that the important ("critical") event in this system is defined as the shedding of sand off the edge of the table. At the outset, the pile grows without incident; the system is "subcritical" – it has not approached its critical state, where the important event will begin to occur. As the pile grows higher, it will also naturally spread out (in a way determined by what geologists refer to as sand's characteristic "angle of repose"), which in turn is determined by the size, shape, uniformity, roughness and resulting frictional characteristics of the grains of sand) until its edge begins to approach the edge of the table. At this point, the system is becoming poised in its critical state – that is, it has now evolved to the point where its critical events will begin to occur.

...at which point vulnerabilities can occur...

An important question now arises: what will be the nature of these critical events? Will sand pour at a constant rate off whichever edge of the table the pile first approaches? Or, instead, will sand cascade more episodically, in pulses, off the table? If it falls in pulses, will they be of more or less uniform size, or will there be a stable distribution of sizes? Will sand begin to fall off one edge, while the pile continues to grow in other directions? Will sand eventually begin to pour off more than one edge?

...bringing the system to a state of supercriticality...

Interestingly, in this and in many systems like it, the answer is that the system will continue to accumulate "energy" (sand, in this case) and will come to exhibit a stable distribution of critical events of various sizes that nearly always follows a logarithmic curve. As the system is forming, the distribution of event sizes is not stable, but instead is evolving. The system has reached "maturity" – referred to as having achieved "supercriticality" – when the size distribution of critical events has stabilised, and the next "cascade" of sand off the edge of the table will behave as a random draw from the power distribution of sizes.

...at which point events will occur...

An important implication of this (quite general) system phenomenon is that, once it has achieved its mature, supercritical state, *the system will generate critical events across a wide range of scales,* with many small events and some modest-sized events and – crucially – with low frequency but at least episodically, *very large events.*

...inevitably leading to system accidents.

An interesting way of interpreting system accidents in the kinds of human systems we are considering here is to think of these formed, self-organised systems as evolving towards a point of supercriticality, at which we will observe critical events of all scales with a random but probabilistically predictable frequency. If we then define a system accident as a critical event beyond a specified scale – how large do the financial ripples have to be before we define an ongoing event as a financial crisis? – system accidents will be an inevitably recurring phenomenon. Viewed this way, the conditions outlined above as leading to system accidents can be seen as the conditions that generate (either self-organised or designed-in) supercriticality.

[12] This concept was originally developed and articulated by Chao Tang Bak and Kurt Wiesenfeld in "Self Organizing Criticality", *Physical Review A,* Vol. 38, No. 1, July 1, 1988.

Forces that push human systems toward supercriticality

Both system complexity and tightness contribute to adverse events.

The question of what forces seem to be generating an increasing array of significant accidents in the systems that surround us thus becomes the question of what forces tend to bring systems towards and into a supercritical state. In both designed and self-organising human systems, the forces seem to be similar. We might divide them broadly into two types: first, there is a tendency for more complex linkages to form (or, in the case of designed systems, to be formed); second, there is a tendency to increase the "tightness" of the linkages between different parts of the system.

Complexity develops as additional linkages develop to successful systems…

Greater complexity may tend to arise over time because the additional linkages constitute ways to take advantage of the flow and activities of the system. In the example of global climate change discussed above, for example, evolution was observed to be a natural force that often generates additional linkages among species, thus increasing the interconnectedness of the overall system. The same is true in economic systems – only here the actions of the agents are often conscious and intentional. Seeing advantages to connecting different aspects of commerce, economic agents create new links that facilitate valuable commercial activity in good times but that also enable the ripple effects that can tear such systems asunder in the face of significant simultaneous disruptions in component subsystems. Information technology systems that enable "just in time" inventory methods in manufacturing, food distribution and other areas of economic life generate benefits to the agents that invent them and to many of the participants in the more complex system that results – but also may make that same system more susceptible to self-reinforcing destructive reverberations. Competitive forces (in ecology and in commerce) can often create self-organising pressures that cause system vulnerabilities to accumulate.

…and technology facilitates greater system interaction.

Greater complexity may also arise over time because it has become technologically more sustainable. In a world without computers, we could not possibly operate a structure as complex as the current air traffic control system; we are enabled to build more complex systems by the advent of new information management capacities. On the one hand, they allow us to construct and manage more complex systems that would be impossible (or would produce catastrophic levels of errors and collisions) in their absence. On the other hand, by enabling the construction of more complex and interdependent systems, they also create conditions that may lead inexorably to periodic system accidents.

Tight system coupling gains efficiency but reduces the buffer zones that protect against system accidents.

Tighter coupling may also tend to increase over time for "efficiency" reasons. What tight coupling means is that there is no buffer zone between the components of the system – and output from one module immediately becomes an input to the next. This is good – and efficient – in the sense that there are no inventories where costly resources are "parked" without providing direct benefits; resources generated by one module are immediately used in the next. The problem arises because this creates the potential for rapid transmission of errors, and thus instability – when one module goes out of its control space (and its outputs similarly depart from the expected range), the next module in line is immediately being driven by inputs that are out of the anticipated range – and it, too, is thus nearly immediately operating out of its control space as well. Reducing the vulnerability thus generated generally requires the construction of a "buffer" of some kind between the components. When module A goes out of control, the flow into the buffer is affected, but the flow out of the buffer can still be regulated, at least for a while, to stay in control – the level in the buffer rises (or falls), so the buffer is being pushed away from its normal state, but module B is not immediately affected.

This has the effect of delaying the transmission of the "error," in effect isolating the error in module A (rather than transmitting it instantly to other modules), and giving system operators or participants (or the intrinsic or designed-in resilience within module A) a chance to rectify the challenge while it is still confined to one module (rather than allowing it to sweep across the whole system). However, such buffers constitute (and consume) real resources – they cost money, and "efficiency" thus often seems to call for reducing or eliminating them. They do not tangibly produce value – instead, they intangibly reduce risk – and as a result their value may be easy to overlook.[13]

Greater complexity and tightness brings systems interaction to the state of supercriticality.

All of these forces can be viewed as the natural result of powerful and systematic economic forces. The competition to create new opportunities tends to create additional linkages; the desire to take advantage of the benefits of greater complexity pushes the development of information and control technologies to enable the (temporary) management of more interdependent systems; the pressure to conserve scarce resources tends to push system designers, participants and engineers to see buffers as costly and superfluous, leading to tighter coupling within the system. Thus, in human systems driven by economic considerations – which is a very large class of human systems indeed – systematic economic forces drive both designed and self-organising systems towards being balanced on the point of supercriticality.

Social management of the risks of system accidents

How well prepared are we, as societies, to confront the rising level of social risk that this phenomenon is generating? Unfortunately, there is good reason to think that neither individuals, organisations nor governments are either very aware of or very well-designed to manage these risks in advance or to cope with their consequences when they arise.

Many complex systems are not systemically regulated.

First, since many of the systems that create large-scale social vulnerabilities are not comprehensively designed, but instead are at least partly self-organising, there is no natural regulatory process that encompasses them. A nuclear power plant is fully designed, and does operate within a nexus of defined authority (in the US, in the form of the Nuclear Regulatory Commission). However, there is no similar comprehensive design nor overall supervisor for the world food production and distribution system, nor for the world manufacturing system, nor for the world financial system.

[13] We are using the term "buffer" here both broadly and loosely. There are, of course, mechanisms other than buffers (narrowly defined as inventories accumulated between the output side of one module and the input side of the next module to which it is linked) that can be used to regulate or mediate the relationships across a system link. The output from one module can be influenced (and, in the extreme, controlled) before it becomes an input to the next module by control gates, switches, impedance mechanisms, flow restrictors, check points, back-flow preventers, circuit breakers, and so on. For example, according to an analysis presented by Ivor van Heerden and Mike Bryan, the levee breaches in New Orleans were along drainage outflow canals and navigation canals where storm surge came back into the outflow canals; simple gates (backflow preventers) costing something in the range of USD 100–150 million have prevented this problem (Ivor van Heerden and Mike Bryan, *The Storm;* New York: Penguin Books, 2006). While such mechanisms can thus be very useful, we should note that the addition of these control devices constitutes yet another complexity in the system that, in addition to keeping it safe from some anomalous interactions, may also make it more vulnerable to others.

Individuals frequently fail to perceive the high degree of system interconnectivity…

Second, participants in systems often don't see what they are involved in as systems, and therefore may not be able (or even interested) to see the interdependencies (and to notice the attendant risks). Participants do not generally have transactions with a large fraction of a system – they tend to be engaged with one or a few components. Consumers of food, for example, interact almost exclusively with the retail end of the distribution system; by and large, they have little natural contact with even the wholesale system, or any of the other components that stand behind it; farmers deal with the other end of the chain (and with suppliers of seed and fertiliser), but see even the very next components down the line from their operation – the wholesale agricultural commodities purchasing and transport systems that buy and carry their products away – only in part and only briefly. Generally, there is little about participation in a complex system that systematically reveals the nature of the larger system – or that even invites interest in it.[14]

…or system vulnerability…

Third, there is widespread lack of understanding of the nature of system accidents, of the risks associated with them, and of the conditions that create vulnerability to them. In the absence of this understanding – and the ability of relevant, large-scale actors to be able to develop a broad, system-level perspective – it is difficult to take any significant steps to minimise the associated social hazards.

…for which governments can also be poorly prepared.

Fourth, large institution-regulating organisations – in a word, governments – do not always tend to think of the things they manage as systems, do not appear to be aware of the conditions that generate vulnerability, and may not have developed the instruments necessary to reduce the hazards that these conditions collectively produce.

System vulnerabilities may ultimately exceed government capacity.

Finally, many of the systems that exhibit the characteristics of vulnerability outlined here are of a scale that transcends existing government structures – as is perhaps best illustrated by the piecemeal and incompletely coordinated attempts to manage the recent (ongoing?) financial crisis and the attempts to coordinate reforms of the system. The financial system is global in scope, self-organising, and strongly driven by the economic forces that tend to push systems toward supercriticality; it was not comprehensively designed (and therefore can easily develop difficult-to-notice linkages that generate significant potential for instability and system-level failure); its current evolution and the development of new linkages is undertaken by self-interested agents who can't or don't see the larger system as a whole (and who might not care if they could or did); and it is not owned or overseen by any single (or even a well-organised and coordinated group of) regulatory overseers. It should be no great wonder, then, that it may have been (and may still be) self-organising itself into a supercritical state in which it is subject to system accidents.

Reducing the risks of system level accidents

What, then, can societies do to reduce the likelihood and cost of system accidents? The analysis above suggests several possible (non-exclusive) approaches:

1) We can become more aware of the presence of complex, nonlinear, tightly-coupled systems – and more alert to their vulnerability and to the scale of potential losses should they be triggered;

[14] Optimistically, however, there appears to be a growing public interest in where our food originates and how it is processed and delivered. Michael Pollan's work, *In Defense of Food: An Eater's Manifesto* (New York: The Penguin Press, 2008), and his other popular publications point to the growing general interest in healthy food and the "locavore" movement, which (at least locally) both simplifies the food system and brings people more directly into contact with the key elements of the system they are participating in.

2) Where potentially vulnerable systems have been identified, we can try to intervene (or induce others to intervene) to eliminate one or more of the preconditions to disaster through any of three broad approaches: (a) by improving diagnostic and control systems, which could allow us to keep some systems more reliably within desired tolerances (but we have to be careful that dependence on these more finely tuned control systems does not simply create an additional complexity that adds to vulnerability instead of resolving it); or (b) more commonly, by identifying areas in the system that are most likely to create or retransmit errors that could spin up into system accidents, and then isolating them through inserting buffer zones or other flow-control devices between them and the modules they are linked to; and/or (c) by directly simplifying systems by breaking them into a larger number of subsystems that are better isolated from one another.[15] More generally, reducing system accident vulnerability means trying to "redesign" either comprehensively designed or self-organising systems so that they exhibit less complexity, have fewer linkages overall and fewer nonlinearities in particular, and have more buffering and controls between modules. The simpler systems thus developed may be more expensive to build and to operate in the short run, but they will be less prone to large-scale accidents because it is easier to isolate difficulties and anomalies within one or a small number of modules rather than having them rapidly transmitted throughout the system (so long as we engineer the interactions among them with appropriate controls and buffers); and
3) We can be more careful to recognise the possibility of large-scale system accidents, and can better prepare to respond to and recover from their correspondingly large-scale consequences.

Conclusion

Strong and durable forces (including most notably the evolution of technology and the increasing drive for short-run economic efficiency) are producing increasingly prevalent conditions for system-level collapses. Wise organisations will recognise this, identify the systems they are embedded in (or managing or overseeing) and are dependent upon, and will do what they can to prevent these hazards from eventuating. They will also organise themselves to be better able to respond to, and to be more resilient in the face of, the unavoidable systemic risks that remain.

Herman B. "Dutch" Leonard is George F. Baker Jr. Professor of Public Management at Harvard University's John F. Kennedy School of Government and Eliot I. Snider and Family Professor of Business Administration and Co-Chair of the Social Enterprise Initiative at Harvard Business School.

Arnold M. Howitt is Executive Director of the Roy and Lila Ash Center for Democratic Governance and Innovation at Harvard University's John F. Kennedy School of Government.

The two authors co-direct the Program on Crisis Leadership at the Kennedy School of Government's Taubman Center for State and Local Government and Ash Center for Democratic Governance and Innovation.

[15] Here, the square law of complexity works in our favour; by creating a larger array of simpler subsystems, each component by itself becomes less internally complex; if we reduce the number of elements within a component by half, we have reduced its internal complexity by a factor of four – we wind up with a larger number of more manageable components. This process may, however, generate a larger number of interactions between the resulting larger number of subsystems, and in those interactions the square law is again acting against us – we need to be careful to engineer the external interactions between the larger number of system modules in a way that mediates the flows between them effectively.