A System-Theoretic, Control-Inspired View and Approach to Process Safety

Nancy G. Leveson
Department of Aeronautics and Astronautics, and Engineering Systems Division
Massachusetts Institute of Technology
and
George Stephanopoulos
Department of Chemical Engineering
Massachusetts Institute of Technology

Accidents in the process industry continue to occur, and we do not seem to be making much progress in reducing them (Venkatasubramanian, 2011). Post mortem analysis has indicated that they were preventable and had similar systemic causes (Kletz, 2003). Why do we fail to learn from the past and make adequate changes to prevent their reappearance? A variety of explanations have been offered; operators' faults, component failures, lax supervision of operations, poor maintenance, etc. All of these explanations, and many others, have been exhaustively studied, analyzed, "systematized" into causal groups, and a variety of approaches have been developed to address them. Even so, they still occur with significant numbers of fatalities and injured people, with significant disruption of productive operations and frequently extensive destruction of the surrounding environment, both physical and social.

Is it true that the problem of ensuring safe process operations is so complex that it defies our technical and managerial abilities to address in an effective manner, or is it that we approach the problem in wrong ways? Should we consider process accidents as "normal" events and factor in our considerations the cost of addressing their consequences? What goes on in the course of engineering activities that addresses process safety issues is not purely formalizable, either in abstract terms or in taxonomic views. In other words, it cannot be converted into an executable algorithm. However, the authors believe that the problem has an inherent structure, which is conducive to effective techniques. Both its structure and its techniques can be taught and learned, and even more significantly they define the scope within which engineers and managers can deploy their personal touch, not only in trivialities, but in deeper considerations of skill and suitability.

The authors of this Perspective are convinced that accidents continue to occur, and near misses multiply in alarming numbers for three basic reasons: (1) Analysis methods used, do not discover all the underlying causes of events. (2) Learning from experience does not work as it is supposed to do. (3) Learning is happening in the wrong places. Consequently, it is imperative that we should re-examine the entrenched beliefs, assumptions and paradigms that underlie the engineering of safe processing systems, and attempt to identify disconnects to the prevailing experience.

It is the intention of this Perspective to offer a system theoretic view and approach to a rational, all-encompassing framework for addressing process safety. It is based on a control-inspired statement of the process safety problem, which is amenable to modern model-predictive control approaches, and can encompass all potential systemic constraints (from those on engineered systems at the processing level, to

those associated with operations management, regulations by governmental agencies or standards by insurance companies, or legislation, governing operating plants), whose violation leads to accidents.

The Perspective starts by questioning the applicability of the accident causation models that have constituted the basis for the development of almost all process engineering tools dealing with process safety, e.g. HAZOP analysis, fault-trees, and event-trees. It then proceeds to offer the "enforcement of safety constraints" as the fundamental underpinning of a comprehensive framework for process safety engineering, both at the development phase of a new process and during operations. Then, Section 2 defines Process Safety as a *system problem* by underlining its fundamental distinction from reliability engineering. Within a system-theoretic framework, it introduces the hierarchical organization of a socio-technical system, which allows for a comprehensive treatment of process safety, and constitutes the framework for the control-inspired view and approach for the engineering of safe processes, which is detailed in Section 3.

It is the authors' expectation/hope that academics and industrial practitioners of Process Systems Engineering (PSE) will find the system-theoretic, control-inspired view and approach to process safety, proposed in this Perspective, a natural setting for thinking of process safety in new and more effective ways. It is also the authors' conviction that the proposed framework may offer an attractive venue for the deployment of numerous methodologies and techniques that have been significantly advanced in other areas of PSE (Stephanopoulos and Reklaitis, 2011), such as; computer-aided modeling and simulation of large and complex processes, model-predictive multivariable control and optimization, large-scale monitoring and diagnosis, planning and scheduling of process operations, all of which would have a material effect in advancing the state-of-the-art in process safety.

1. The Prevailing Premise: Accident Causation Models and Their Shortcomings

In Process Safety the prevailing assumption is:

Accidents are caused by chains of directly related failure events.

This assumption implies that working backward from the loss event and identifying directly related predecessor events (usually failures of process components, or human errors) will identify the "root cause" for the loss. Using this information, either the "root cause" event is eliminated or an attempt is made to stop the propagation of events by adding barriers between events, by preventing individual failure events in the chain, or by redesigning the system so that multiple failures are required before propagation can occur (putting "*and*" gates into the event chain). Figure 1 shows a typical chain-of-events model for a tank rupture. The events are annotated in the figure with standard engineering "fixes" to eliminate the event.

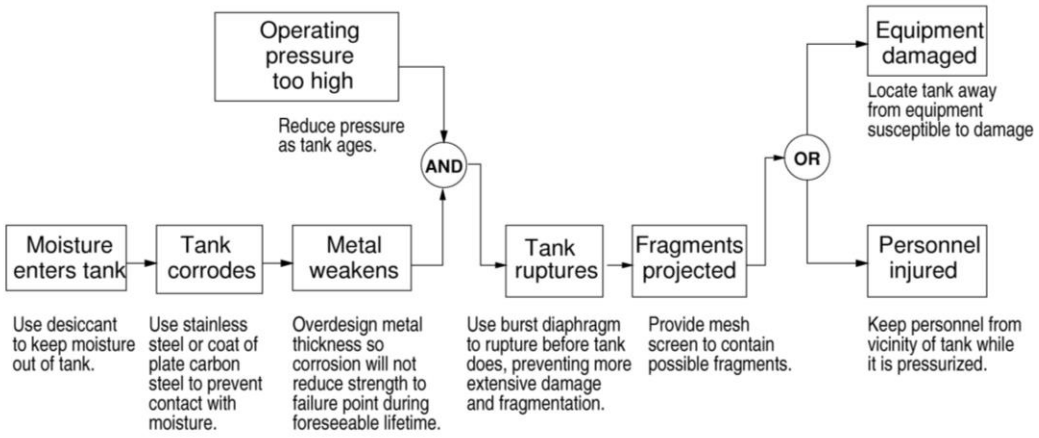

Figure 1: An Example Chain-of-Events Model for a Tank Rupture from [Leveson, 2012]

Accident causation models constitute the foundational piece on which many of the process safety methodologies and engineering tools, such as Event-Trees, Fault-Trees analysis [Lapp and Powers, 1977; Erickson, 2011], Hazard and Operability Analysis (HAZOP) [Kletz, 1999; Venkatasubramanian et. al., 2000], risk analysis (AIChE/CCPS, 2009), have been constructed and deployed. However, this model has several serious drawbacks, such as the following:

Oversimplifies causality and the accident process: Whether in the process design mode (i.e. when safety considerations are taken into account for the design of a safer process), or the post mortem accident analysis mode, the application of the chain-of-events model is based on the classical assumption that cause and effect must be directly related. Consequently, these models are limited in that they are unable to go beyond the designers' or investigators' collective current knowledge, i.e. they cannot find causes that they do not already know.

Furthermore, most current accident models and accident analysis techniques suffer from the limitation of considering only the events underlying an accident and not the entire *accident process*. Accidents are often viewed as some unfortunate coincidence of factors that come together at one particular point in time and lead to the loss. This belief arises from too narrow a view of the causal time line. However, processing systems are not static. Rather than accidents being a chance occurrence of multiple independent events, they tend to involve a *migration to a state of increasing risk over time*. A point is reached where an accident is inevitable (unless the high risk is detected and reduced) and the particular events involved are somewhat irrelevant: if those events had not occurred, something else would have led to the loss. This concept is reflected in the common observation that a loss was "*an accident waiting to happen*".

Understanding and preventing or detecting system migration to states of higher risk requires that our accident models consider the *processes* involved in accidents and not simply the events and conditions: Processes control a sequence of events and describe system and human behavior as it changes and adapts over time (perhaps as a result of feedback or a changing environment), rather than considering individual events and human actions.

Excludes many of the systemic factors in accidents and indirect or non-linear interactions among events: Accident causation models oversimplify the causality because they exclude systemic factors, which have only an indirect relationship to the events and conditions in the chain of events. A few attempts have been made to include systemic factors, but they turned out to be severely limited in achieving their goals. In accident causation models no account is made for common causes of the failures of the barriers or the other types of events in the chains. These "systemic" accident causes can defeat multiple barriers and

other design techniques that are assumed to be independent. In such cases, the popular Swiss Cheese Model (Reason, 1990) is totally inadequate to capture the propagation of events leading to an accident.

It does not account for the hierarchical interaction of processes-management-regulation-legislation: Accident causation is a complex process involving the entire socio-technical system including legislators, government regulatory agencies, industry associations and insurance companies, company management, technical and engineering personnel, operators, etc. To understand why an accident has occurred, the entire process needs to be examined, not just the proximate events in the event chain. Otherwise, only symptoms will be identified and fixed, and accidents will continue to recur.

It does not provide effective treatment for common causes: Migration to higher-risk states may occur as a result of a single common cause, e.g. competitive or financial pressures that force people to cut corners or to behave in more risky ways [Rasmussen 1997]. As an example, consider the Bhopal accident. None of the safety devices, for example, the vent scrubber, flare tower, water spouts, refrigeration system, alarms, and monitoring instruments worked at the time of the accident. At first glance, the failure of all these devices at the same time appears to be an event of extremely small probability or likelihood. But these "failure" events were far from independent. Financial and other pressures led to reduced maintenance of the safety devices, turning off safety devices such as refrigeration to save money, hiring less qualified staff, and taking short cuts to increase productivity. An audit two years before the accident noted many of the factors involved, such as nonoperational safety devices and unsafe practices, but nothing was done to fix them.

1.1 A System-Theoretic View of Process Safety

More effective process safety analysis methodologies and techniques that avoid the limitations of those based on chain-of-events models are possible, if they are grounded on systems thinking and systems theory. Systems theory dates from the 1930s and 1940s and was a response to the limitations of the classic analysis techniques in coping with the increasingly complex systems being built [Checkland, 1981]. Systems theory is also the theoretical basis for system engineering and Process Systems Engineering (PSE), as practiced extensively by chemical engineers (Stephanopoulos and Reklaitis, 2011).

In the traditional *analytic reductionist approach* of classical chemical engineering science, processing systems are broken into distinct unit operations and other operating components, such as the elements of control loops, and safety devices. The behavior of the individual physical elements can be modeled and analyzed separately and their behavior is decomposed into events over time. Then, the behavior of the whole system is described by the behavior of the individual elements and the interactions among these elements. A set of underlying assumptions assures that this is a reasonable approach for designing and analyzing the behavior of many processing systems and has been in practice for a long time. The underlying assumptions are: (a) Separation of the process into its components is feasible, i.e. each component or subsystem operates independently and analysis results are not distorted when these components are considered separately. (b) Processing components or behavioral events are not subject to feedback loops and other non-linear interactions, and that the behavior of the components is the same when examined singly as when they are playing their part in the whole. Note that the elements of a feedback loop (i.e. sensor, actuator, and controller) are viewed as components of the overall processing system, not as constituents of a processing component. (c) The principles governing the assembly of the components into the whole are straightforward, that is, the interactions among the subsystems are simple enough that they can be considered separate from the behavior of the subsystems themselves.

In contrast to the above, the system approach focuses on the processing system as a whole, and does not decompose its behavior into events over time. It assumes that some properties of the processing system can only be treated adequately in their entirety, taking into account all facets related not only to the technical and physical-chemical underpinnings of the processing operations, but also the human, social,

legislative and regulatory aspects surrounding the process itself [Ramo, 1973]. These system properties derive from the relationships among the parts of the system: how the parts interact and fit together [Ackoff, 1971]. Thus, the system approach concentrates on the analysis and design of the whole as distinct from its components, and provides a means for studying emergent system properties, such as process safety [Leveson, 2009].

Using the system approach as a foundation, new types of accident analysis (both retroactive and proactive) can be devised that go beyond simply looking at events and can identify the processes and systemic factors behind the losses, and also the factors (reasons) for migration toward states of increasing risk. This information can be used to design controls that prevent hazardous states by changing the design of the processing system to prevent or control the hazards, prevent migration of the operational state to higher-risk domains and, in operational systems, detect the increasing risk before a loss occurs.

2. **Process Safety is a System Problem**

Process Safety is not part of the mission or reason for the existence of a chemical plant. Its mission, its reason of existence is to produce chemicals. To be safe in terms of not exposing bystanders and the surrounding environment to destructive effects of unleashed toxins and/or shock waves is *a constraint* on how the mission of a chemical plant can be achieved, where *by constraint we imply limitations on the behavioral degree of freedom of the system components* (definition of a constraint in system theory). This seemingly trivial observation and statement has far reaching ramifications on how Process Safety should be viewed and taught, and how safe processes should be designed and operated. The reason for this assertion is simple:

> *Accidents occur when these process safety-imposed constraints are violated, and given that these constraints are imposed on the operational state of a chemical plant as a system, one concludes that process safety is a problem that must be addressed within the scope of an operating plant seen as a system.*

In the following paragraphs of this section we will further amplify the logical consequences of viewing process safety as a system problem.

2.1 Process Safety is different from Reliability Engineering

Process Safety currently rests on an underlying assumption that *Process Safety is increased by increasing the reliability of the individual system components*. If components do not fail, then accidents will not occur.

However, a high-reliability chemical process, i.e. a process with highly reliable engineered components (vessels, piping, connectors, pumps, sensors, control valves, control algorithms, etc.) is not necessarily a safer process. Dysfunctional interactions among the process components can lead to unsafe operations, while all components are perfectly functioning, as intended. For example, consider the case of an accident that occurred in a batch chemical reactor in England (Kletz, 1982). The design of this system is shown in Figure 1. The computer was responsible for controlling the flow of catalyst into the reactor and also the flow of water into the reflux condenser to remove heat from the reactor. Sensor inputs to the computer were provided as warning signals of any problems in various parts of the plant. The specifications of the monitoring and control algorithms required that if a fault was detected in the plant, the controlled inputs should be left at their current values, and the system should raise an alarm. On one occasion, the computer received a signal indicating a low oil level in a gearbox. The computer reacted as its specifications required: It sounded an alarm and left the controls as they were. This occurred when catalyst had just been added to the reactor and the computer-based control logic had just started to increase the cooling water flow to the reflux condenser; the cooling water flow was therefore kept at a

low rate. The reactor overheated, the relief valve lifted, and the contents of the reactor were discharged into the atmosphere.

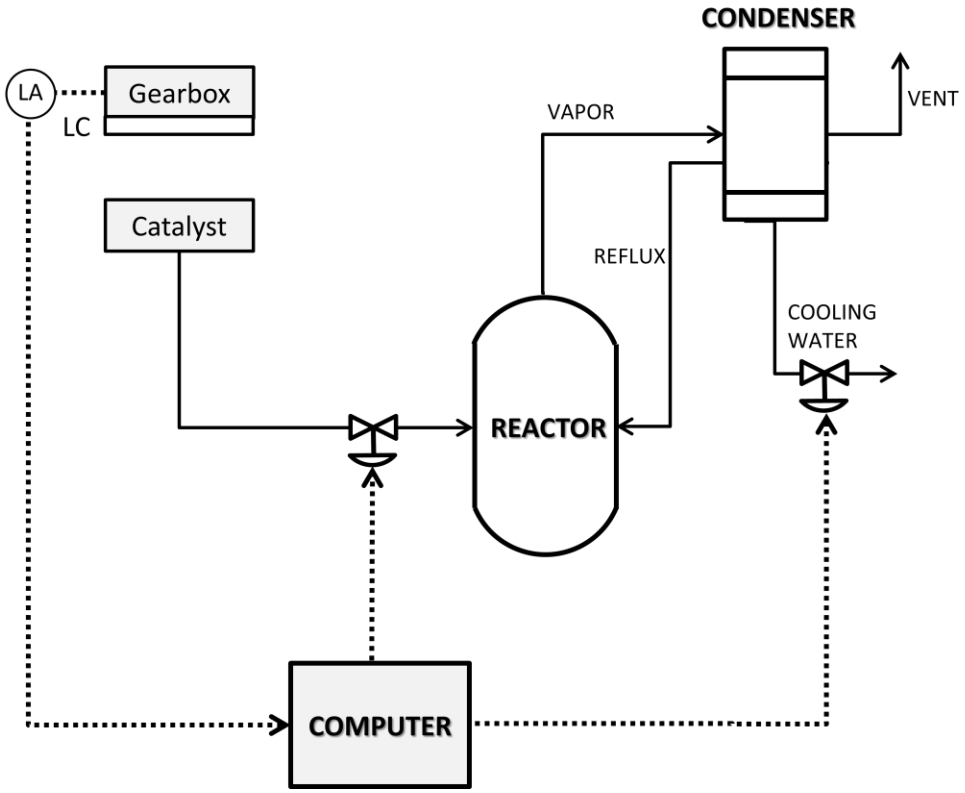

Figure 2. Batch reactor system (adapted from Kletz, 1982)

There were no component failures involved in this accident: the individual components, including the software, worked as specified, but together they created a hazardous system state. Merely increasing the reliability of the individual components or protecting against their failure would not have prevented the loss. Prevention required identifying and eliminating or mitigating unsafe interactions among the system components. Indeed, most software-related accidents have been system accidents—they stem from the operation of the software, not from its *lack* of operation and usually that operation is exactly what the software engineers (mistakenly) intended. Thus event models as well as system design and analysis methods that focus on classic types of failure events will not apply to software. Confusion about this point is reflected in many fault trees containing useless (and misleading) boxes that say *"Software Fails."* Software is the design of a machine abstracted from its physical realization, for example, the logical design of a multivariable control system, separated from any physical design to implement that logic in hardware. What does it mean to talk about an abstraction or a design failing?

Another example comes from the post mortem analysis of the events that led to the 2005 explosion of the isomerization unit at BP's Texas City refinery. The record has indicated that there were malfunctioning sensors, stuck valves, and violation of operating procedures by operators, all of which can be seen as "component failures". If one were to accept this tantalizingly attractive explanation of the accident, one would not have uncovered the *systemic dysfunctional interactions* at higher levels of management, which led to the simultaneous failure of several components.

Safety and reliability are *different* system properties. *Reliability is a component property* and in engineering is defined as the probability that a component satisfies its specified behavioral requirements over time and under given conditions. Failure rates of individual components in chemical processes are fairly low, and the probability of simultaneous failure of two or more components is very low. On the

other hand, *Process Safety is a system property* and can be defined as absence of accidents, where an *accident* is defined as an event involving an unplanned and unacceptable loss (Leveson, 1995). One does not imply nor require the other—a system can be reliable and unsafe or safe and unreliable. In some cases, these two system properties are conflicting, i.e., making the system safer may decrease reliability and enhancing reliability may decrease safety. For example, increasing the burst-pressure to working-pressure ratio of a tank often introduces new dangers of an explosion in the event of a rupture (Leveson, 2005).

As chemical processes have become more economical to operate, their complexity has increased commensurably; many material recycles, heat and power integration, frequent changes of optimal operating points, integration of multivariable control and operational optimization. The type of accidents that result from *dysfunctional interactions* among the various process components is becoming the more frequent source of accidents. In the past, the designs of chemical processes were intellectually manageable (serial processes with a few recycles, operating at fixed steady-states over long periods of time), and the potential interactions among its various components could be thoroughly planned, understood, anticipated, and guarded against. In addition, thorough testing was possible and could be used to eliminate design errors before system use. Highly efficient modern chemical processes no longer satisfy these properties and system design errors are increasingly the cause of major accidents, even when all components have operated reliably, i.e. have not failed.

2.2. Process Safety should be viewed within a socio-technical hierarchical framework

A chemical process is not operating as a purely engineered system, driven only by the physical and chemical/biological phenomena in its operations, and its safety cannot be viewed solely in terms of its technical components alone. Many other functions have a direct or indirect effect on how a process is operated, monitored for process safety, assessed for risk, and evaluated for compliance to a set of regulatory constraints (e.g. environmental, process safety regulations). It is operated by human operators, it is serviced and repaired by maintenance personnel, and it is continuously upgraded and improved by process engineers. Operational managers of process unit areas or whole plants and managers of personal and process safety departments deploy and monitor process performance metrics, execution of standard operating procedures, and compliance with health-safety-environmental regulations [Olivea et. al., 2006; Baker Panel, 2007; Hopkins, 2009; Urbina, 2010]. Higher up in the hierarchy, company-wide groups define rules of expected behavior, compile best practices and safety policy standards, receive and analyze incident reports, and assess risks. Even higher in the hierarchy, local, state, or federal legislative and/or regulatory bodies define, deploy, and monitor the compliance of a set of rules, all of which are intended to protect the social and physical environment in which the process operates.

Indeed, as Lederer (1968) observed, system safety should include non-technical aspects of paramount significance in assessing the risks in a processing system: "System safety covers the entire spectrum of risk management. It goes beyond the hardware and associated procedures to system safety engineering. It involves: attitudes and motivation of designers and production people, employee/management rapport, the relation of industrial associations among themselves and with government, human factors in supervision and quality control, documentation on the interfaces of industrial and public safety with design and operations, the interest and attitudes of top management, the effects of the legal system on accident investigations and exchange of information, the certification of critical workers, political considerations, resources, public sentiment and many other non-technical but vital influences on the attainment of an acceptable level of risk control. These non-technical aspects of system safety cannot be ignored".

The idea of modeling socio-technical systems, using process-control concepts is not a new one.
Jay Forrester in the 1960s, for example, created *System Dynamics* using such an approach (Forrester,1961). Industrial engineering models often include both the management and technical aspects

of systems. As one example, Johansson (Suokas, 1988) describes a production system as four subsystems: physical, human, information, and management. The physical subsystem includes the inanimate objects—equipment, facilities, and materials. The human subsystem controls the physical subsystem. The information subsystem provides flow and exchange of information that authorizes activity, guides effort, evaluates performance, and provides overall direction. The organizational and management subsystem establishes goals and objectives for the organization and its functional components, allocates authority and responsibility, and generally guides activities for the entire organization and its parts.

To account for the hierarchical organization of the various socio-technical functions, which have a bearing on the operation of a processing plant, and thus on its safety, Rasmussen (1997) proposed the structured hierarchy shown in Figure 3. Aimed at risk management (AIChE/CCPS, 2009), at all levels their model focuses on information flow. Also, at each level, and between levels, Rasmussen models the associated events, their initiation, and their effects through an event-chain modeling language, similar to the cause-consequence models, described in Section 1.1, which combine the use of traditional fault-trees and events-trees. In addition, he focuses on the downstream part of the chain, following the occurrence of the hazard, a fairly common attitude in the process industry. Finally, it should be noted that his model focuses on operations—process engineering design activities are treated as input to the model and are not a central part of the model itself.

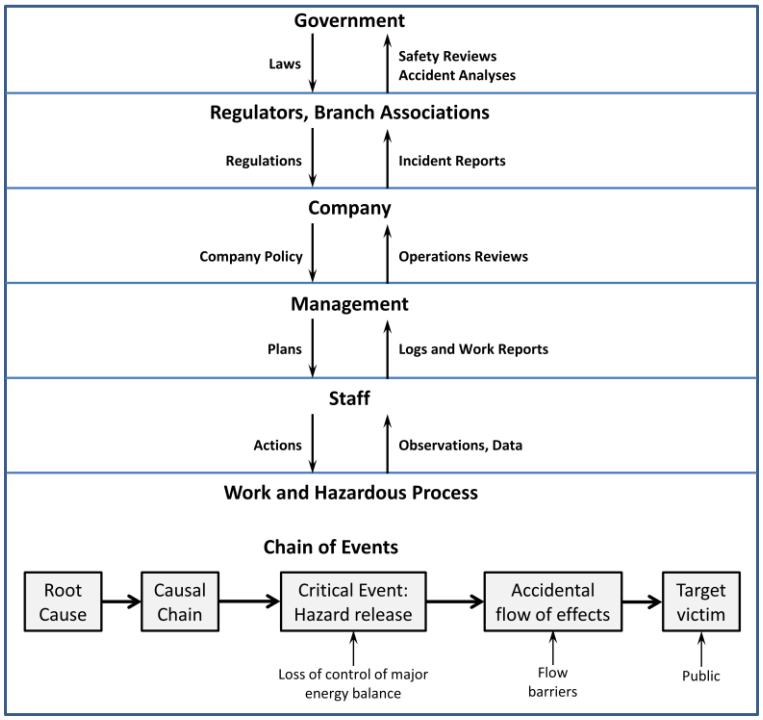

Figure 3. Rasmussen's socio-technical model of system operations and risk assessment (adapted from Rasmussen, 1997).

Rasmussen 's hierarchical organization of the socio-technical factors, related to process safety, is clearly in the right direction and offers an effective framework for treating process safety in a system-theoretic manner. However, the preservation of linear accident causation models as their fundamental premise, leads to similar weaknesses as those discussed in Section 1. In Section 3 we will present an alternative approach, which, while it maintains Rasmussen's hierarchical socio-technical model of system operations and risk assessment, it employs a control-inspired framework with the *safety constraint violation* principle as its foundational element for a comprehensive and exhaustive treatment of process safety.

3. A Control-Inspired Engineering Approach to Process Safety

Process safety is an emergent property of systems that arises from the interaction of system components. Determining whether a plant is acceptably safe, for example, is not possible by examining a single valve in the plant. In fact, statements about the "safety of the valve", without information about the context in which that valve is used, are meaningless. Safety can only be determined by the relationship between the valve and the other plant components—that is, in the context of the whole. Therefore, it is not possible to take a single system component in isolation and assess its safety. A component that is perfectly safe in one system may not be when used in another.

Emergent properties are defined by a set of constraints, which relate the behavior of the system's processing components. Accidents result from interactions among components that violate the safety constraints. These violations may result from inadequate monitoring of the safety constraints, absence or inadequacy of control elements, which would provide sufficient corrective feedback action.

> *A control-inspired view of process safety suggests that accidents occur when external disturbances, component failures, or dysfunctional interactions among processing components are not adequately handled by the existing control systems, leading to a violation of the underlying safety constraints.*

In this view, process safety is a control problem and is managed by a properly designed control structure, which is imbedded in the adaptive socio-technical hierarchical system of Figure 3.

Process Safety during process development, design and engineering. The control structure that ensures the satisfaction of the safety constraints should be designed when process safety is considered during the development, design and engineering of a processing plant. It should provide the definition of all classical elements in the specification of a plant-wide control structure, i.e.

1. control objectives (i.e. the complete set of safety constraints),
2. set of measurements, which monitor the state of the control objectives,
3. set of manipulations, which can ensure the accomplishment of desired values for the control objectives, and
4. model-based controller logic that relates the measurements to the manipulations in ensuring the satisfaction of the control objectives.

In Figure 4 (Leveson, 2004), the left-hand side of the diagram shows the upward flow of monitoring signals and the downward flow of feedback control signals, identified during the development, design, and engineering of a processing system. It is in this phase of defining the scope of the control structures that accident causation models exhibit their most pronounced inadequacy; fault-trees and event-trees cannot identify all the pertinent safety constraints that arise from the dysfunctional interactions of processing components, or from the interaction of management functions and processing operations. The reason, as we have discussed earlier, is simple; they cannot account for such interactions in the absence of a system-wide treatment of the hierarchical organization of all factors that affect the safety of the process.

Process Safety Management during process operation. The process leading up to an accident during plant operation can be described as an adaptive feedback function, which fails to ensure the satisfaction of a safety constraint, due to changes over time in operational performance and/or management functions to meet a complex set of operating goals and values (right-hand side of Figure 4). In Figure 4, the right-hand side of the diagram shows the upward flow of monitoring signals and the downward flow of feedback control signals during the operation of chemical plant. It interacts with the left-hand side of the hierarchical organization. Consequently, the control structure conceived during process development, design and engineering, should not be viewed as a static design. Instead, it should be viewed as a continually evolving system that adapts to the changes introduced to the processing system itself and its

environment. Therefore, instead of defining process safety management as a concerted effort to prevent component failures, we should focus its actions in monitoring the state of safety constraints and ensuring that they are satisfied, as operating conditions change, different regulations are imposed on the operation of the plant, or different structure and culture are introduced in the management of the plant.

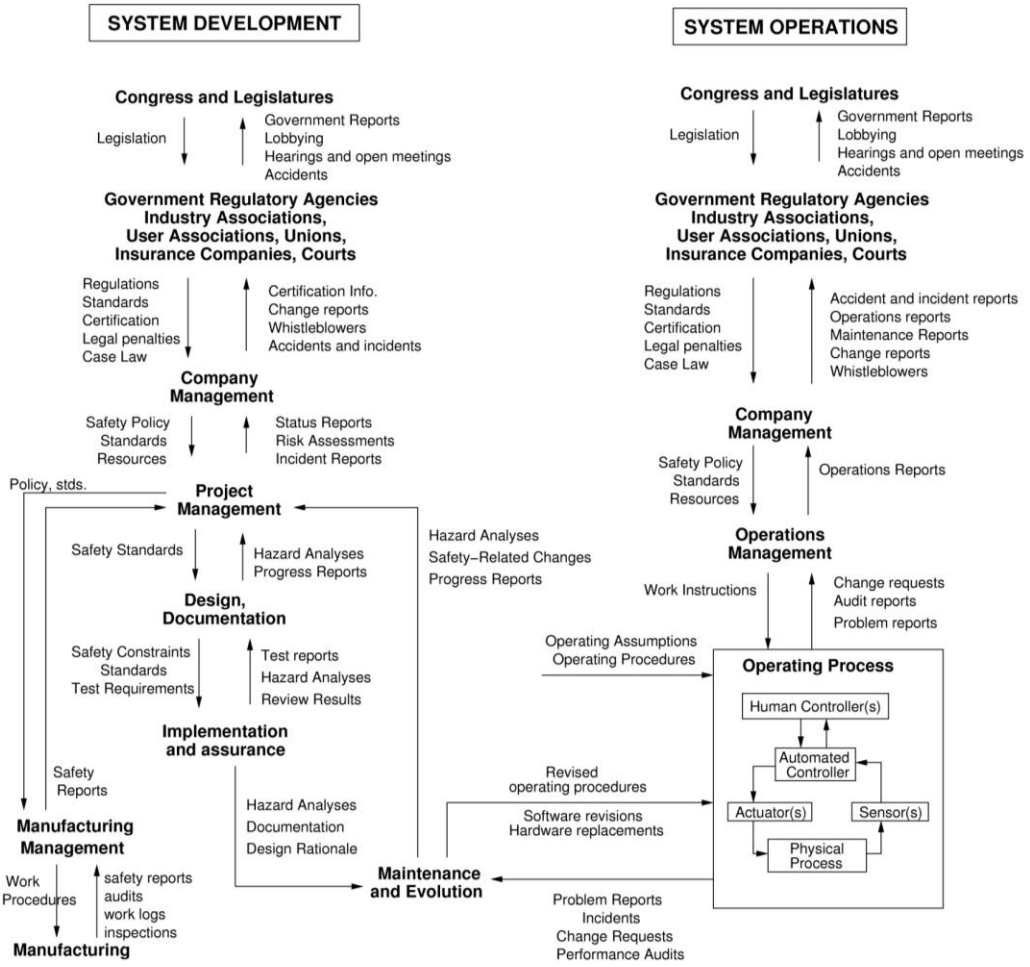

Figure 4. The hierarchical organization of control structures for the monitoring and control of safety constraints during the development and operation of a chemical process (Leveson, 2004).

3.1 Safety Constraints: The Pivotal element in Process Safety

In the prevailing framework of accident causation models the basic element is the failure of a component. In the control-inspired, system-theoretic view of process safety, the basic element is a safety constraint. Accidents occur because safety constraints were never imposed during process development and design, or, in case that they were imposed, because they were inadequately monitored or controlled at each level of the hierarchical socio-technical organization of Figure 3.

But, what exactly are the constraints? Obvious safety constraints in processing plants are, (a) the restriction that hydrocarbons to air ratio is outside the explosion range, and (b) chemicals must be under positive control at all times. Other technical constraints may involve restrictions on the pressures of vessels with gaseous components, levels of liquids in processing units, loads on compressors, operating temperatures in reactors or separators, or flows throughout the plant. Traditional operating envelopes are manifestations of these safety constraints. However, all of these constraints are local, restricted to the

operation of a single unit, or a cluster of units with a common processing goal, e.g. the condenser and reboiler in conjunction with the associated distillation column. It is crucial that safety constraints involving several processing units be satisfied, such as; material and energy balances in a dynamic setting, should be obeyed at all time-points, and for all units, plant sections, and the entire plant. Monitoring these balances over time ensures that one will be able to observe the migration of operating states towards situations of higher risk. Unfortunately, in very few plants one will encounter computer-aided systems that monitor these balances, attempt to diagnose and reconcile any deviations, and alert human operators for further actions. Usually, the enforcement of these constraints is left to the human operators, who observe the trends of critical variables and "compute" the balances in their own mind. In the Texas City accident, the migration of the material and energy balances in the isomerization unit towards higher-risk values took several hours to reach critical points, and no one (automatic system or human operator) observed it until it was too late.

Emergent properties, that relate the operations of processing units that are physically removed from each other, are normally not constrained by safety constraints, because the accident causation models do not reveal such restrictions. The violation of material or energy balances over multi-unit sections of chemical plants is a typical example of an emergent property that is often overlooked. Furthermore, emergent properties resulting from management functions are not constrained, since they have never been the explicit goal of a comprehensive process safety treatment. For example, in the Texas City isomerization explosion the repeated violation of safe operating procedures by the startup operators never constituted a violation of an important safety constraint in the eyes of the supervisory management.

As the operation of chemical processes is increasingly controlled by software systems, the ensuing complexity, in identifying and enforcing safety constraints, increases exponentially. This is what Leveson has called the *curse of flexibility* (Leveson, 1995). Physical constraints restrict the values of physical quantities and thus impose discipline on the development, design, construction, operation of a chemical plant. They also control the complexity of the processing plants that are being built. With model-predictive control software, we can simultaneously vary the values of hundreds of flow-controlling valves to regulate the values of hundreds of measured outputs, and with modern real-time optimization algorithms we can optimize process operations by varying the values of hundreds of control set points. *Controlling the limits of what is possible to accomplish with software systems is different from what can be accomplished successfully and safely.* The limiting factors change from the structural integrity and physical constraints on materials to limits on human intellectual capabilities. The accident caused by the correct deployment of a software-based control system in the batch reactor of Figure 1, is a manifestation of the dangers lurking in the increasing usage of software systems in the control and optimization of processing operations. In this example, the primary concern is not the "failure" of the software control system, but the lack of appropriate safety constraints on the behavior of the software system. Clearly, the solution is to identify the required constraints and enforce them in the software and overall system design.

The interplay between human operators and safety constraints is crucial in ensuring the satisfaction of safety constraints (Cook, 1996). In times past, the operators were located close to process operations, and this proximity allowed a sensory perception of the status of the process safety constraints via direct "measurement", such as vibration, sound, and temperature. Displays were directly linked to the process via analog signals and thus a direct extension of it. As the distance between the operators and the process grew longer, due to the introduction of electromechanical and then digital measurement, control, and display systems, the designers had to synthesize and provide a "virtual" image of the process operations' state. The monitoring and control of safety constraints became more indirect, through the monitoring and control of the "virtual" process safety constraints.

Thus, modern computer-aided monitoring and control of process operations introduces a new set of safety constraints that the designers must account for. Designers must anticipate particular operating situations and provide the operators with sufficient information to allow the monitoring and control of critical safety constraints. This is possible within the system-theoretic view of a plant's operation, but it requires careful

design. For example, a designer should make certain that the information an operator receives about the status of a valve is related to the valve's status, i.e. open or close, not on whether power had been applied to the valve or not, as happened in the Three Mile Island accident. Direct displays of the status of all safety constraints would be highly desirable, but present practices do not promise that it will be available any time soon.

3.2 Controlling Safety Constraints: A Model-Predictive Control approach

The system-theoretic, control-inspired view of process safety, advanced in this Perspective, is grounded on the modern principles of process Model-Predictive Control. Thus, instead of decomposing the process into its structural elements and defining accidents as the result of a flow of events, as prevailing methodologies do, it describes processing systems and accidents in terms of a hierarchy of adaptive feedback control systems, as shown in Figure 4.

At each level of the hierarchy a set of feedback control structures ensures the satisfaction of the control objectives, i.e. of the safety constraints. Figure 5 shows the structure of a typical feedback control loop and the process models associated with the design of the controllers. The dotted lines indicate that the human supervisor may have direct access to system state information (not provided by the computer) and may have ways to manipulate the controlled process other than through computer commands. In general, in each feedback loop there exist two types of controllers; an *automated controller*, and a *human supervisor* (as *human controller*). Both require models for the deployment of the model-predictive controllers. To appreciate the type of models needed, we have to digress and discuss certain fundamental principles of model-predictive control.

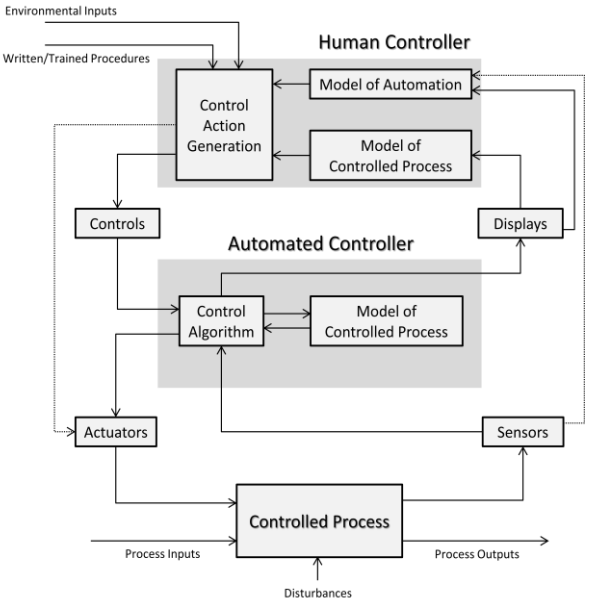

Figure 5. A typical control loop and the associated model-predictive controllers to ensure the satisfaction of the safety constraints.

In general, effective control of safety constraints requires that the following four conditions are satisfied (Ashby, 1956; Leveson, 2004):

1. The controller must have a goal or goals (e.g., to maintain the validity of the safety constraint),
2. The controller must be able to affect the state of the system,
3. The controller must be (or contain) a model of the process, and
4. The controller must be able to ascertain the state of the system.

To satisfy Condition 3, every model-predictive controller incorporates in its design two models:

(a) A process model that relates its measured output to the manipulated control input (e.g. a valve that regulates a flow).
(b) A model of the external disturbances and how they affect the value of the regulated output. If the external disturbances are unmeasured, an estimate must be provided.

Whether the process models are embedded in the control logic of an automated, model-predictive controller, or in the mental model maintained by a human controller, they must contain the same type of information:

i. The current state of the safety constraint,
ii. The relationships between the state of the safety constraint (output) and the two inputs, i.e. manipulated input, and external disturbance(s),
iii. The ways that the state of the safety constraint can change.

For the feedback control of safety constraints, associated with the physical and chemical/biological phenomena taking place in operating processing units, the questions of process model development, state estimation and design of feedback model-predictive controllers have received quite satisfactory answers with impressive results in industrial practice (Stephanopoulos and Reklaitis, 2011). These models are quantitative and can be effectively adapted in time, through the judicious use of measured process inputs and outputs. However, for the feedback control of safety constraints that require human supervision, the associated models are much simpler, they tend to be qualitative/logical (e.g. causal graphs), or involve simple quantifiable measures (e.g. dead-times, time constants, gains). Systematic development of satisfactory models for human controllers lags seriously behind. In addition, not providing appropriate feedback to the operators can result in incorrect operating behavior, which is frequently blamed on the operator and not on the system design. As just one example, at Texas City there were no sensors above the maximum permitted level of liquid in the ISOM tower and the operators were therefore unaware of the dangerous state of the system (Baker Panel, 2007).

Furthermore, as Figure 5 suggests, human controllers (e.g. human operators) interacting with automated controllers, in addition to having a model of the controlled process, must have a model of the automated controller's behavior, in order to monitor and supervise it. In the chemical industry the number of accidents and near-misses, caused by inaccurate mental models of the controller's behavior, is significant.

If a computer is acting in a purely supervisory role, i.e. the computer does not issue control commands to the process actuators, but instead its software provides decision-making tools to the human operators, then the logic of the software system providing the advice should contain a model of the process, as discussed earlier. As Leveson (2004) has observed, "Common arguments that in this design the software is not safety-critical are not justified—it is still a critical part of the functioning of the control loop and software errors can lead to accidents."

While the previous discussion has primarily focused on the criticality of process models, one should not underestimate the importance of models associated with the other elements of the control loop in Figure 5, e.g. sensors, actuators, displays of the state of process operations (including posting of alarms), or interfaces with automated computer control systems. The latter two are particularly important, as a series of accidents in chemical plants have demonstrated.

3.3 The Crux of the Matter: Managing control flaws in a hierarchically structured system

As Figure 4 suggests, legislation established at the highest level is converted into regulations, standards, and certification procedures, which in turn, at a company level, define company policies on safety, and operating standards, which are then converted into work instructions and operating procedures at the

operations management level. In other words, constraints established at a higher level, define behavior at a lower level. Thus, what we observe as "behavior" of engineered or human systems at the level of an operating plant, has been decisively shaped by decisions made at higher levels. In this view, process safety arises as an emergent property of a complex system, and should be treated as such within the framework of the hierarchical organization shown in Figure 4.

Audits and reports on operations, accidents, problem areas, gained at the lowest level of process operations provides important information for the adaptation of work instructions, maintenance and/or operating procedures, which in turn may lead to adaptation of company policies on safety and standards. In case of high profile destructive accidents, or a series of reported events, new tighter regulations may be imposed, or new legislation may be introduced to empower tighter regulations.

The feedback loops, depicted in Figure 4, are expected to operate effectively and prevent the occurrence of destructive accidents, but they do not, and accidents with severe losses continue to occur. We can attribute the occurrence of accidents to a series of control structure flaws that lead to the violation of the safety constraints, that the control systems are assumed to monitor and regulate. Table 1 lists the classification of these control system flaws, which are organized in three groups, as discussed in the following paragraphs, and Figure 6 shows the distribution of these “flaws” to the various elements of the control structure.

---

1. Inadequate Enforcement of Safety Constraints (Control Objectives)
    a. Unidentified hazards
    b. Inappropriate, ineffective, or missing control actions for identified hazards
        i. Control algorithm does not enforce safety constraints
            – Flaws in the controller design process
            – Process changes without appropriate changes (adaptations) in the control algorithm (asynchronous evolution)
            – Incorrect modification of adaptation of the control algorithm
        ii. Process Models inconsistent, incomplete, or incorrect
            – Flaws in the model generation process
            – Flaws in updating/adapting process models (asynchronous evolution)
            – Time lags and measurement inaccuracies not accounted for
        iii. Inadequate coordination among controllers and decision makers
            – Overlap of boundaries and areas of responsibility
2. Inadequate execution of control actions
    a. Flaws in the communication channels of measurement and actuating signals
    b. Inadequate operation of actuators
    c. Time lag
3. Inadequate or missing feedback loops
    a. Not provided in control system design
    b. Communication flaws
    c. Inadequate sensor operation (incorrect, or no information provided)
    d. Time lag

Table 1. Classification of flaws leading to violation of safety constraints and associated hazards

---

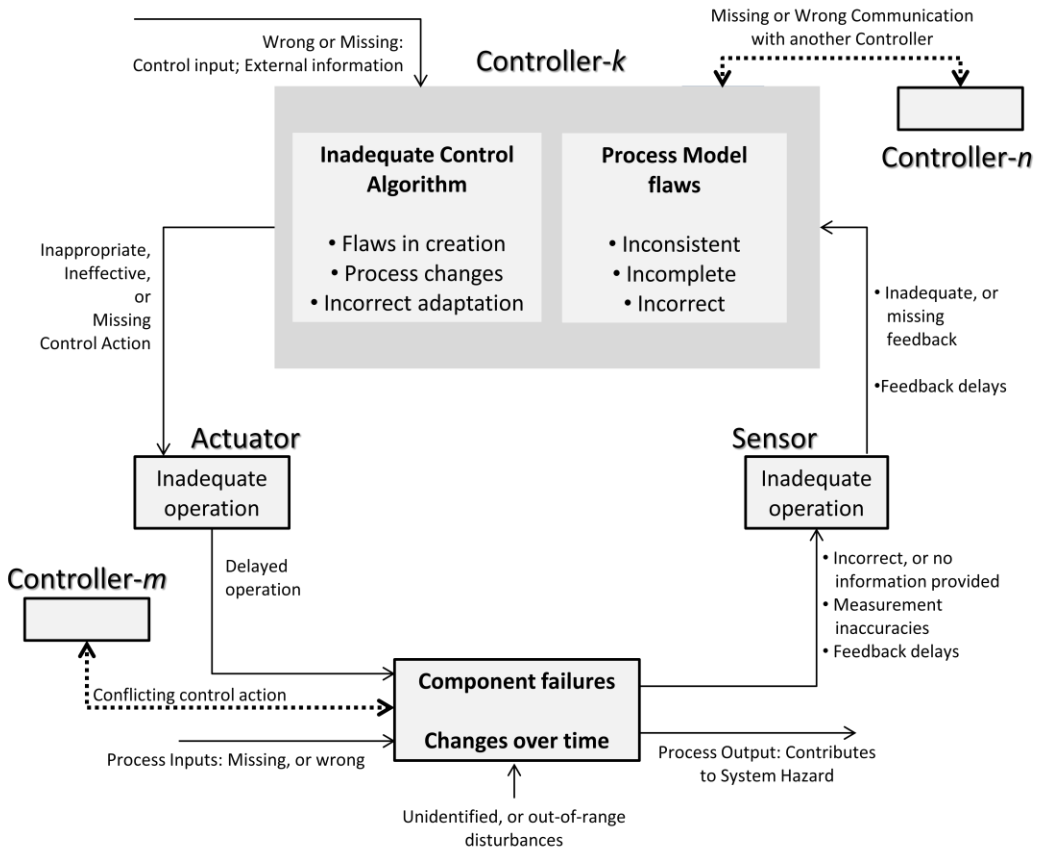

Figure 6. Distribution of the control "flaws" in the elements of the control structure.

Inadequate enforcement of safety constraints: This can occur either because the safety constraints associated with specific hazards were not identified, or because the control actions do not adequately enforce the satisfaction of the imposed safety constraints. The latter may result from *flawed control algorithms*, *inconsistent or incorrect models* used to describe process behavior or effects of external disturbances, and/or *inadequate coordination* among multiple controllers.

Unidentified safety constraints is a well-appreciated pitfall that engineers try to avoid as vigorously as possible. A system-theoretic view of process safety within the framework of the hierarchical system shown in Figure 4, imposes a discipline that minimizes the number of potentially missed safety constraints.

Flawed control algorithms may be the result of flawed initial designs, changing process characteristics (e.g. catalyst deactivation, fouling of heat exchangers, leaks in flow systems), poor or inadequate maintenance/adaptation by humans entrusted with this task.

The reliance of automatic or human controllers on process models, as discussed earlier, is quite critical in ensuring safety constraints. In fact, inconsistent or incorrect process models can be found in a significant percentage of accidents or near misses in chemical industry. When the process model employed by an automatic controller or a human controller does not represent the true process behavior, any corrective action is bound to be ineffective, or even detrimental. The most common sources of inconsistency in process models are: (i) Changes in process characteristics, as discussed above, (ii) the models inability to account for failures in processing components, (iii) the modeling errors in inferential control systems (where secondary measurements are used to infer the values of regulated variables), and (b) incompleteness in terms of defining process behavior for all possible process states, or all possible

disturbances. Clearly no model can be complete in an absolute sense, but it should be complete enough to capture the anticipated states that a process can go through during operation.

Inadequate coordination among different controllers, automated and/or human, can result from overlapping boundaries of responsibility (who is really in charge of control?). They may involved only automatic controllers, such as is the common case of conflicting PID loops, or overlapping responsibilities of several departments such as, maintenance and repairs, process engineering, and operations supervision.

Inadequate execution of control actions: In this case, faults in the transmission or implementation of control commands (e.g. failure of an actuating valve) lead to the violation of safety constraints. A common flaw during the development of the safety system is that safety-related information (e.g. HAZOP studies, fault-trees, event-trees) was inadequately communicated to the designer of the overall safety system.

Inadequate or missing feedback loop: Missing feedback loops on safety constraints could result from poor safety system design, while inadequacies in feedback systems may be traced to inadequacies of the channels carrying measurement and actuating signals. In all cases, one wonders whether the designers of the safety constraint control loops (automatic or human) were given all the necessary information about the process and the desired operating objectives. Alternatively, one may pose the question as to whether the safety system designers were given all the pertinent process models.

4. Implications of the System-Theoretic Approach to Process Safety

From the discussion in Section 3, it is clear that accidents occur when (a) important safety constraints have not been identified and thus not controlled, or (b) control structures or control actions, for identified safety constraints, do not enforce these constraints. In this section we will discuss the implications of these conclusions on certain important aspects of process safety engineering.

4.1 Identifying Potential Hazards and Safety Constraints

Nagel and Stephanopoulos (1996) proposed an approach for the identification of potential hazards in a chemical process. Their approach uses a system theoretic framework, which allows for the identification of all potential hazards, i.e. it is complete, in contrast to other methods such as, HAZOP, fault-trees, or event-trees. In this approach, *inductive reasoning* is employed to identify, (a) all chemical and/or physical interactions, which could lead to a hazard, and (b) the requisite conditions that would enable the occurrence of these hazards. Then, *deductive reasoning* was used to convert these enabling conditions for hazards into *process design* and/or *operational "faults"*, which would constitute the causes for the occurrence of the hazards.

Within this framework, the concept of a "fault" is identical to the notion of the "safety constraint", as has been defined in earlier sections. Consequently, the proposed system-theoretic approach of the Perspective is perfectly positioned to provide methods for the complete identification of potential hazards at the level of engineered systems, e.g. processing units, sensors, actuators, and controllers. In addition, it can also extend to higher levels of the socio-technical organization and, using analogous methodologies, identify the complete set of potential hazards and associated safety constraints.

4.2 Monitoring Safety Constraints and Diagnosing Migration to High Risk Operations

Methodologies on monitoring and diagnosing the operational state of a chemical process have received a lot of attention in chemical engineering (Calandranis et al., 1990; Venkatasubramanian et al., 2003a, b and c) by PSE researchers, both academic and industrial practitioners. However, they have always been

considered as self-standing systems, not part of a broader process safety program. Within the framework of the system-theoretic framework for process safety, proposed in this Perspective, monitoring and diagnosis tasks become essential elements of the adaptive control structure that ensures the satisfaction of the safety constraints. Integrating the two into one comprehensive system, that monitors the migration of operating states to higher risk domains, and diagnosing the reasons for this migration, becomes a natural task for process systems engineers. This integration is highly attractive as a research problem and highly promising in creating engineering tools with significant impact on process safety.

4.3 Causal Analysis: Learning from Accidents Using Systems Theory

After an accident or an incident the ensuing analysis is heavily biased by hindsight. Minimizing this bias is essential in "learning" the systemic factors that led to the events. The first step is getting away from the "blame game" (who is to blame for the accident/incident) and focusing on the "why" it happened and "how" it can be prevented in the future.

After an accident/incident it is easy to see what people "should have", "could have", or "would have" done to avoid it, but it is nearly impossible to understand how the world looked to someone not having the knowledge of the outcome. Simply finding and highlighting the mistakes people did, explains nothing, and emphasizing what they did not do or should have done does not explain why they did what they did. Thus, based on the prevailing models of accident causation, the teams that analyze accidents/incidents, attempt to establish cause-and-effect chains of events, which would explain the outcomes. In doing so they may fall into one or more of the following traps: They (a) tend to oversimplify causality because they can start from the outcome and reason backwards to the presumed "causes"; (b) overestimate peoples' ability to foresee the outcome, because it is already known; (c) overrate "violations" of rules or procedures; (c) misjudge the prominence or relevance of data/information presented to people at the time; and (d) match outcomes with actions, i.e. if the outcome was bad, actions leading to it must have been bad.

A post-accident/incident investigation, within a system-theoretic view, takes a very different approach. The purpose of the analysis is to: (a) Reveal if the proper safety constraints were in place at all levels of the socio-technical organization of the process safety system, and if they were not, why they were not, and how can they be established to avoid future occurrences. (b) Identify the weaknesses in the safety control structures that allowed the accident/incident to occur. Thus, rather than judging people for what they did or did not do, the emphasis is on explaining why it made sense for people to do what they did, and what changes are required in the control structures that ensure the satisfaction of safety constraints. The answers may be found in the list of control flaws shown in Table 1; unhandled environmental disturbances or operating dictates; unhandled or uncontrolled component failures; dysfunctional (unsafe) interactions among processing components; degradation of control structures over time; inadequate coordination of multiple controllers (automated and/or human).

4.4 Addressing Human Errors

Human error is considered as the most frequent reason for accidents in chemical industry, and the prevailing attitude for a long time has been to reward "correct" behavior and punish "incorrect" actions. What has not been sufficiently appreciated is the fact that human behavior is heavily influenced by the environment in which humans operate, i.e. the human environment of a plant's management structure and culture (e.g. emphasis on cost-cutting efficiency), and the specific milieu of engineered systems, which were developed by the process developers/designers (e.g. sizes and operating conditions of processing units, logic of automated controllers in place, schedules of operating procedures, features of interfaces to computer-aided control systems, etc.).

Within the scope of the prevailing chain-of-events approach, it is usually very difficult if not impossible to find an "event" that precedes and is causal to the observed operator behavior. For example, it is nearly

impossible to link, in a clear way, particular design features of the processing units or automated controllers to operator behavior. Furthermore, instructions and written procedures on how to start-up a plant, or switch its operation to a new state, both results of detailed engineering work, are almost never followed exactly, as operators strive to become more efficient and productive, and deal with time and other pressures (Rassmussen, 1997).

Within a system-theoretic framework, the handling of operator behavior is based on two elements: (a) The "Operator's Mental Model" of the process, and (b) its relationship to the "Designer's Model", and the model of the "Actual System" (Figure 7). Process designers deal with systems as ideals or averages, and they provide procedures to operators with respect to this ideal. Processes may deviate from the ideal through manufacturing and construction variances, or through evolution over time. Operators must deal with the existing system and change their mental models (and operational procedures), using operational experience and experimentation. While procedures may be updated over time, there is necessarily a time lag in this updating process and operators must deal with the existing system state. Based on current information, the operators' actual behavior may differ from the prescribed procedures. When the deviation is correct (the designers' models are less accurate than the operators' models at that particular instant in time), then the operators are considered to be doing their job. When the operators' models are incorrect, they are often blamed for any unfortunate results, even though their incorrect actions may have been reasonable given the information they had at the time.

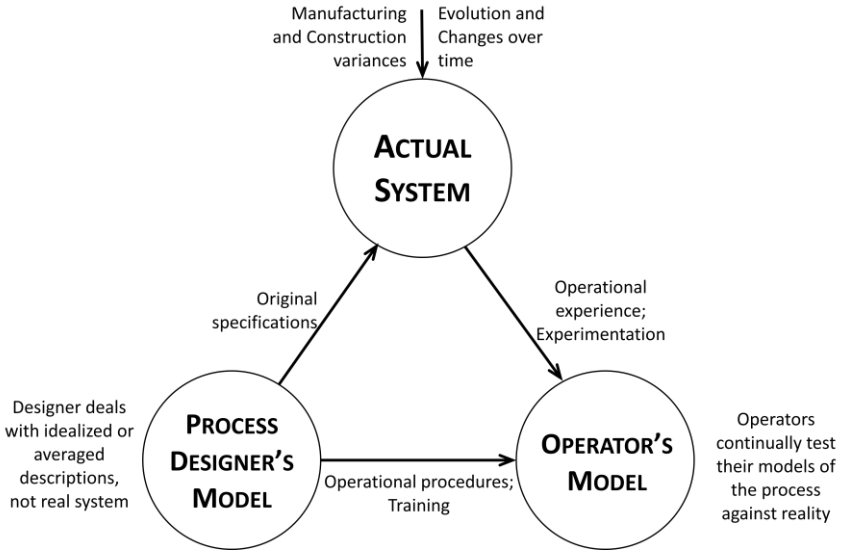

Figure 7. The role of mental models in operations (adapted from Leveson et al., 2009)

The system-theoretic framework is particularly effective in handling the interactions among the three different views of a process, as shown in Figure 7. It captures all the elements that need to be considered in defining the (a) safety constraints that should be controlled, as discrepancies arise among the three models, (b) the measurements that need to be made, in order to monitor the development of these discrepancies, and (c) the control actions that need to be taken in order to minimize the discrepancies.

5. Summary

Shifting the focus of discussion from component failures to violations of safety constraints has been the primary objective of this Perspective. Once this view has been accepted, then, process safety engineering can be established within a broader and comprehensive framework, fully supported by the vast array of

system-theoretic methodologies, at the center of which is the *design and operational integrity of a control structure* that ensures the satisfaction of the safety constraints. For the systematic deployment of these ideas, Leveson (2004) developed STAMP (*Systems-Theoretic Accident Model and Processes*), and more recently (Leveson, 2012 and 2013), STPA (*System-Theoretic Process Analysis*), which has been implemented with significant degrees of success in a variety of industries, e.g. aeronautics and aerospace, nuclear power, chemical processes, oil and gas production, and pharmaceutical.

Within the discipline of chemical engineering, Process Systems Engineering (PSE) can be credited with the superlative improvements in the design and operation of chemical plants over the last 50 years (Stephanopoulos and Reklaitis, 2011). Unfortunately, PSE's impact on process safety engineering has been lagging behind in ground-breaking contributions, and has focused mostly on monitoring and diagnostic methodologies. The reason is simple: The absence of a comprehensive framework that would allow the full deployment of powerful PSE methodologies. The accident-causation models used have been inadequate in providing the required comprehensive framework.

This situation can change, once the system-theoretic framework outlined in this Perspective is used in the design and operation of safer chemical processes. The array of interesting research topics that exist should be able to attract a large number of academic researchers and industrial practitioners.

References

American Institute of Chemical Engineers. Center for Chemical Process Safety, *Guidelines for developing quantitative safety risk criteria*, A John Wiley and Sons, New York, 2009.

Ackoff, R.L., "Towards a System of Systems Concepts", *Management Science* 17 (11):661-671, 1971

Ashby, W.R. *An Introduction to Cybernetics*, London: Chapman and Hall, 1956

Baker panel, *The Report of the BP U.S. Refineries Independent Safety Review Panel*, January 2007.

Calandranis, J., Nunokowa, S., Stephanopoulos, G., "DiAD-Kit/BOILER: On-Line Performance Monitoring and Diagnosis", *Chemical Engineering Progress*, p. 60-68 (January 1990).

Checkland, P. *Systems Thinking, Systems Practice*, New York: John Wiley & Sons, 1981

Cook, R.I. Verite, "Abstraction, and Ordinateur Systems in the Evolution of Complex Process Control", *3rd Annual Symposium on Human Interaction and Complex Sysems*, Dayton, Ohio, August 1996.

Erickson, C. A. II, *Fault Tree Analysis Primer*, 2011

Forrester, J. *Industrial Dynamics*, Waltham, MA: Pegasus Commuications, 1961

Hopkins, A., "Failure to Learn The BP Texas City Refinery Disaster", CCH Publishing, Australia, 2009.

Kletz, T. A., "Human Problems with Computer Control", *Plant/Operations Progress* 1(4), October 1982
Lederer, 1968

Kletz, T.A., *Hazop and Hazan*, Institution of Chemical Engineers, Rugby, UK, 1999.

Kletz, T. A., *Still going wrong! : case histories of process plant disasters and how they could have been avoided*, Elsevier Science & Technology Books, 2003

Lapp, S. A. and Powers, G. J., "Computer-aided Synthesis of Fault Trees", *IEEE Transactions on Reliability*, 26(1), p 2-13, 1977.

Leveson, N., *Safeware: System Safety and Computers*, Addison Wesley (1995).

Leveson, N., "A New Accident Model for Engineering Safer Systems", *Safety Science*, 2004, 42(4),

Leveson, N., Marais, K., Dulac, N., Carroll, J., "Moving Beyond Normal Accidents and High Reliability Organizations: An Alternative Approach to Safety in Complex Systems", *Organizational Studies* , Vol. 30, Feb/Mar 2009, Sage Publishers, pp. 227-249.

Leveson, N., "Applying Systems Thinking to Analyze and Learn from Events", *Safety Science*, 2010, 49(1), 55

Leveson, N., *Engineering a Safer World: Applying Systems Thinking to Safety*, MIT Press (January 2012).

Leveson, N., *An STPA Primer*, http://sunnyday.mit.edu/STPA-Primer-v0.pdf , 2013.

Nagel, C., Stephanopoulos, G., "Inductive and Deductive Reasoning: The Case of Identifying Potential Hazards in Chemical Processes", in *Intelligent Systems in Process Engineering: Paradigms from Design and Operations*, edited by C. Han and G. Stephanopoulos, *Academic Press*, 1996.

Olivea, C., O'Connora, T. M., and Mannan, M. S., "Relationship of safety culture and process safety", *J. Hazardous Materials*, 130(1-2), pp. 133-140, 2006

Ramo, S. The Systems Approach, in Systems Concepts: *Lectures on Contemporary Approaches to Systems*, ed. Ralph F. Miles, Jr, New York: John Wiley & Sonds, 1973

Rasmussen, J., "Risk Management in a Dynamic Society: A Modelling Problem", *Safety Science* 27(2/3): 183-213, 1997

Reason, J., "The contribution of latent human failures to the breakdown of complex systems", Philosophical Transactions of the Royal Society (London), series B. 327: 475-484, 1990.

Stephanopoulos, G., Reklaitis, G.V., "Process Systems Engineering: From Solvay to Modern Bio- and Nanotechnology. A History of Development, Successes and Prospects for the Future", *Chemical Engineering Science*, 66, 4272–4306, 2011.

Suokas, J. "Evaluation of the Quality of safety and risk analysis in the chemical industry", *Risk Analysis*, 8(4):581-591, 1985

Urbina, I., "Inspector General's Inquiry Faults Regulators", *New York Times*, May 24, 2010

Venkatasubramanian, V., Zhao, J. and Viswanathan, S., "Intelligent Systems for HAZOP Analysis of Complex Process Plants", *Computers and Chemical Engineering*, 24 (9-10), 2000, pp. 2291 – 2302.

Venkatasubramanian, V., Rengaswamy, R., Yin, K., "A review of process fault detection and diagnosis: Part I: Quantitative model-based methods", *Computers & Chemical Engineering*, 27, 293–311, 2003a.

Venkatasubramanian, V., Rengaswamy, R., Kavuri, S. N., "A review of process fault detection and diagnosis: Part II: Qualitative models and search strategies", *Computers & Chemical Engineering*, 27, 313-326, 2003b.

Venkatasubramanian, V., Rengaswamy, R., Kavuri, S. N., Yin, K., "A review of process fault detection and diagnosis: Part III: Process history based methods", *Computers & Chemical Engineering*, 27, 327-346, 2003c.

Venkatasubramanian, V., "Systemic Failures: Challenges and Opportunities in Risk Management in Complex Systems", *AIChE J.*, ***57,*** p. 2, 2011