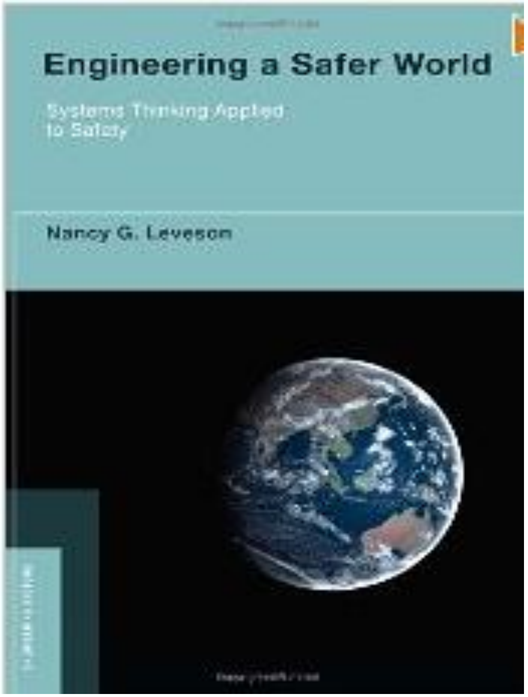
Welcome to the

STAMP/STPA “Workshop”



Engineering a Safer World

Systems Thinking Applied to Safety

Nancy G. Leveson

Introduction

- Attendance:
  - Nearly 250 attendees
  - From 19 countries
  - And nearly every industry
- Sponsored by
  - Engineering Systems Division,
  - Aeronautics and Astronautics Department
  - Industrial Liaison Program.

Outline

1. The Problem

2. STAMP: A New Accident Model

3. STPA: A New Hazard Analysis Technique Built on STAMP

4. CAST: Structured Accident Analysis

The Problem

The first step in solving any problem is to understand it.

We often propose solutions to problems that we do not understand and then are surprised when the solutions fail to have the anticipated effect.

Why need a new approach?

"Without changing our patterns of thought, we will not be able to solve the problems we created with our current patterns of thought."

Albert Einstein

- Traditional safety engineering approaches developed for relatively simple electro-mechanical systems
- Accidents in complex, software-intensive systems are changing their nature
- Role of humans in systems is changing
- We need more effective techniques for these new systems

Changes in the Last 50 Years

- Use of software has created new causes of accidents
- Role of humans in systems and in accidents has changed
- Increased recognition of importance of management and social factors in accidents
- Fast pace of technological change
  - Learning from experience (“fly-fix-fly”) no longer as effective
  - Introduces “unknowns” and new paths to accidents
  - Faster time to market means less testing and analysis
- Increasing complexity
- Decreasing tolerance for single accidents

The Starting Point: Questioning Our Assumptions

“It’s never what we don’t know that stops us, it’s what we do know that just ain’t so.”

(Attributed to many people)

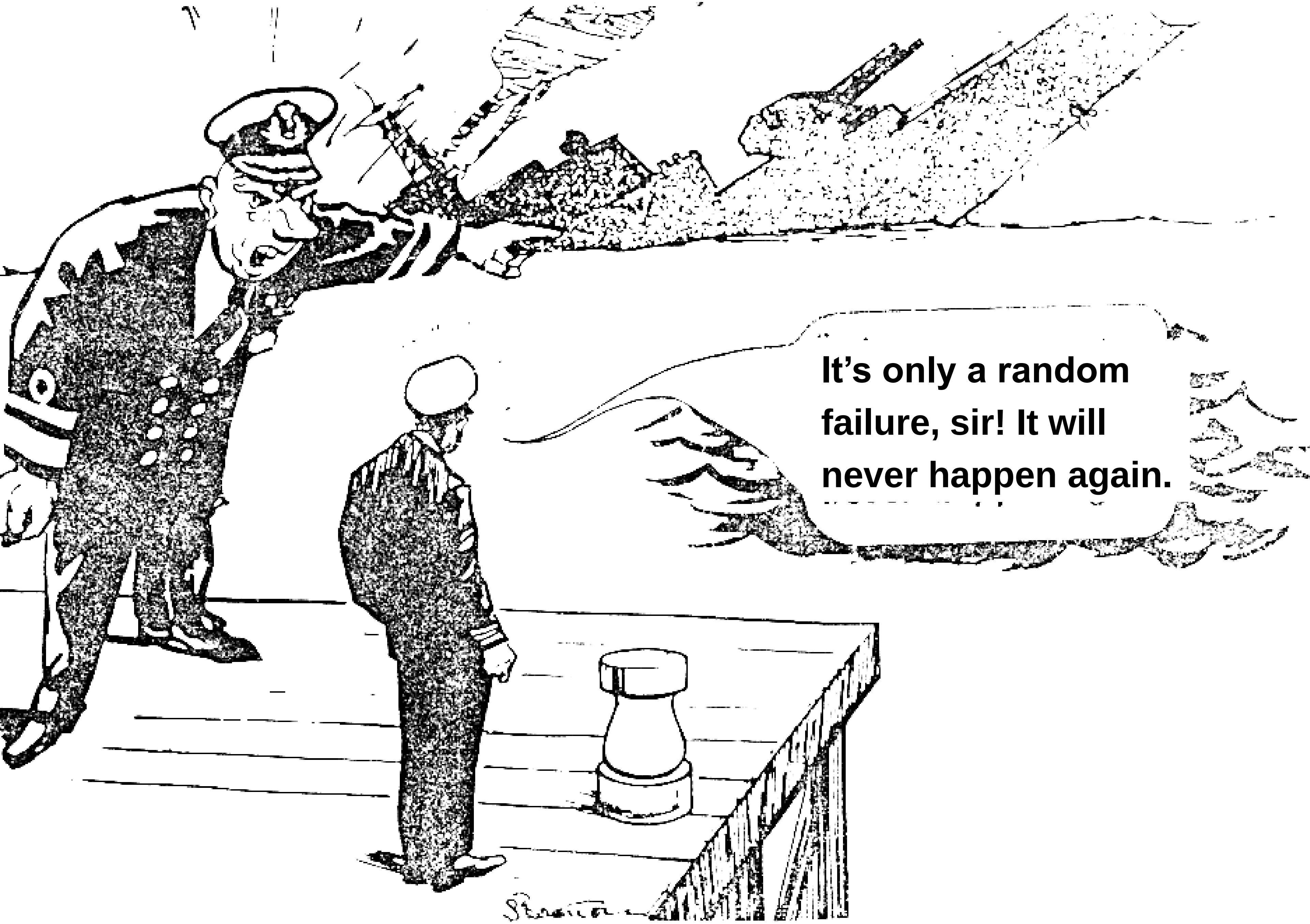
What are some of the things we know about safety that just ain’t so?

Assumption 1

- Accidents are caused by component failures.
- Therefore, safety is increased by reducing component failures (i.e., increasing reliability)
- If components don’t fail, accidents will not occur

Is This True?

- Many accidents occur without any component “failure”
  - Caused by equipment operation outside parameters and time limits upon which reliability analyses are based.
  - Caused by interactions of components all operating according to specification.
- Highly reliable components are not necessarily safe

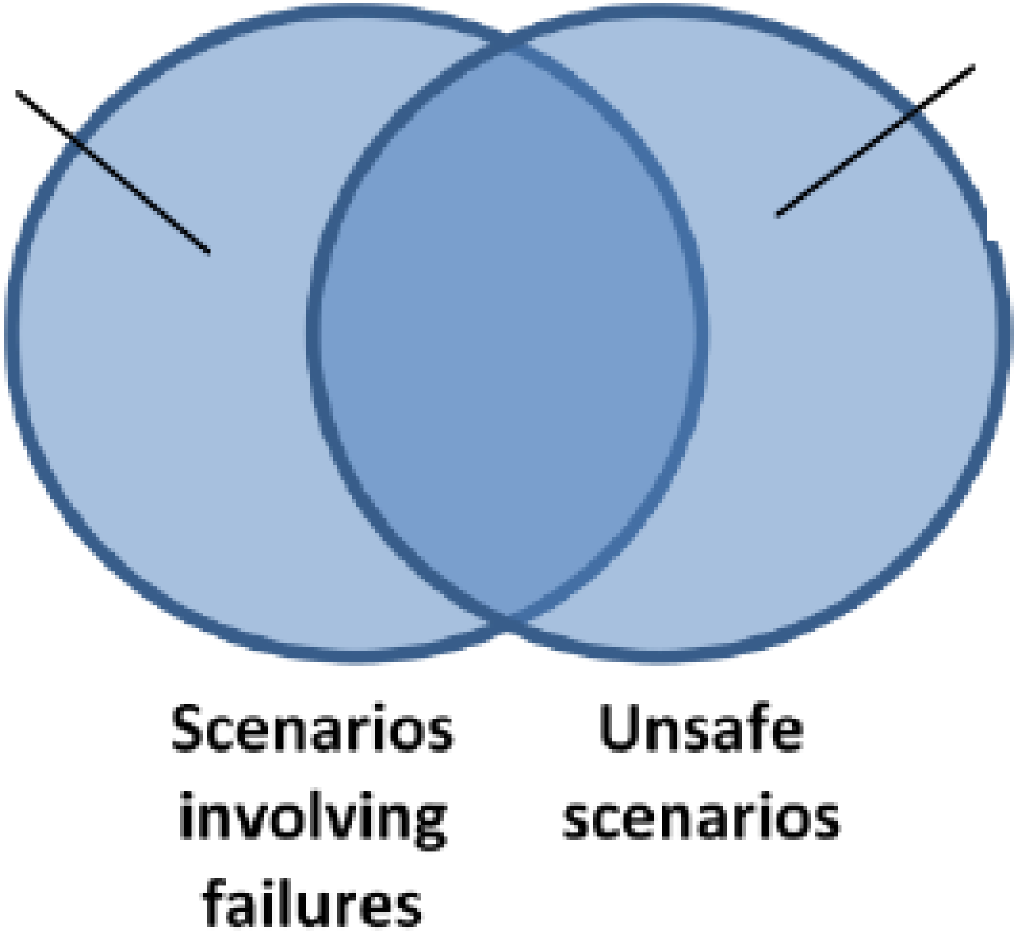

It’s only a random failure, sir! It will never happen again.

Types of Accidents

- Component Failure Accidents
  - Single or multiple component failures
  - Usually assume random failure
- Component Interaction Accidents
  - Arise in interactions among components
  - Components may not have “failed”
  - Exacerbated by introduction of computers and complexity

Interactive Complexity

- Critical factor is intellectual manageability
    - A simple system has a small number of unknowns in its interactions (within system and with environment)
    - Interactively complex (intellectually unmanageable) when level of interactions reaches point where can no longer be thoroughly
        - Planned
        - Understood
        - Anticipated
        - Guarded against


Scenarios involving failures
Unsafe scenarios

Assumption 1

- ~~Accidents are caused by component failures.~~
- ~~Therefore, safety is increased by reducing component failures~~
- ~~If components don’t fail, accidents will not occur~~

- High component reliability is neither necessary nor sufficient for safety.

Assumption 1b

- Highly reliable software is safe.

Software-Related Accidents

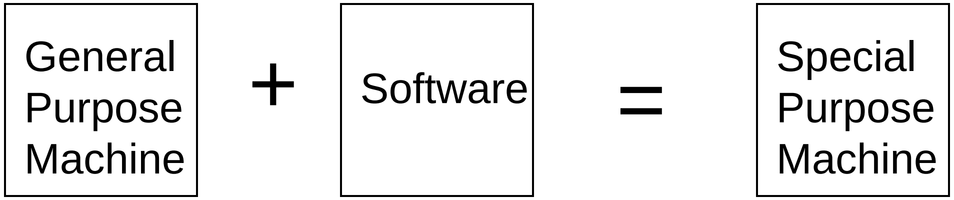
- Are usually caused by flawed requirements
  - Incomplete or wrong assumptions about operation of controlled system or required operation of computer
  - Unhandled controlled-system states and environmental conditions

- Merely trying to get the software “correct” or to make it reliable will not make it safer under these conditions.

Software-Related Accidents (2)

- Software may be highly reliable and “correct” and still be unsafe:
  - Correctly implements requirements but specified behavior unsafe from a system perspective.
  - Requirements do not specify some particular behavior required for system safety (incomplete)
  - Software has unintended (and unsafe) behavior beyond what is specified in requirements.

The Computer Revolution

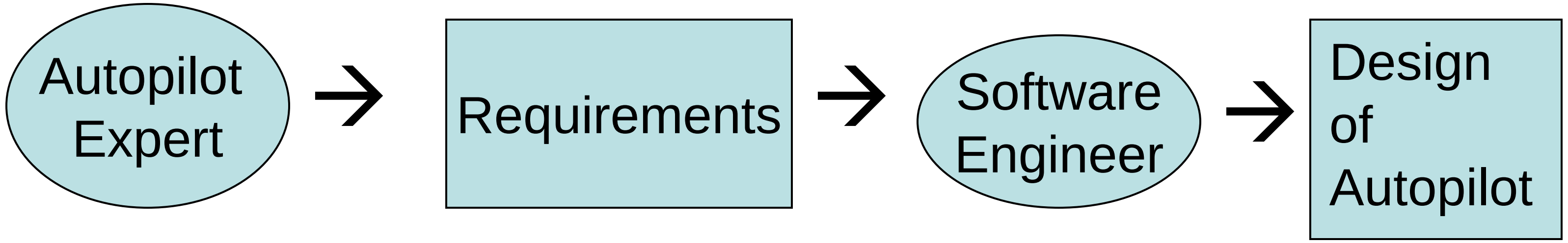


General Purpose Machine + Software = Special Purpose Machine

- Software is simply the design of a machine abstracted from its physical realization
- Machines that were physically impossible or impractical to build become feasible
- Design can be changed without retooling or manufacturing
- Can concentrate on steps to be achieved without worrying about how steps will be realized physically

Abstraction from Physical Design

- Software engineers are doing physical design



- Most operational software errors related to requirements (particularly incompleteness)
- Software "failure modes" are different
  - Usually does exactly what you tell it to do
  - Problems occur from operation, not lack of operation
  - Usually doing exactly what software engineers wanted

Safety vs. Correctness

- Safety involves more than simply getting the software "correct":

  Example: altitude switch

  1. Signal safety-increasing ➔

     Require any of three altimeters report below threshold

  2. Signal safety-decreasing ➔

     Require all three altimeters to report below threshold

- Software is very different from hardware.
- We cannot just apply techniques developed for hardware and expect them to work.
- We need something new that fits software properties.

Assumption 1b

- ~~Highly reliable software is safe.~~

- Highly reliable software (correctly implements its requirements) is not necessarily safe

- Increasing software reliability (correctness) will have only minimal impact on system safety

Assumption 2

- Accidents are caused by chains of failure events.
- We can understand accidents and assess risk by looking only at the direct relationships between the events leading to the loss

Jerome Lederer (1968)

“Systems safety covers the total spectrum of risk management. It goes *beyond the hardware* and associated procedures of systems 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.”

Direct Causality No Longer Adequate to Understand Accidents

- **Interactive Complexity**: Arises in complex and indirect interactions among system components
- **Non-linear complexity**: Cause and effect not related in an obvious way
- **Dynamic complexity**: Related to changes over time
- **Decompositional complexity**: Related to how decompose or modularize our systems
- Others ??

Assumption 2

- ~~Accidents are caused by chains of directly related failure events.~~
- ~~We can understand accidents and assess risk by looking at the chains of events leading to the loss~~

- Accidents are complex processes involving the entire socio-technical system.
- Traditional event-chain models cannot describe this process adequately

Assumption 3

- Most accidents are caused by operator error.
- Better training, rewarding good behavior and punishing bad behavior will eliminate accidents or reduce them significantly.

Human Error: Traditional View

- Operator error is cause of most incidents and accidents
- So do something about human involved (fire them, retrain, admonish)
- Or do something about humans in general
  - Marginalize them by putting in more automation
  - Rigidify their work by creating more rules and procedures

Human Error: New View
(Sydney Dekker, Jens Rasmussen, etc.)

- Human error is a symptom, not a cause
- All behavior affected by context (system) in which occurs
- Role of operators in our systems is changing
  - Supervising rather than directly controlling
  - Systems are stretching limits of comprehensibility
  - Designing systems in which operator error inevitable and then blame accidents on operators rather than designers
- To do something about error, must look at system in which people work:
  - Design of equipment
  - Usefulness of procedures
  - Existence of goal conflicts and production pressures

Cali American Airlines Crash

Cited probable causes:

- Flight crew’s failure to adequately plan and execute the approach to runway 10 at Cali and their inadequate use of automation
- Failure of flight crew to discontinue the approach into Cali, despite numerous cues alerting them of the inadvisability of continuing the approach
- Lack of situational awareness of the flight crew regarding vertical navigation, proximity to terrain, and the relative location of critical radio aids.
- Failure of the flight crew to revert to basic radio navigation at the time when the FMS-assisted navigation became confusing and demanded an excessive workload in a critical phase of flight.

Assumption 3

- ~~Most accidents are caused by operator error.~~
- ~~Better training, rewarding good behavior and punishing bad behavior will eliminate accidents or reduce them significantly.~~

- Operator error is a product of the environment in which it occurs.
- To reduce operator “error” we must change the environment in which the operator works.

Assumption 4

- Probabilistic risk analysis based on event chains is the best (only?) way to assess and communicate about safety

Assumption 4

- ~~Probabilistic risk analysis based on event chains is the best (only?) way to assess and communicate about safety~~

- Risk and safety may be best understood and communicated in ways other than probabilistic risk analysis.

Assumption 5

- Most accidents occur from the chance simultaneous occurrence of random events

Evolution and Adaptation

- Most major accidents arise from a slow migration of the entire system toward a state of high-risk (Jens Rasmussen)
  - A socio-technical system is a dynamic process continually adapting to achieve its ends and to react to changes in itself and its environment
  - Systems and organizations migrate toward accidents (states of high risk) under cost and productivity pressures in an aggressive, competitive environment
  - Need to control and detect this migration

Assumption 5

- ~~Most accidents occur from the chance simultaneous occurrence of random events~~

- Systems tend to migrate toward states of higher risk
- Hypothesis:
  - Such migration is predictable and hazardous changes can either be
    - Prevented by appropriate system design and management of change procedures or
    - Detected during operations using leading indicators of increasing risk

Assumption 6

- Assigning blame is necessary to learn from and prevent accidents or incidents.
- If we can identify the “root cause,” then we can prevent future accidents.

Impediments to Learning from Accidents and Incidents

- Filtering and subjectivity in accident reports
- “Blame is the enemy of safety”
  - Focus on “who” and not “why”
- “Root cause” seduction
  - Believing in a “root cause” appeals to our desire for control
  - Leads to a sophisticated “whack a mole” game
  - Fix symptoms but not process that led to loss
  - Same accident happening over and over again

Impediments to Learning (2)

- Oversimplification
- Almost always there is:
  - Operator “error”
  - Flawed management decision making
  - Flaws in the physical design of equipment
  - Safety culture problems
  - Regulatory deficiencies
  - Etc.

Three Levels of Analysis

- What (events)
  - e.g., explosion
- Who and how (conditions)
  - e.g., bad valve design, operator did not notice something
- Why (systemic factors)
  - e.g., production pressures, cost concerns, flaws in design process, flaws in reporting process, etc.
  - Why was safety control structure ineffective in preventing the loss?

Assumption 6

- ~~Assigning blame is necessary to learn from and prevent accidents or incidents.~~

- Blame is the enemy of safety.

- Focus should be on understanding how the system behavior as a whole contributed to the loss and not on who or what to blame for it.

So What Do We Need to Do?
"Engineering a Safer World"

- Expand our accident causation models
- Create new, more powerful and inclusive hazard analysis techniques
- Use new system design techniques
  - Safety-driven design
  - Improved system engineering
- Improve accident analysis and learning from events
- Improve control of safety during operations
- Improve management decision-making and safety culture

Accident Causality Models

- Underlie all our efforts to engineer for safety
- Explain why accidents occur
- Determine the way we prevent and investigate accidents
- May not be aware you are using one, but you are
- Imposes patterns on accidents

“All models are wrong, some models are useful”

George Box

Chain-of-Events Model

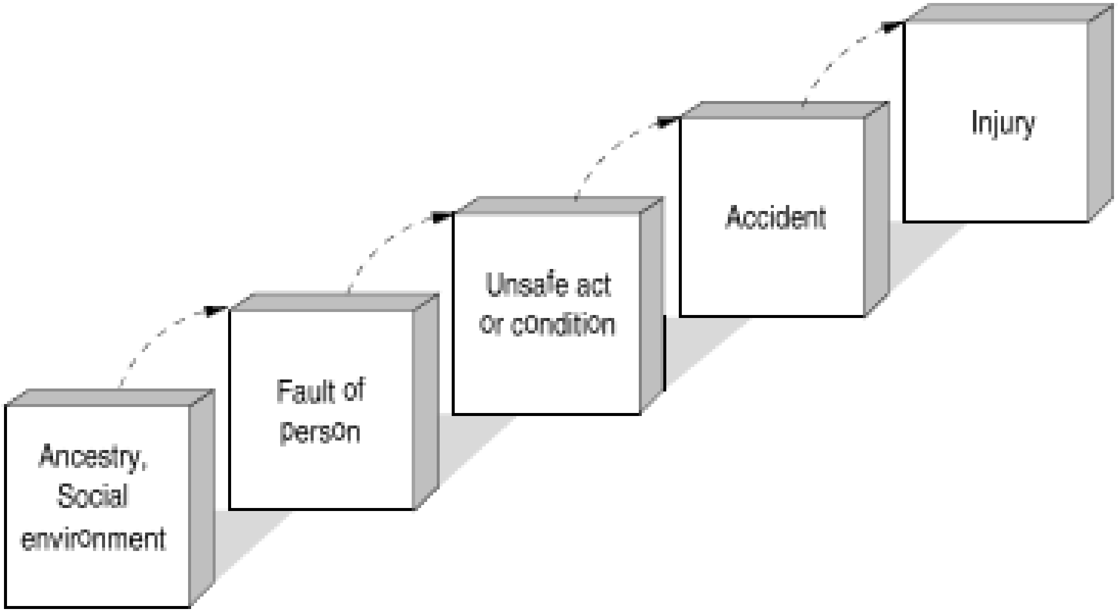
- Explains accidents in terms of multiple events, sequenced as a forward chain over time.
  - Simple, direct relationship between events in chain
- Events almost always involve component failure, human error, or energy-related event
- Forms the basis for most safety engineering and reliability engineering analysis:

  e,g, FTA, PRA, FMECA, Event Trees, etc.

  and design:

  e.g., redundancy, overdesign, safety margins, ….

Heinrich’s Domino Model (1931)



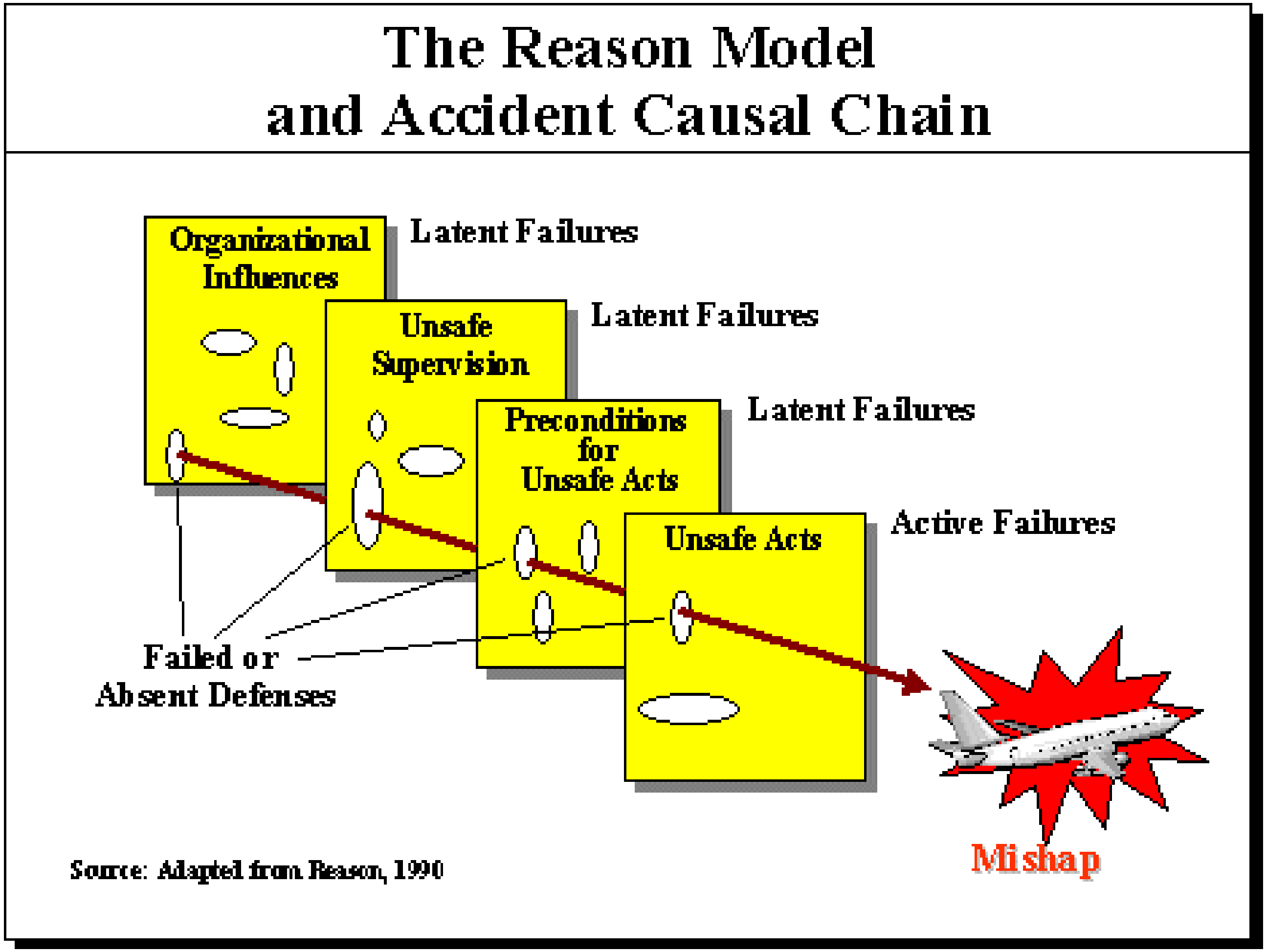
Note: focus on direct causality and human error

The Domino Model in action

Variants of Domino Model

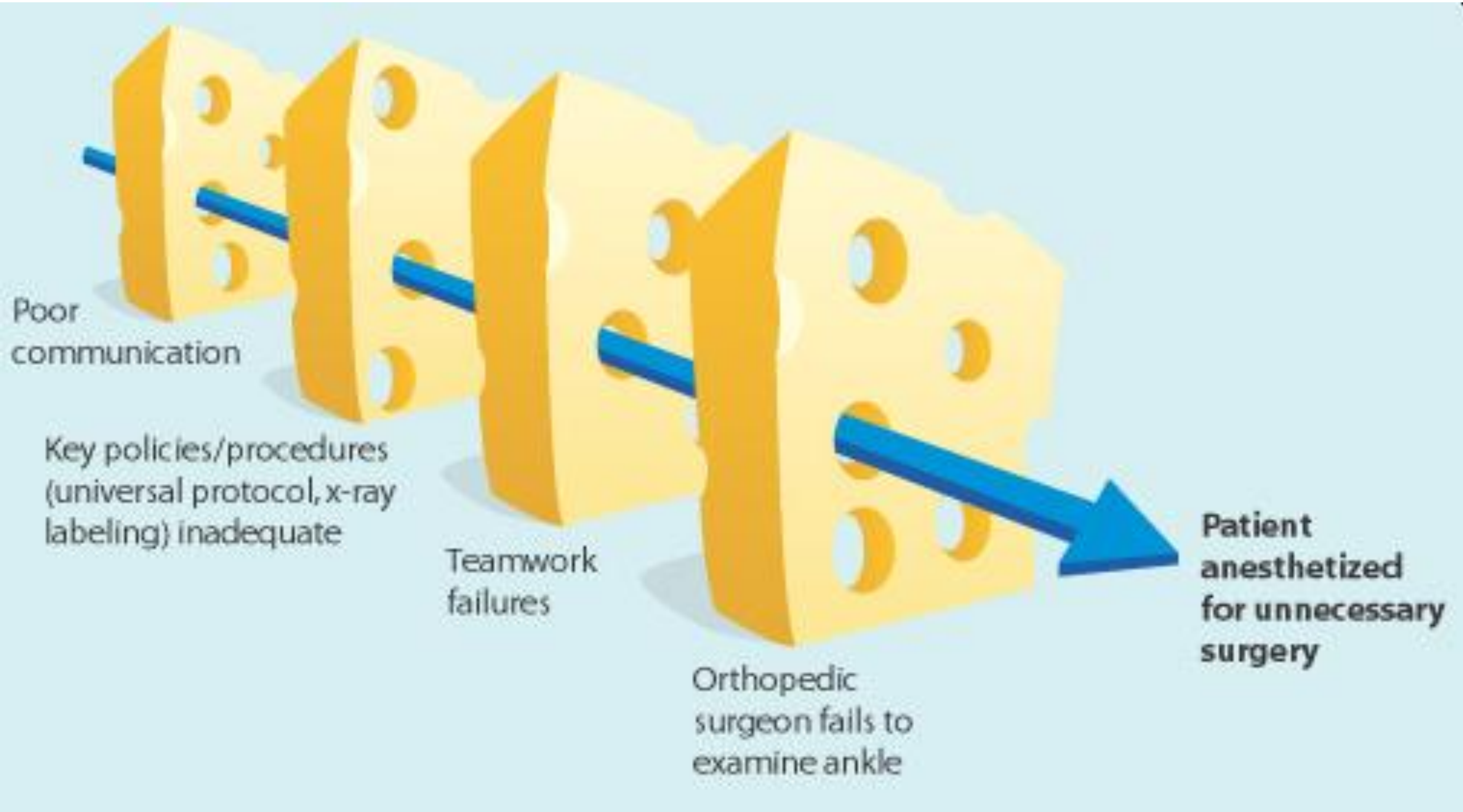
- Bird and Loftus (1976)
  - Lack of control by management, permitting
  - Basic causes (personal and job factors) that lead to
  - Immediate causes (substandard practices/conditions/errors), which are the proximate cause of
  - An accident or incident, which results in
  - A loss.
- Adams (1976)
  - Management structure (objectives, organization, and operations)
  - Operational errors (management or supervisor behavior)
  - Tactical errors (caused by employee behavior and work conditions)
  - Accident or incident
  - Injury or damage to persons or property.

Reason Swiss Cheese



The Reason Model
and Accident Causal Chain

Source: Adapted from Reason, 1990



Poor communication

Key policies/procedures (universal protocol, x-ray labeling) inadequate

Teamwork failures

Orthopedic surgeon fails to examine ankle

Patient anesthetized for unnecessary surgery

Swiss Cheese Model Limitations

- Ignores common cause failures of defenses (systemic accident factors)

- Does not include migration to states of high risk: an alternative is the “Mickey Mouse Model”

- Assumes accidents are random events coming together accidentally

  “High-consequence, low probability events”

- Assumes some (linear) causality or precedence in the cheese slices.

Limitations of Chain-of-Events Causation Models

- Oversimplifies causality
- Excludes or does not handle
  - Component interaction accidents (vs. component failure accidents)
  - Indirect or non-linear interactions and complexity
  - Systemic factors in accidents
  - Human “errors”
  - System design errors (including software errors)
  - Adaptation and migration toward states of increasing risk

STAMP
(System-Theoretic Accident Model and Processes)

- A new, more powerful accident causation model
- Based on systems theory, not reliability theory
- Treats accidents as a dynamic control problem (vs. a failure problem)
- Includes
  - Entire socio-technical system (not just technical part)
  - Component interaction accidents
  - Software and system design errors
  - Human errors

Safety as a Control Problem

- Safety is an emergent property that arises when system components interact with each other within a larger environment
  - A set of constraints related to behavior of system components (physical, human, social) enforces that property
  - Accidents occur when interactions violate those constraints (a lack of appropriate constraints on the interactions)
- Goal is to control the behavior of the components and systems as a whole to ensure safety constraints are enforced in the operating system.

Safety as a Control Problem (2)

- Accidents are not simply an event or chain of events but involve a complex, dynamic process
- Events are the result of the inadequate control
  - Result from lack of enforcement of safety constraints in system design and operations
  - Migration of systems to states of higher risk
- A change in emphasis:

~~“prevent failures”~~

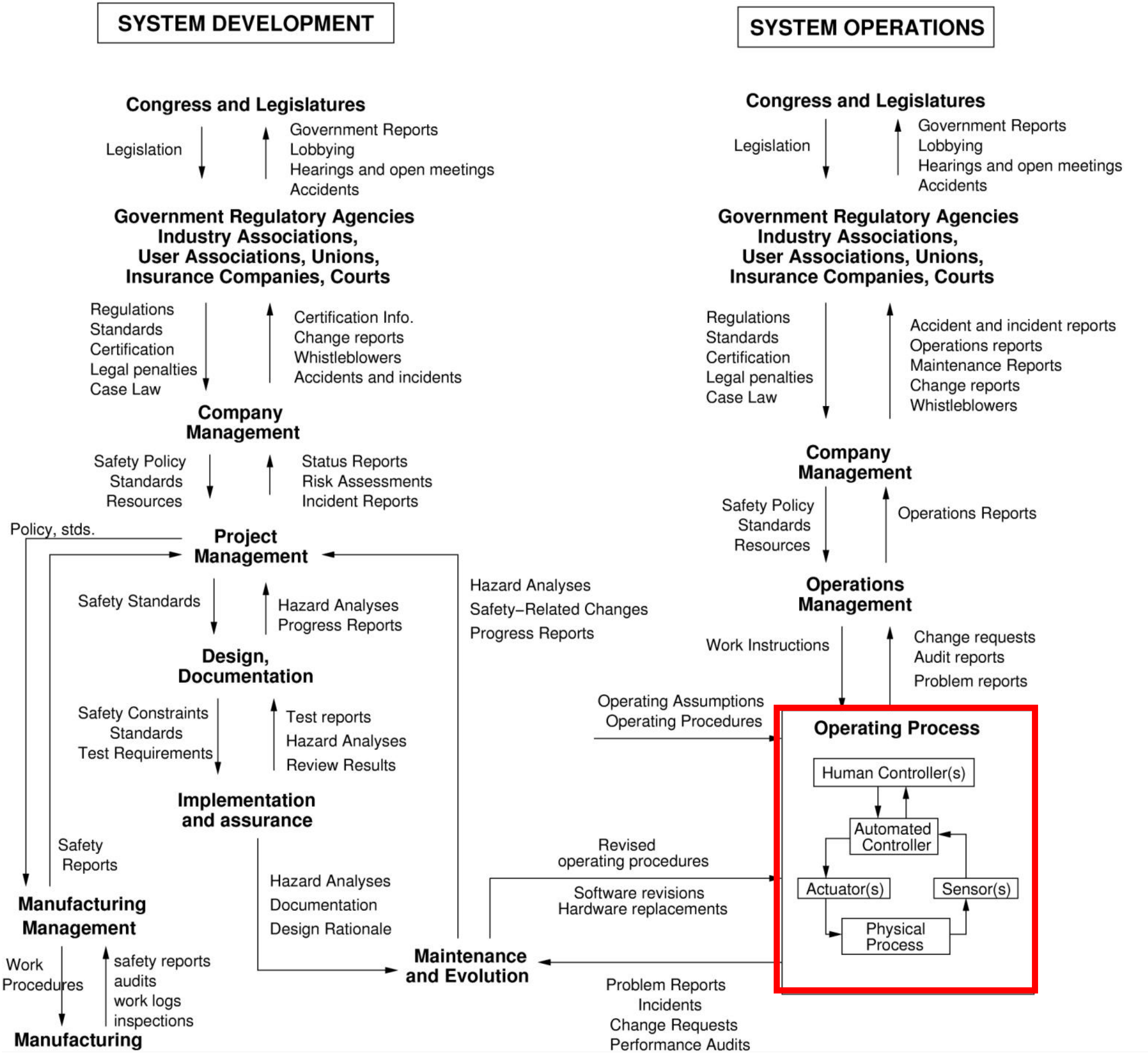
↓

“enforce safety constraints on system behavior”

STAMP

- Treat safety as a dynamic control problem rather than a component failure problem.
  - O-ring did not control propellant gas release by sealing gap in field joint of Challenger Space Shuttle
  - Software did not adequately control descent speed of Mars Polar Lander
  - Temperature in batch reactor not adequately controlled in system design
  - Public health system did not adequately control contamination of the milk supply with melamine
  - Financial system did not adequately control the use of financial instruments
- Events are the <u>result</u> of the inadequate control
  - Result from lack of enforcement of safety constraints in system design and operations

Example Safety Control Structure



SYSTEM DEVELOPMENT

Congress and Legislatures

Legislation

Government Reports
Lobbying
Hearings and open meetings
Accidents

Government Regulatory Agencies
Industry Associations,
User Associations, Unions,
Insurance Companies, Courts

Regulations
Standards
Certification
Legal penalties
Case Law

Certification Info.
Change reports
Whistleblowers
Accidents and incidents

Company Management

Safety Policy
Standards
Resources

Status Reports
Risk Assessments
Incident Reports

Policy, stds.

Project Management

Safety Standards

Hazard Analyses
Progress Reports

Hazard Analyses
Safety–Related Changes
Progress Reports

Design, Documentation

Safety Constraints
Standards
Test Requirements

Test reports
Hazard Analyses
Review Results

Implementation and assurance

Safety Reports

Hazard Analyses
Documentation
Design Rationale

Manufacturing Management

Work Procedures

safety reports
audits
work logs
inspections

Manufacturing

Maintenance and Evolution

SYSTEM OPERATIONS

Congress and Legislatures

Legislation

Government Reports
Lobbying
Hearings and open meetings
Accidents

Government Regulatory Agencies
Industry Associations,
User Associations, Unions,
Insurance Companies, Courts

Regulations
Standards
Certification
Legal penalties
Case Law

Accident and incident reports
Operations reports
Maintenance Reports
Change reports
Whistleblowers

Company Management

Safety Policy
Standards
Resources

Operations Reports

Operations Management

Work Instructions

Change requests
Audit reports
Problem reports

Operating Assumptions
Operating Procedures

Operating Process

Human Controller(s)

Automated Controller

Actuator(s)

Sensor(s)

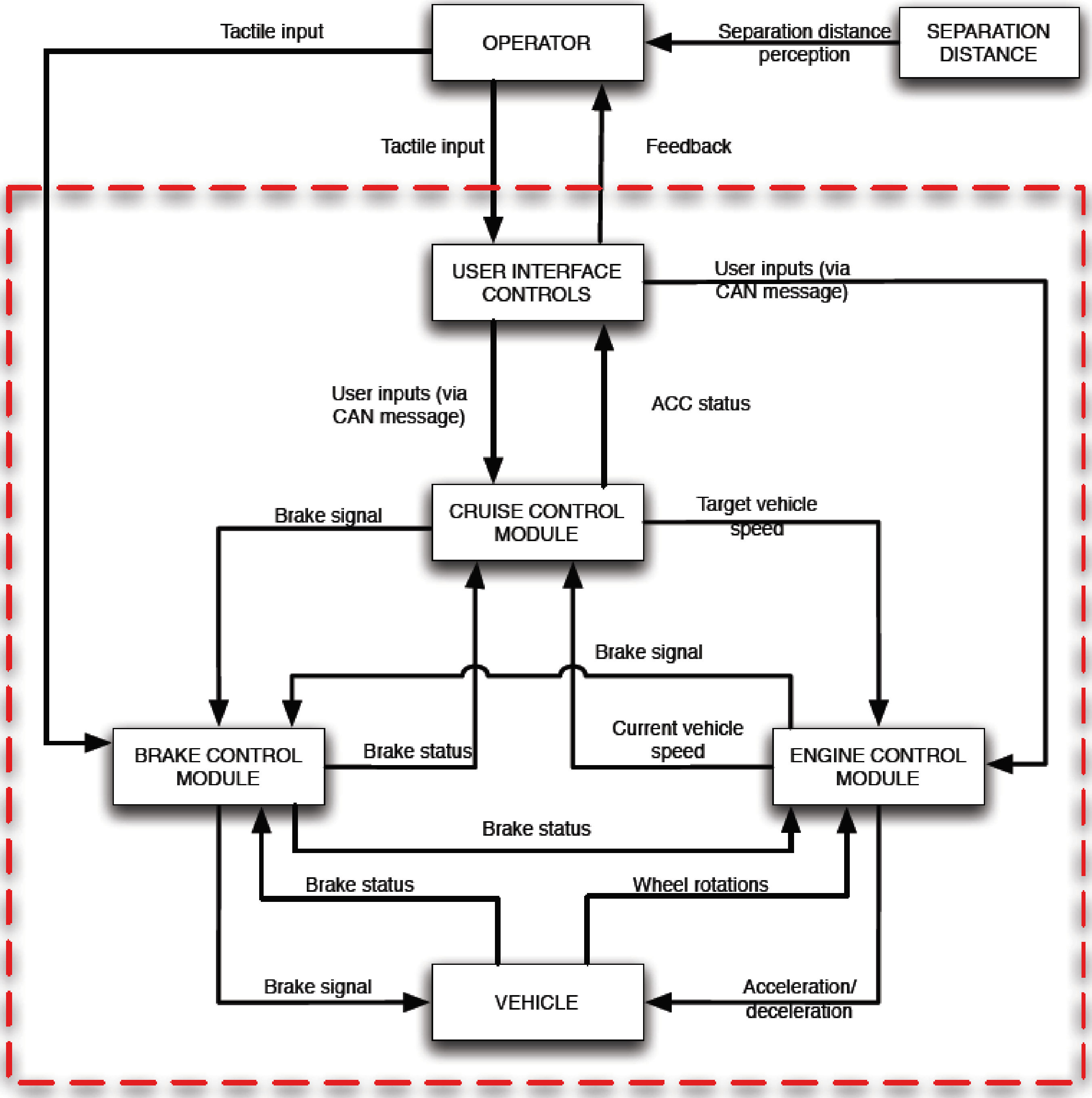
Physical Process

Revised operating procedures

Software revisions
Hardware replacements

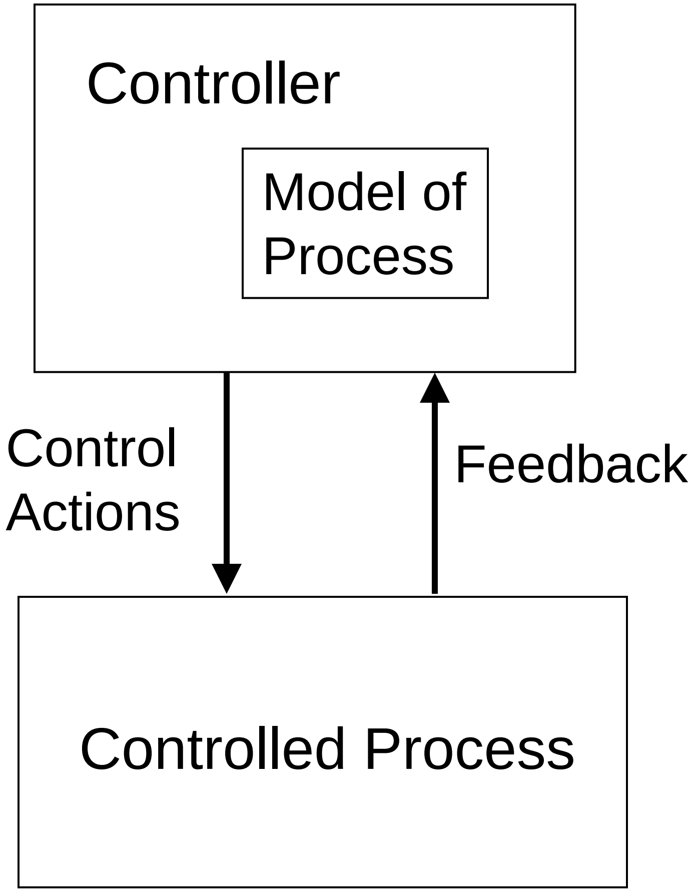
Problem Reports
Incidents
Change Requests
Performance Audits

LEVEL 1



OPERATOR
SEPARATION DISTANCE
Separation distance perception
Tactile input
Tactile input
Feedback
USER INTERFACE CONTROLS
User inputs (via CAN message)
User inputs (via CAN message)
ACC status
CRUISE CONTROL MODULE
Brake signal
Target vehicle speed
Brake signal
BRAKE CONTROL MODULE
Brake status
Current vehicle speed
ENGINE CONTROL MODULE
Brake status
Brake status
Wheel rotations
Brake signal
VEHICLE
Acceleration/ deceleration

Control processes operate between levels of control



Accidents occur when model of process is inconsistent with real state of process and controller provides inadequate control actions

Feedback channels are critical
- -- Design
- -- Operation

Relationship Between Safety and Process Models

- How do they become inconsistent?
  - Wrong from beginning
  - Missing or incorrect feedback
  - Not updated correctly
  - Time lags not accounted for

  Resulting in

  Uncontrolled disturbances

  Unhandled process states

  Inadvertently commanding system into a hazardous state

  Unhandled or incorrectly handled system component failures

Relationship Between Safety and Process Models (2)

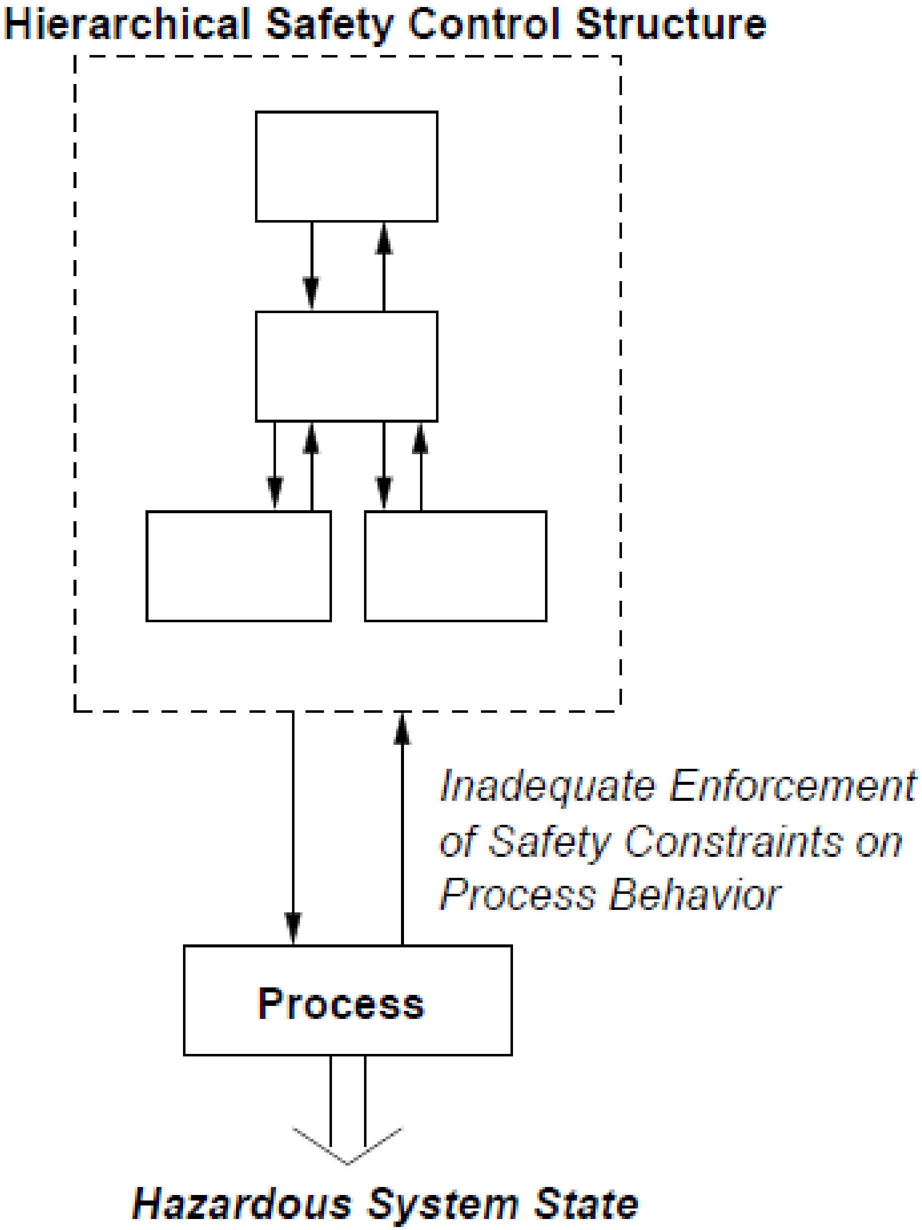
- Accidents occur when models do not match process and
  - Required control commands are not given
  - Incorrect (unsafe) ones are given
  - Correct commands given at wrong time (too early, too late)
  - Control action stops too soon or applied too long

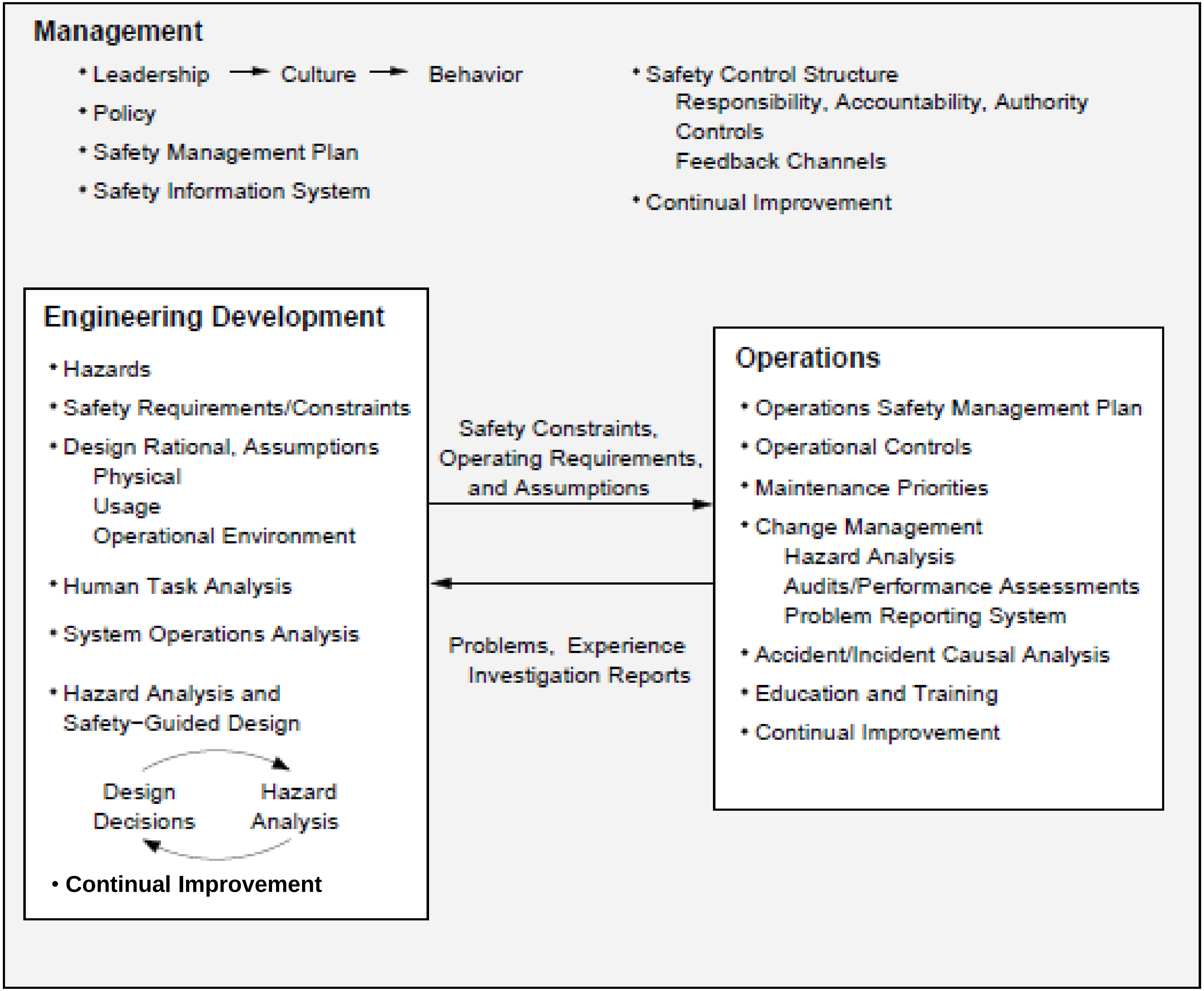
Explains software errors, human errors, component interaction accidents …

Summary: Accident Causality in STAMP

- Accidents occur when
  - Control structure or control actions do not enforce safety constraints
    - Unhandled environmental disturbances or conditions
    - Unhandled or uncontrolled component failures
    - Dysfunctional (unsafe) interactions among components
  - Control structure degrades over time (asynchronous evolution)
  - Control actions inadequately coordinated among multiple controllers

Accident Causality Using STAMP


Hierarchical Safety Control Structure
Inadequate Enforcement of Safety Constraints on Process Behavior
Process
Hazardous System State



Management

- Leadership → Culture → Behavior
- Policy
- Safety Management Plan
- Safety Information System
- Safety Control Structure
  Responsibility, Accountability, Authority
  Controls
  Feedback Channels
- Continual Improvement

Engineering Development

- Hazards
- Safety Requirements/Constraints
- Design Rational, Assumptions
  Physical
  Usage
  Operational Environment
- Human Task Analysis
- System Operations Analysis
- Hazard Analysis and Safety−Guided Design

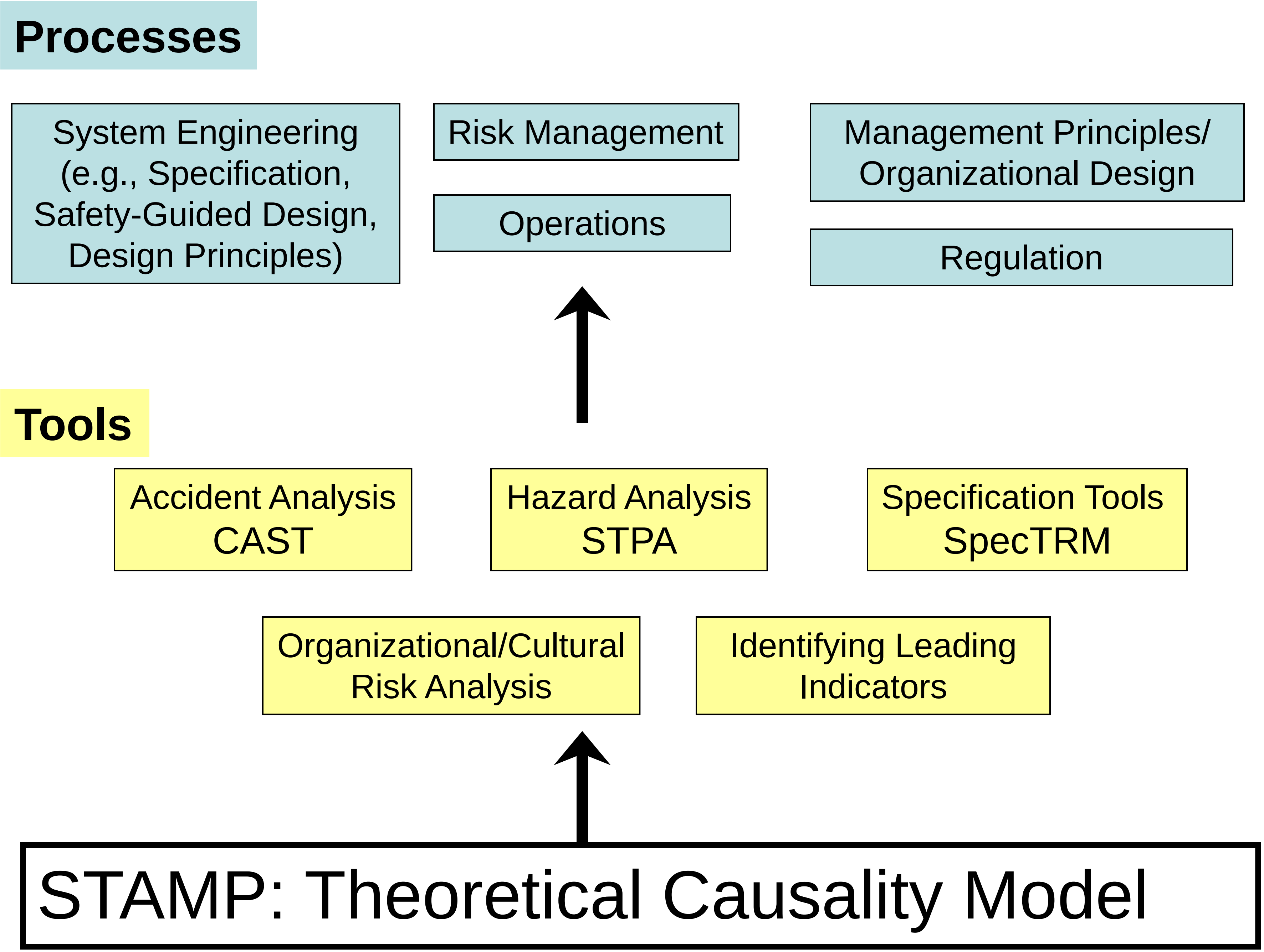
- **Continual Improvement**

Safety Constraints, Operating Requirements, and Assumptions

Problems, Experience Investigation Reports

Operations

- Operations Safety Management Plan
- Operational Controls
- Maintenance Priorities
- Change Management
  Hazard Analysis
  Audits/Performance Assessments
  Problem Reporting System
- Accident/Incident Causal Analysis
- Education and Training
- Continual Improvement



Processes

System Engineering (e.g., Specification, Safety-Guided Design, Design Principles)

Risk Management

Operations

Management Principles/ Organizational Design

Regulation

Tools

Accident Analysis CAST

Hazard Analysis STPA

Specification Tools SpecTRM

Organizational/Cultural Risk Analysis

Identifying Leading Indicators

STAMP: Theoretical Causality Model

PSAS: Partnership for Systems Approaches to Safety

- Evaluate current practices and potential new ones
- Solve real problems, not just abstract or theoretical ones.
  - Suggested by and supported by industrial and governmental partners.
  - Mentoring and internships by graduate students in industry and government
- Newsletters and other information dissemination channels about activities, results, etc., including early access to thesis abstracts and results.
- Sponsored research

PSAS: Partnership for Systems Approaches to Safety

- Educational activities including short classes and workshops for PSAS partners.
- Knowledge and information sharing
- Annual conference
- Visitors from industry, government, and other research institutions
- Collaboration with like-minded researchers around the world
- Take a global perspective

Educational Initiatives

- New system safety track in the ESD master’s degree
- System safety emphasis possible in ESD and Aero/Astro Ph.D. programs
- Professional master’s programs participate in PSAS projects
- New undergraduate class on system safety
- Industry classes and continuing education

Faculty

- Prof. Nancy Leveson (Aero/Astro and ESD)
- Prof. Joseph Sussman (Civil Engineering and ESD)
- Prof. John Carroll (Sloan School of Management and ESD)
- Dr. Qi Hommes (ESD)

Current Research in PSAS

- **Aviation**:
  - Certification of safety in NextGen (NASA Aviation Safety Program)
  - Certification of IMA (Integrated Modular Avionics): (with Embraer engineers and FAA, NASA)

- **Spacecraft (JAXA)**:
  - Evaluation of STPA on the HTV
  - Design for safety of a NASA/JAXA scientific satellite
  - Using STPA in early architectural trades for the planned JAXA Crew Vehicle

Current Research Projects

- **Healthcare**:
  - A Systems Theoretic Application to Design for the Safety of Medical Diagnostic Devices.
  - Quality Control in Medical Manufacturing
  - The Role of Culture/Social/Legal Systems on Medical Device Safety in China
  - Safety Certification of Digital-Intense Systems in Radiation Therapy (PSI)
  - Learning from Safety-Relevant Events in Hospitals: The Role of Mental Models

Current Research Projects

- **Nuclear Power Plants**
  - Certification of digital shutdown systems in NPPs (NRC)
- **Automobiles**
  - Using STPA to Analyze the Safety of Electronic Throttle Control Systems
  - Applying STPA to Adaptive Cruise Control
- **Oil and Gas (Petrochemicals) and Energy**
  - Developing Leading Indicators for Process Safety
  - Power Plant Gas Turbine Accident Investigation in China

Current Research Projects

- **Defense**
  - Coast Guard Helicopter Night Rescue Training Accident Investigation
  - Prevention of fratricide in the Patriot Missile System
  - A Systems Approach to Cyber Security
- **Railroads**
  - Application of CAST and STPA to Railroad Safety in China

Current Research Projects

- **General**
  - Corporate Governance and Management Decision Making about Safety
  - System Engineering Aspects of Safety
  - Applying STAMP for Automation Decision Making in a Manufacturing Plant Quality Inspection Station (Continental Tires and the MIT Portugal Program)
  - Integrating Safety into ILF’s System Engineering process using the guidelines of STAMP (ILF and Heriot Watt University, Edinburgh)
  - Using STAMP to Understand the Recent Financial Crisis