A Systems Approach to Risk Management Through Leading Indicators

Nancy Leveson

MIT

Goal

- To identify potential for an accident before it occurs
- Underlying assumption:
  - Major accidents not due to a unique set of random, proximal events
  - Instead result from
    - Migration of organization to state of heightened risk over time
    - As safeguards and controls relaxed over time
    - Due to conflicting goals and tradeoffs
- Detection not enough---need to embed in a risk management program

Current State of the Art: Industry

- Much effort, particularly in petrochemicals
  - Trying to find generally applicable indicators

    (e.g., maintenance backlogs, minor incidents, equipment failure rates, surveys on employee culture (assumes that all or most accidents caused by employee misbehavior)
  - Tend to focus on occupational safety
  - Identification of leading indicators from hazard analysis
    - Use standard techniques so limited types of causes
    - Use likelihood to reduce scope of search
      - May result in overlooking low likelihood events

Current State of the Art: Research

- Identify precursors to accidents
  - Use incident reporting system (only look for things that have happened before)
  - Use probabilistic risk analysis
  - Hazard analysis (but limited scenarios)
  - Organizational precursors
    - Identify small number of factors (5 or 6)
    - Quantitative risk analysis (fault trees, Bayesian analysis)
    - Lack a model of framework on how and why accidents occur

Current State of the Art: PRA

- Risk and Risk Assessment
  - Little data validating PRA or methods for calculating it
  - Other problems
    - May be significant divergence between modeled system and as-built system
    - Interactions between social and technical part of system may invalidate technical assumptions underlying analysis
    - Effectiveness of mitigation measures may change over time
  - Why are likelihood estimates inaccurate in practice?
    - Important factors left out (operator error, flawed decision making, software) because don't have probability estimates
    - Non-stochastic errors involved in leading indicator events
    - Heuristic biases

Heuristic Biases

- Confirmation bias (tend to deny uncertainty and vulnerability)
- Construct simple causal scenarios
  - If none comes to mind, assume impossible
- Tend to identify simple, dramatic events rather than events that are chronic or cumulative
- Incomplete search for causes
  - Once one cause identified and not compelling, then stop search
- Defensive avoidance
  - Downgrade accuracy or don’t take seriously
  - Avoid topic that is stressful or conflicts with other goals

Controlling Heuristic Biases

- Cannot eliminate completely but can reduce
- Use structured method for identifying, detecting, and managing leading indicators
  - Following a structured process and rules to follow can diminish power of biases and encourage more thorough search
  - Concentrate on plausibility (vulnerability) rather than likelihood
    - Think about whether an assumption could fail, not whether likely to do so
    - Concentrate on causal mechanisms vs. likelihood
- Use worst case analysis (vs. “design basis accident”)

Assumption-Based Leading Indicators

- Hypothesis: Useful leading indicators can be identified based on
  - Assumptions underlying safety engineering practices and
  - Vulnerability of those assumptions

  (rather than likelihood of loss events)
- Monitor assumptions on which safety of system was assured to find
  - Assumptions that were originally incorrect
  - Those that have become incorrect over time

Definitions

- **Leading Indicator**

  1. Warning sign used to detect when a safety-related assumption is broken or dangerously weak and action needed to prevent an accident
  2. Warning signal that validity or vulnerability of an assumption is changing

- **Shaping Actions**

  – Used to maintain assumptions, prevent hazards, and control migration to states of higher risk, e.g.,
    - Interlocks
    - Dessicant to prevent corrosion
    - Design human operation to be easy and hard to omit

  – Feedforward control

- **Hedging (Contingency) Actions**
  - Prepare for possibility an assumption will fail
  - Generate scenarios from broken assumptions (worst case analysis) to identify actions that might be taken
  - Feedback control
  - Examples:
    - Performance audits
    - Fail-safe design (e.g., protection and shutdown systems)
- **Signposts**
  - Points in future where changes in safety controls (shaping and hedging actions) may be necessary or advisable
  - Examples: New construction or known future changes may trigger a planned response or MOC

- **Assumption Checking**
  - Checking whether assumptions underlying safety design are still valid
  - (Signposts identified during design and development and specific responses specified)
  - Monitor operations to determine if assumptions still valid
  - Might focus on signposts or on assumptions that have not been adequately handled by shaping and hedging actions
  - Accidents often occur after a change
    - Signposts used for planned or expected changes
    - Assumption checking used for detecting unplanned and potentially unsafe change

Assumptions about Why Accidents Occur

- Starting point for identifying assumption-based leading indicators

- Development and Implementation
  - Inadequate hazard analysis (assumptions about system hazards or hazard analysis process do not hold)
  - Inadequate design of control and mitigation measures for identified hazards perhaps due to
    - Inadequate engineering knowledge
    - Inappropriate assumptions about operations

Assumptions about Why Accidents Occur (2)

- Operations
  - Controls designers assumed would exist are not adequately implemented or used
  - Controls are implemented but changes over time violate assumptions underlying original design of the controls
    - New hazards arise with changing conditions, not anticipated during design and development, or dismissed as unlikely to occur
    - Physical controls and mitigation measures degrade over time in ways not accounted for in design
    - Components (including humans) behave differently over time (violate assumptions made during design)
    - System environment changes over time

Assumptions about Why Accidents Occur (3)

- Management
  - Safety management system design is flawed
  - Safety management system does not operate as was designed (assumed) perhaps because (examples)
    - Safety culture degrades over time
    - Behavior of those making safety-critical decisions may be influenced by competitive, financial, or other pressures.
- To prevent accidents must
  - Eliminate or reduce occurrence of these causes
  - Response may take form of shaping or hedging actions
  - Leading indicators program can be used to try to detect them before an accident occurs.

Vulnerability vs. Likelihood

- Vulnerability: assessment of whether assumption could plausibly fail during lifetime of system, not specific probability of that happening
- If assumption vulnerable, then should protect against it in some way
  - Vulnerability may change over time
  - Need to identify when vulnerability has changed
- Helps reduce heuristic biases:
  - Whether assumption could fail to hold, not whether likely to
  - Concentrate on causal mechanisms rather than likelihood

Is Vulnerability just a Proxy for Likelihood?

- Difference is potential for error
  - Not assigning a probability or a relative category
    - Frequent, probable, occasional, remote, improbable, impossible
    - Categories usually undefined oir poorly defined
  - Instead only two categories: possible or impossible
  - If likelihood not zero, then needs to be included in leading indicators program
  - Does not mean costly controls must be used
  - Does mean that cannot dismiss it at beginning of development and never consider again (until first accident)

Identifying Safety-Critical Assumptions

- Process based on STAMP model
  - Specify assumptions made during engineering development
  - Use STPA to identify safety constraints and controls needed as well as specific assumptions underlying shaping and hedging actions designed to prevent hazards/losses
  - Real examples and detailed explanation in paper
  - Identify severity and vulnerability
  - Generate leading indicators along with
    - Associated assumptions
    - How will be checked
    - When will be checked
    - Hedging actions to take if indicator is true

Checking Leading Indicators

- Many if not most can be handled through shaping and hedging actions. Then need to check these are effective.
- Accident/incident analysis process and error reporting system
- Periodic performance audits
- Signposts
- Monitoring program
  - Dokas: EWaSAP (Early Warning Sign Analysis using STPA

Managing a Leading Indicators Program

- Integrate into risk management program
- Communicate to decision makers when fail
- Develop detailed action plans and triggers for implementing them before assumptions found to be invalid
  - To lessen denial and avoidance behavior
  - To overcome organizational and cultural blinders
- May need to assign responsibility to independent organization and not project managers or those with conflicting pressures
- Periodically revisit list of leading indicators. Establish a continuous improvement process

Feasibility Considerations

- Most assumptions identified and considered during development so just need to document them.
- I've done it for TCAS II (technical) and NASA ITA program (management)
- Dokas has successfully used EWaSAP on industrial systems
- One reason for not handling worst case and just likely case is hazard analysis cost (remove through PHA).
  - STPA turning out to be much cheaper than older methods
  - Can create protection for where assumed risk even if do not know all the causes. May not be most efficient but also not very likely to be needed,