STAMP: Where To From Here?

Nancy Leveson

MIT

Outline

- Quick Intro
- Evaluations and Comparisons
- Where to from here?
- Conclusions

Our current tools are all 50-65 years old but our technology is very different today

1940 1950 1960 1970 1980 1990 2000 2010 2020

FMEA

FTA HAZOP ETA

Bow Tie (CCA) FTA + ETA

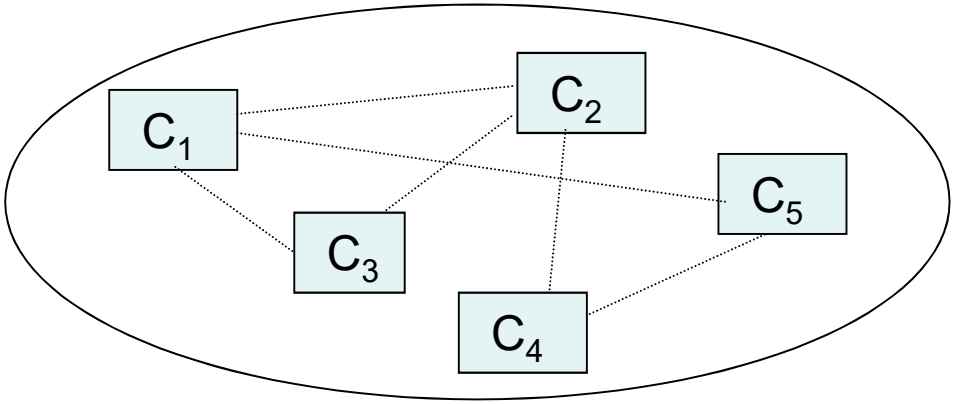
- Introduction of computer control
- Exponential increases in complexity
- New technology
- Changes in human roles

Assumes accidents caused by component failures

Analytic Reduction ("Divide and Conquer")

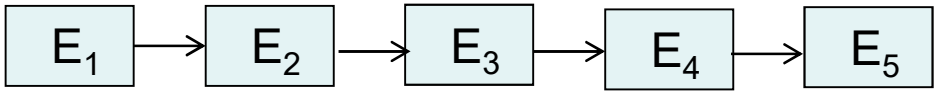
1. Divide system into separate parts

Physical/Functional: Separate into distinct components



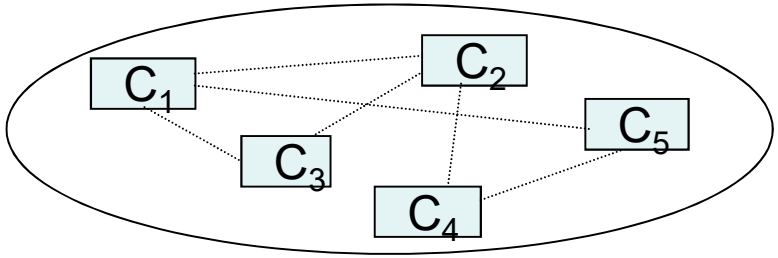
Components interact
In direct ways

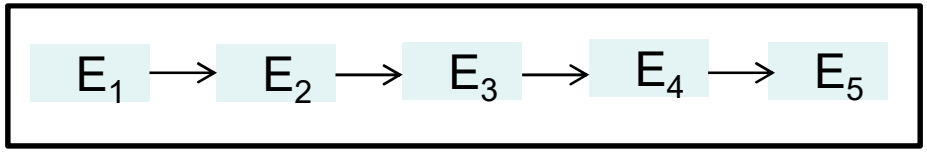
Behavior: Separate into events over time



Each event is the direct
result of the preceding event

Analytic Reduction (2)





2. Analyze/examine pieces separately and combine results

- Assumes such separation does not distort phenomenon
  - ✓ Each component or subsystem operates independently
  - ✓ Components act the same when examined singly as when playing their part in the whole
  - ✓ Components/events not subject to feedback loops, indirect, and non-linear interactions
  - ✓ Interactions can be examined pairwise

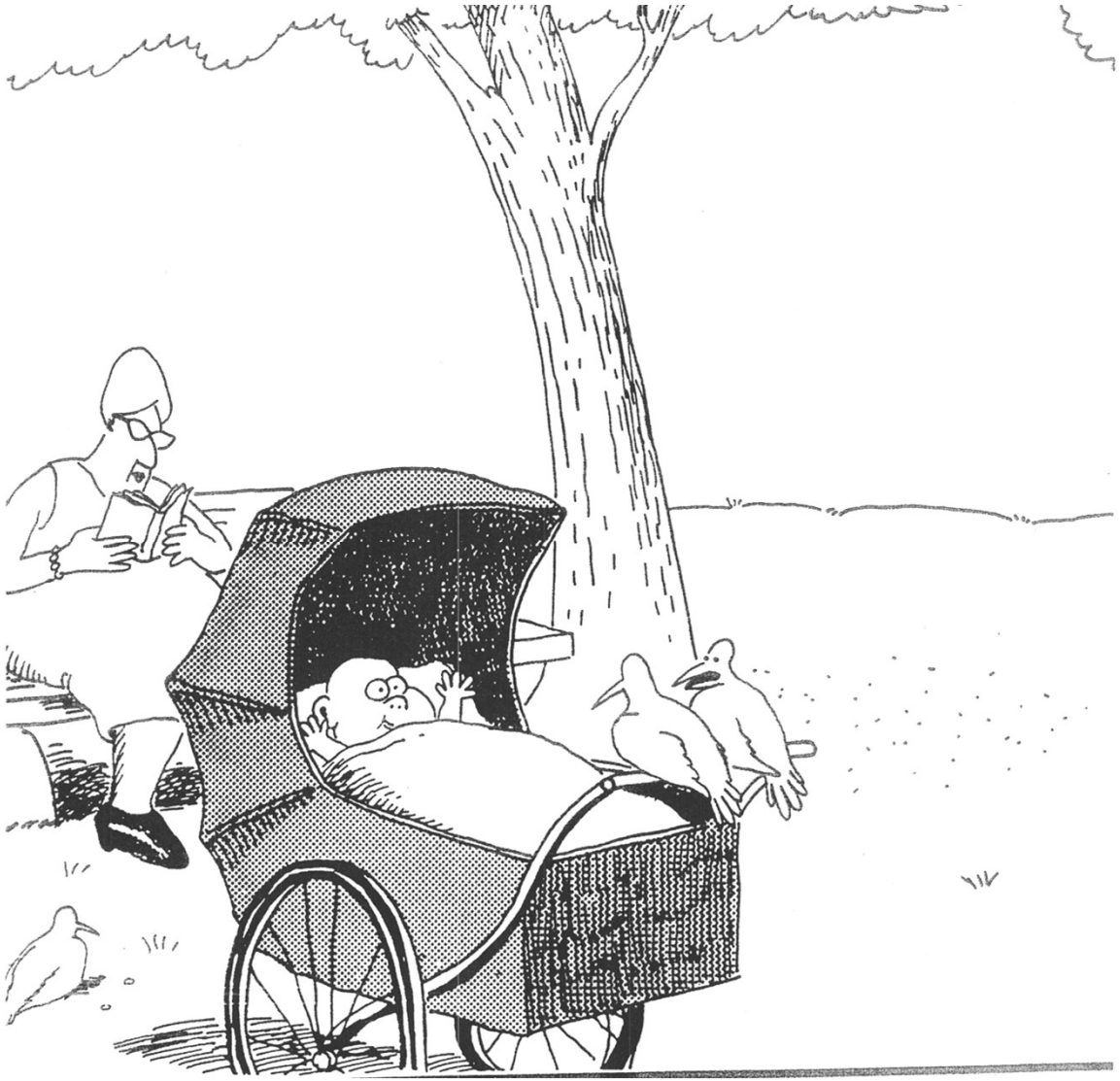
Bottom Line

- These assumptions are no longer true in our
  - Tightly coupled
  - Software intensive
  - Highly automated
  - Interconnected

  engineered systems today
- Need a new theoretical basis
  - *System theory* can provide it

Standard Approach does not Handle

- Component interaction accidents
- Systemic factors (affecting all components and barriers)
- Software and software requirements errors
- Human behavior (in a non-superficial way)
- System design errors
- Indirect or non-linear interactions and complexity
- Migration of systems toward greater risk over time (e.g., in search for greater efficiency and productivity)



It’s still hungry … and I’ve been stuffing worms into it all day.

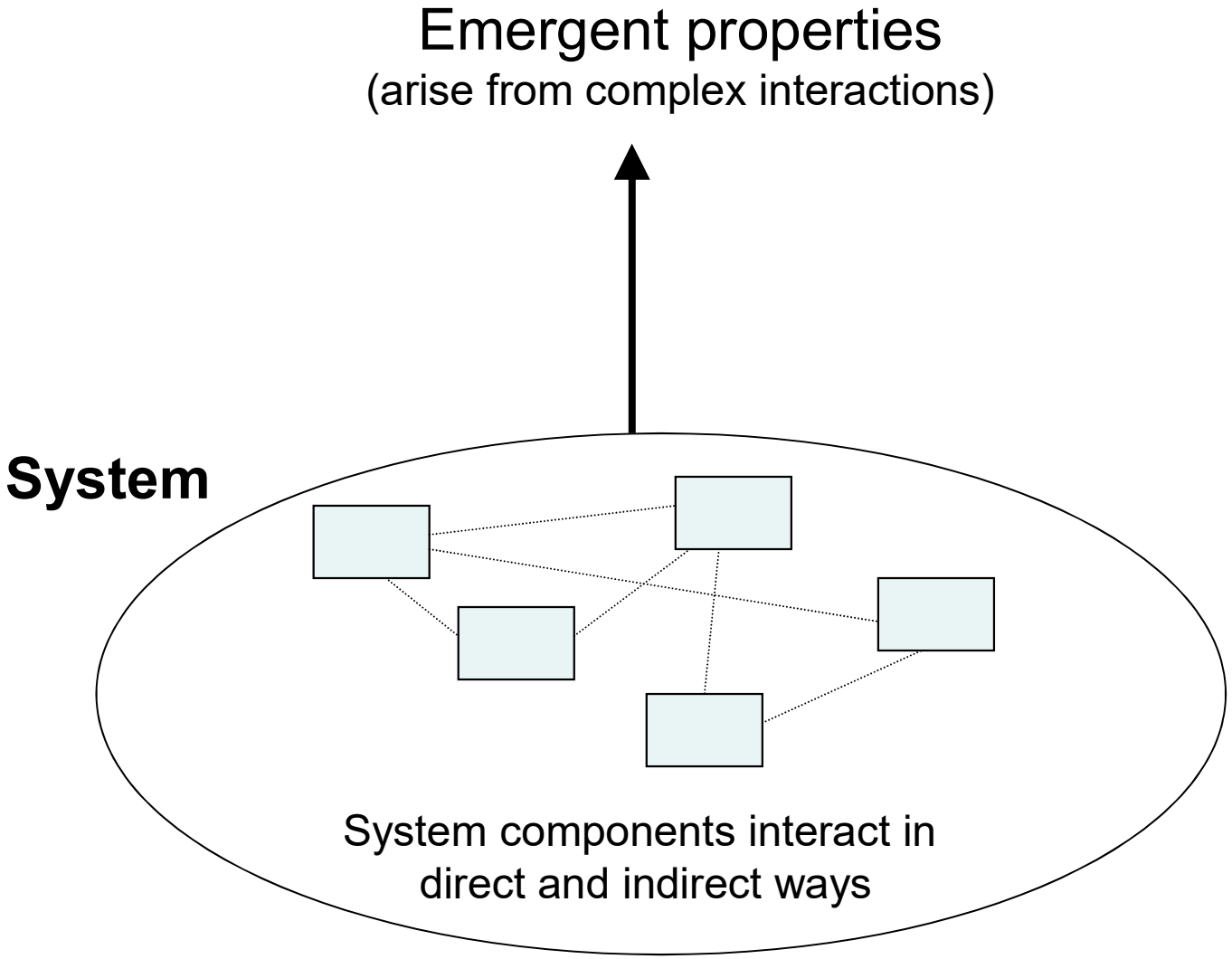
Systems Theory

- Focuses on systems taken as a whole, not on parts taken separately
- Emergent properties
  - Some properties can only be treated adequately in their entirety, taking into account all social and technical aspects

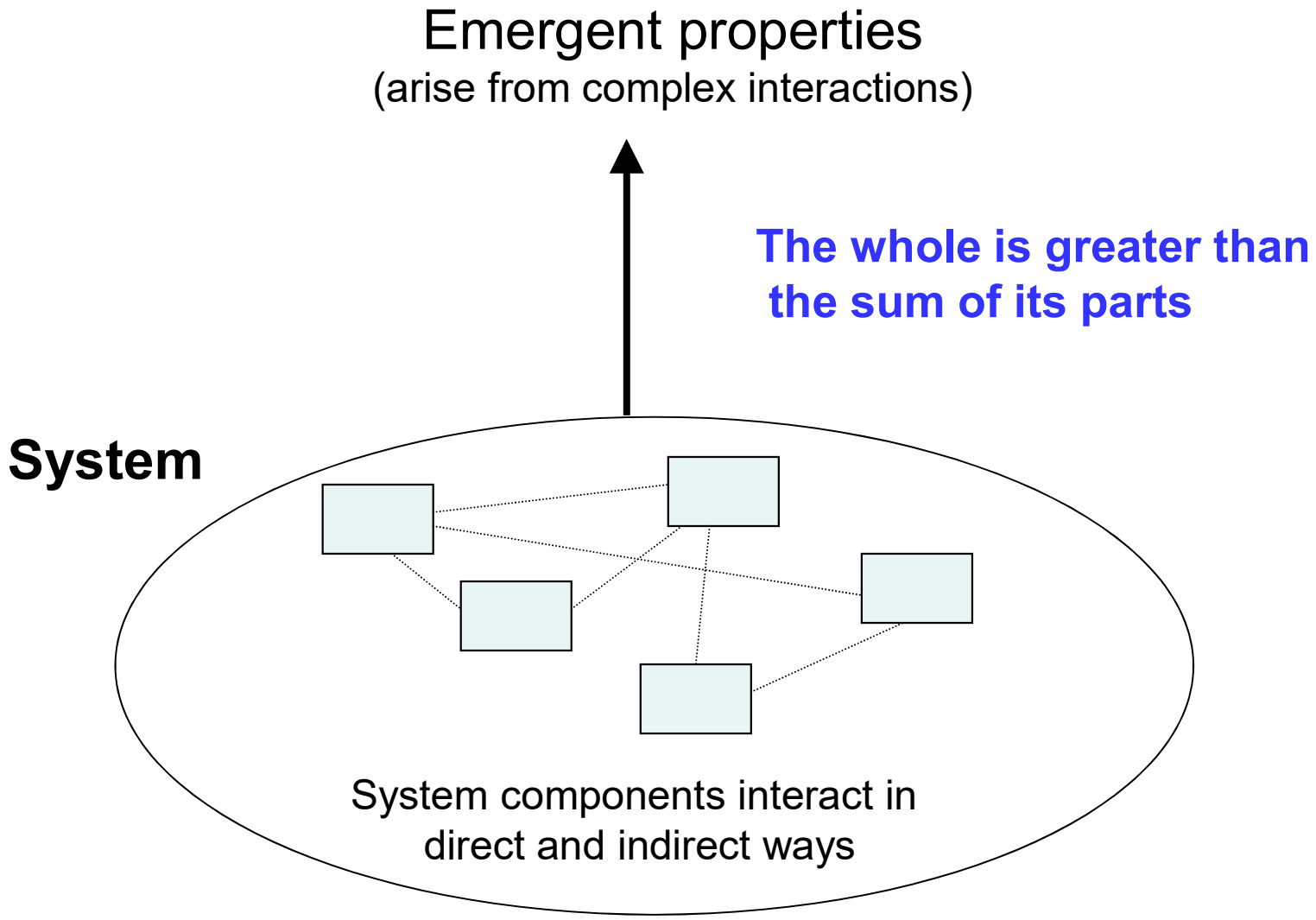
    "The whole is greater than the sum of the parts"

  - These properties arise from relationships among the parts of the system

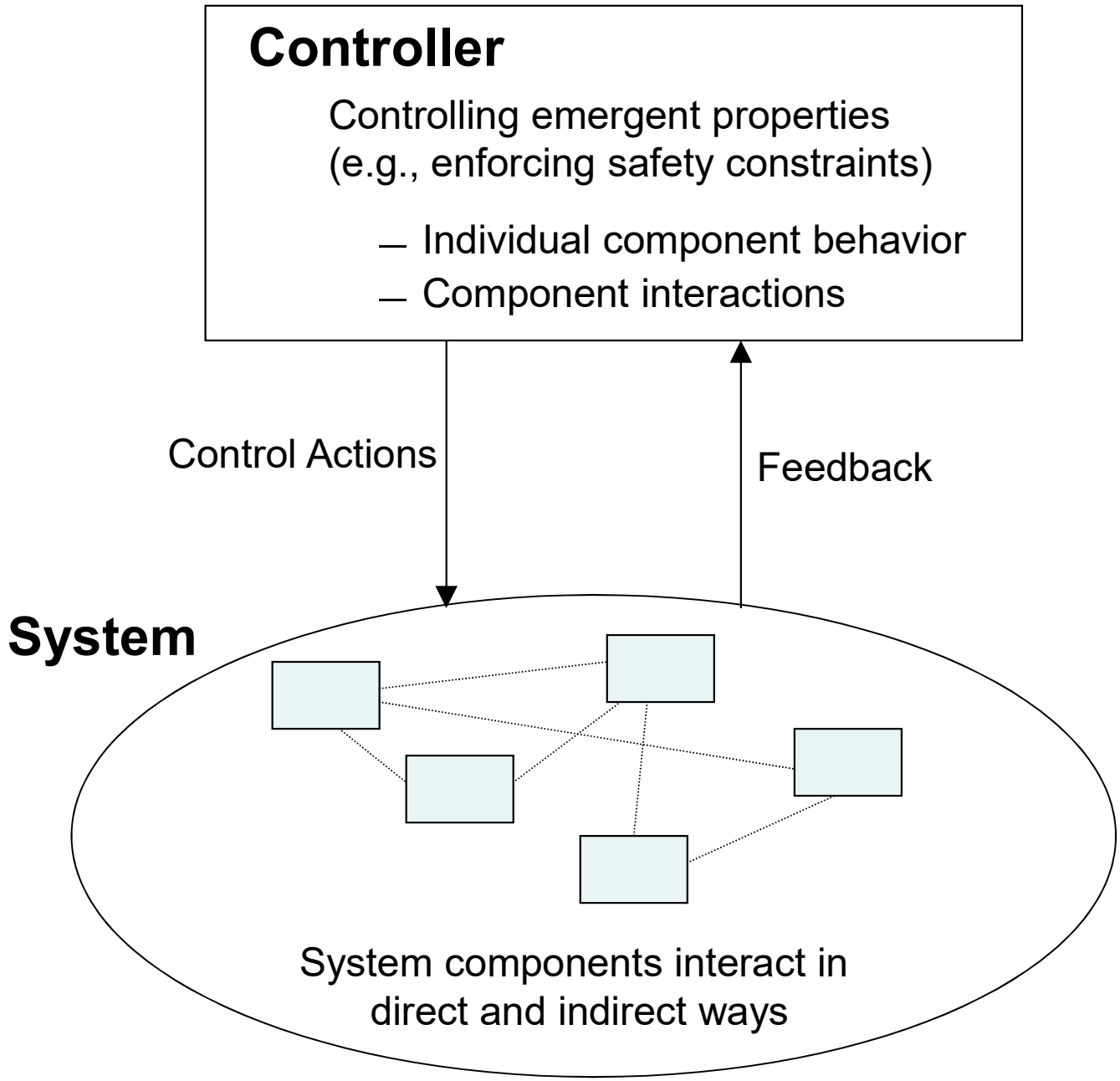
    How they interact and fit together



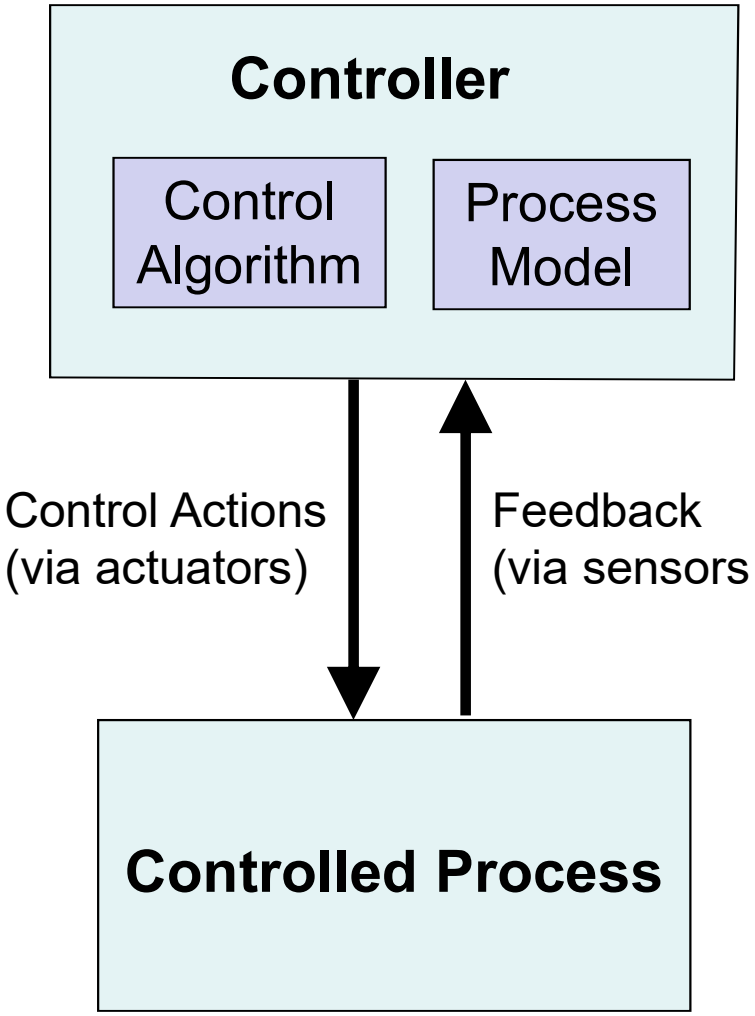
Safety and security are emergent properties



Safety and security are emergent properties


Controller
Controlling emergent properties
(e.g., enforcing safety constraints)
– Individual component behavior
– Component interactions
Control Actions
Feedback
System
System components interact in
direct and indirect ways

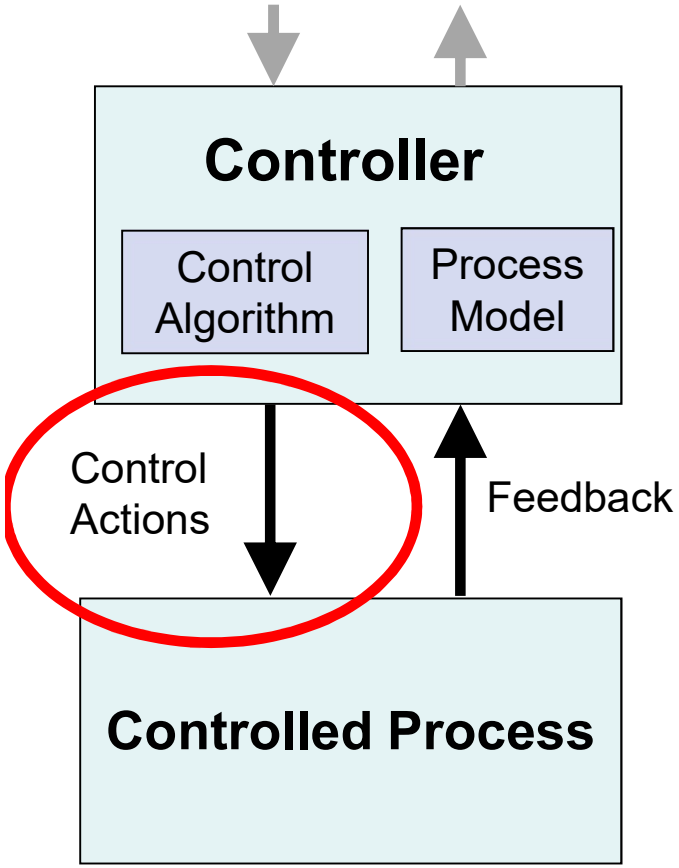
Role of Process Models in Control



- Controllers use a **process model** to determine control actions
- Software/human related accidents often occur when the process model is incorrect
- Captures software errors, human errors, flawed requirements …

(Leveson, 2003); (Leveson, 2011)

Unsafe Control Actions



Four types of unsafe control actions

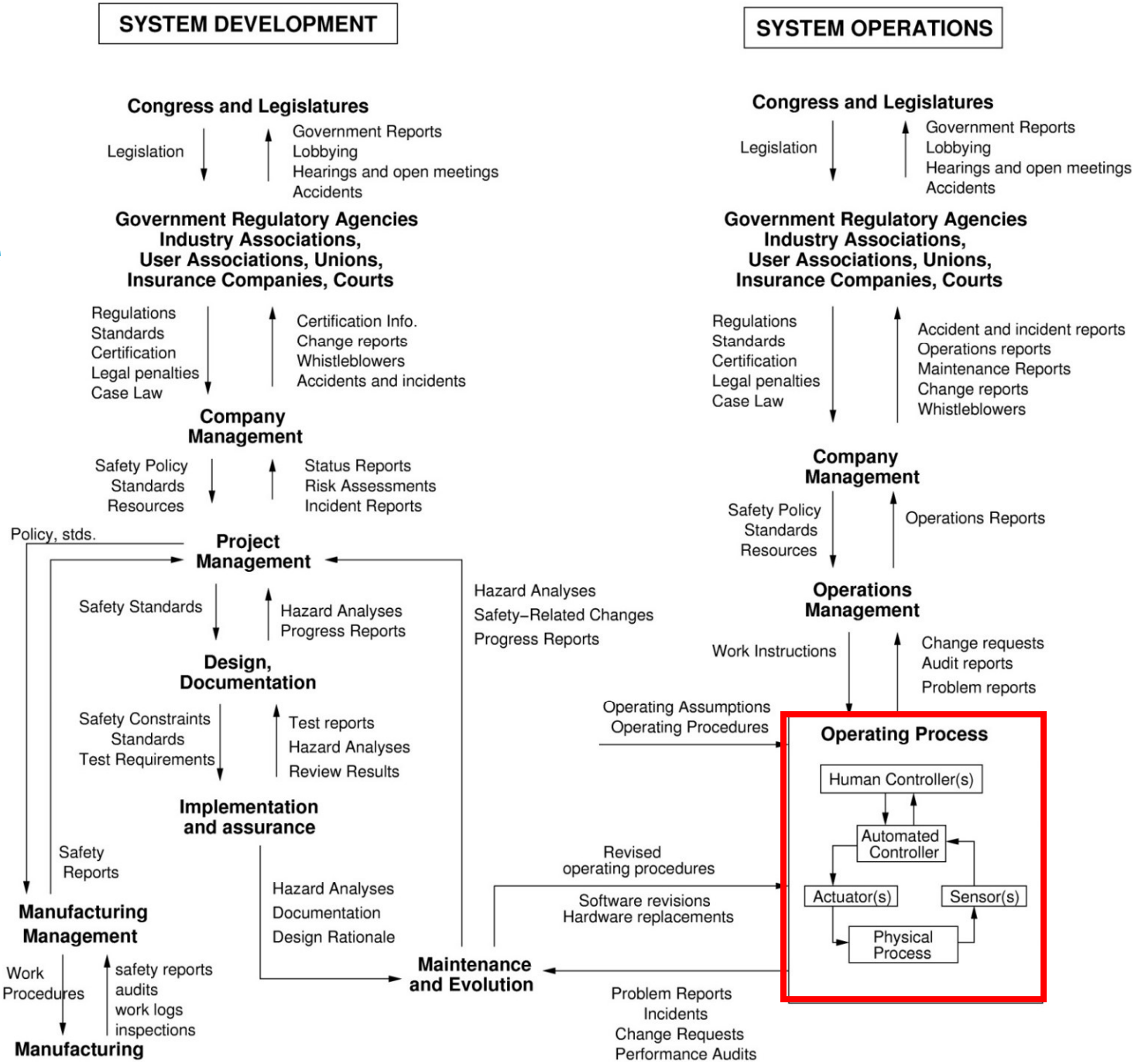
1) Control commands required for safety are not given
2) Unsafe commands are given
3) Potentially safe commands but given too early, too late
4) Control action stops too soon or applied too long (continuous control)

Analysis:

1. Identify potential unsafe control actions
2. Identify why they might be given
3. If safe ones provided, then why not followed?

(Leveson, 2012)

Example Safety Control Structure

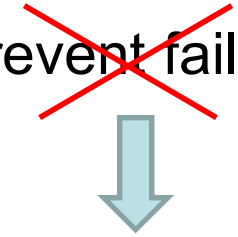


STAMP (System-Theoretic Accident Model and Processes)

- Expands the traditional model of the accident causation (cause of losses)
  - Not just a chain of directly related failure events
  - Losses are complex processes
- Defines safety/security as a **dynamic control** problem (vs. failure problem)
  - Works on very complex systems
  - Includes software, humans, organizations, safety culture …
  - Based on **systems theory** (vs. reliability theory)
- Allows creating new tools (STPA, accident analysis, leading indicators, etc.) and partially automated analysis

Safety as a Dynamic Control Problem (STAMP)

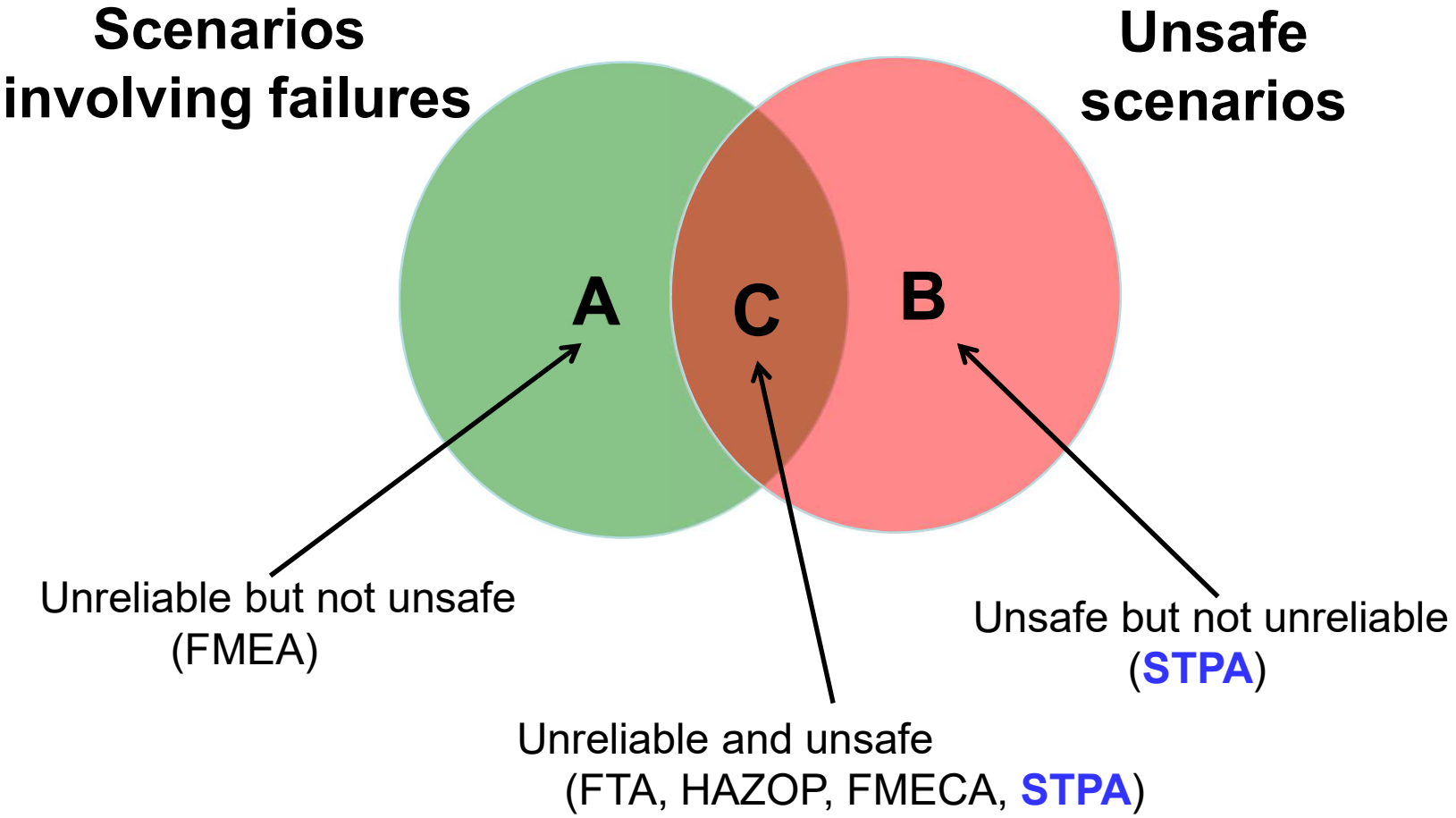
- Can be used in early concept formation stage
- Applies to security too:
  - Combined safety/security analysis
  - Just requires adding additional causal scenarios
- A change in emphasis:



“enforce safety/security constraints on system behavior”

(note that enforcing constraints might require preventing failures or handling them but includes more than that)

Confusing Safety and Reliability



Preventing Component or Functional Failures is Not Enough

Evaluation: Does it Work?

Is it Practical?

- STPA has been or is being used in a large variety of industries
  - Automobiles (about 60-80% of car companies are using it)
  - Aviation (including UAVs, airlines)
  - Space Systems
  - Air Traffic Control
  - Defense
  - Medical Devices and Hospital Safety
  - Chemical plants
  - Oil and Gas
  - Nuclear and Electrical Power
  - Particle Accelerators (CERN)
  - Finance
  - Universities
  - Mining

Does it Work?

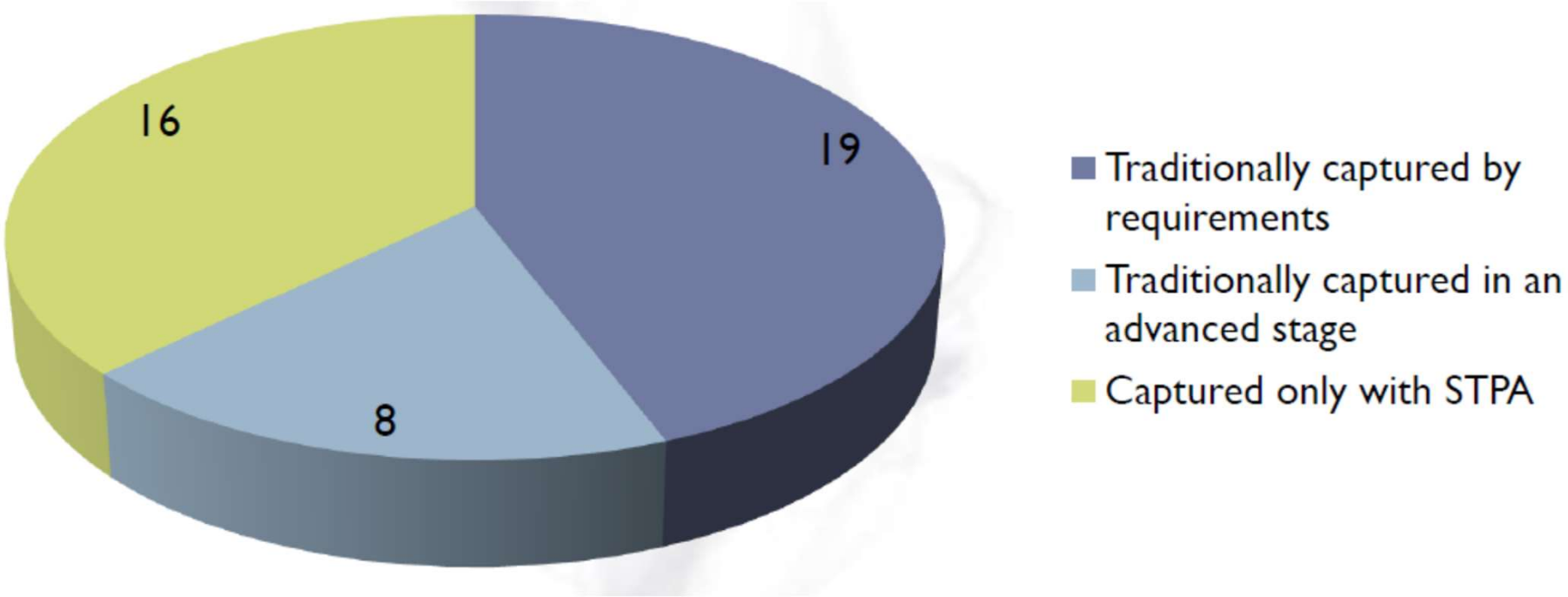
- Most of these systems are very complex (e.g., the new U.S. missile defense system)
- In all cases where a comparison was made (to FTA, HAZOP, FMEA, ETA, etc.)
  - STPA found all the same hazard causes as the old methods
  - Plus it found more causes than traditional methods
  - In some evaluations, found accidents that had occurred that other methods missed
  - Cost was orders of magnitude less than the traditional hazard analysis methods
  - Same results for security evaluations

Some Comparisons

- Embraer Air Management System
  - 3.5 months
  - Identified 200+ safety constraints (requirements) and 700+ design recommendations to eliminate or mitigate hazards (satisfy the safety constraints).
- U.S. Navy Vessel with Dynamic Positioning System
  - Compared STPA results with official FTA/FMEA (STPA commissioned after 2 serious accidents during test)
  - All failures identified by FTA/FMEA identified by STPA plus lots of “non-failure” hazard causes
  - STPA identified scenarios never corrected. Put into service and collided with nuclear submarine (cause was identified by STPA)

More Comparisons

- Embraer Aircraft Smoke Control System requirements captured by STPA



- EPRI Nuclear Power Plant Comparison
  - Compared FTA, FMEA, ETA, HAZOP and STPA
  - Only STPA found accident that had occurred in plant. Analysts were not told about it

And More

- Blackhawk Helicopter: STPA compared with official FTA/FMEA
  - FTA/PHA identified some “hazards” as “marginal” (and thus not considered further) that STPA found led to catastrophic accidents.
  - Causal factors of FTA/FMEA limited to component failures
  - STPA identified non-failure scenarios that could lead to a hazardous state that were not identified by FTA/FMEA. Also found the component failure scenarios.
  - More information about causal scenarios from STPA results led to more cost/effective mitigation measures even for failures (beyond redundancy).

And Even More

- U.S. Air Force hazard analysis in flight testing vs. STPA

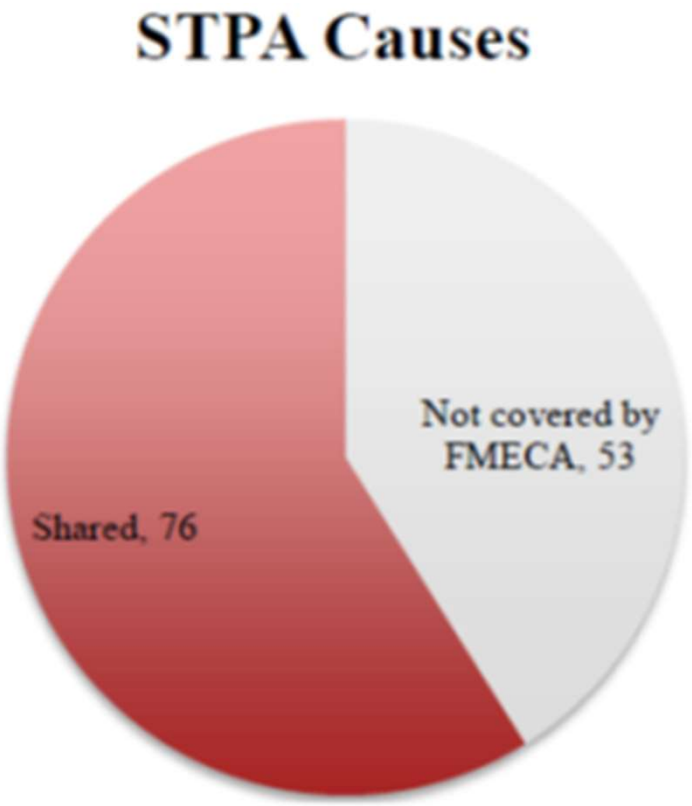
Traditional	STPA
2 Effects	6 Accidents
1 Test Hazard (actually a mishap)	4 System Hazards
3 Causes	392 Unsafe Control Actions
13 Minimizing Procedures *- 8 THA minimizing procedures* *- 5 general minimizing procedures*	46 Minimizing Procedures *- 14 developing influences* *- 10 settings/configurations* *- 22 operating procedures*
Nothing identified to control hazard exposure (test hazard was a mishap)	8 Corrective Actions
1 Accident-Corrective Action	7 Recovery Actions

- In-Trail Procedure (NextGen/Open Skies) DO-312
  - Overlooked critical scenarios that STPA identified
  - Dismissed scenarios as “no safety effect” that STPA identified as critical
  - Human error oversimplified and superficial compared to STPA. Treated as random vs. identifying causal factors so could be reduced.
- U.S. Ballistic Missile Defense System
  - Used STPA just prior to deployment and field testing.
  - Two people, 5 months
  - Found so many paths to inadvertent launch that deployment delayed 6 months to fix them

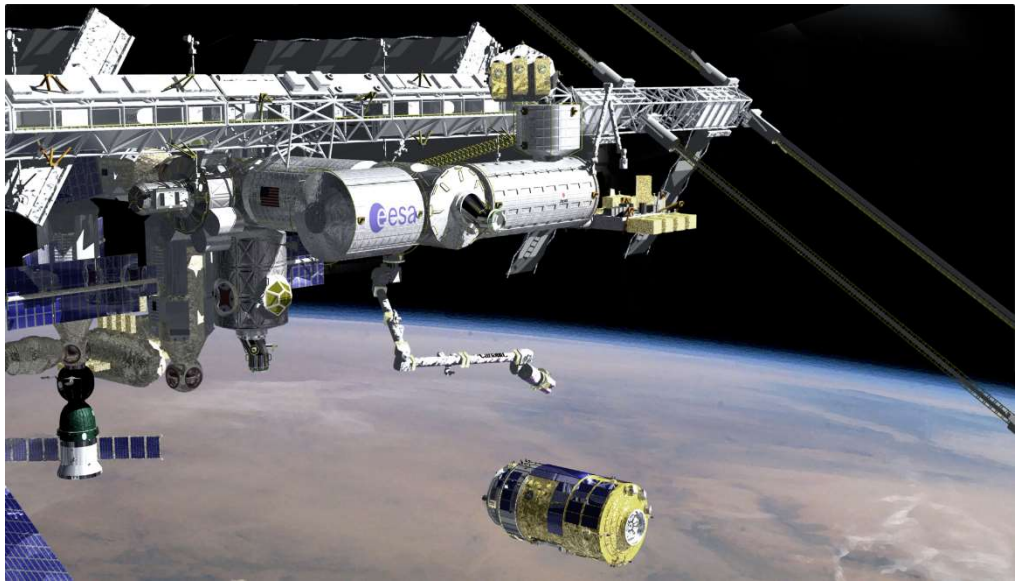
- Quality at Cummins Engine
  - Design teams found more causal factors for quality losses using STPA than FMEA or FTA.
- Range Extender System for Electric Vehicles (Valeo)
  - FTA/CPA took 3 times effort of STPA, found less
- Medical Device (Class A recall)

FMECA	STPA
70+ causes of accidents	175+ causes accidents (9 related to adverse event)
Team of experts	Single semi-expert
Time dedication: months/years)	Time: weeks/month
Identified only single fault causes	Identified complex causes of accidents

- Automotive Electric Power Steering System



- HTV Unmanned Japanese Spacecraft
  - STPA found all causes found by FTA plus a lot more



Some Recent Additions to STPA

- More sophisticated human factors analysis
- Coordination between human and computer controllers (shared control)
- Organizational/managerial analysis
- (Probably more I’ve forgotten)

What do we still need to do?

Better Ways to Make Risk-Based Decisions

- PRA does not work. May make things worse.
- STPA can provide information. Now need ways to make decisions from it. Maybe alter the definition of risk, which now requires a value for likelihood/probability. Could we somehow instead use vulnerability?
  - Probability is unknowable; requires reading the future
  - SUBSAFE is an existence case that PRA is not required
  - What should we use instead?
- Comparative studies: how good are the decisions made from different risk analysis approaches?

Get STPA/CAST Into Standards

- Improved methods need to be in standards or people cannot or will not use them.
- Standards can take 10-20 years to update
  - So always behind the state of the art
  - Write standards without specifying “how”, just “what”
- Standards can be a tugboat or an anchor. At the least they should be neutral.

Create a Handbook for Using STPA/CAST

- Need to define a practical, step-by-step process
  - Industry is starting to create these themselves
  - More is needed in operationalizing STPA and CAST
  - Then need to validate it on real projects
- John and Nancy?

Practicalities for Industrial Use

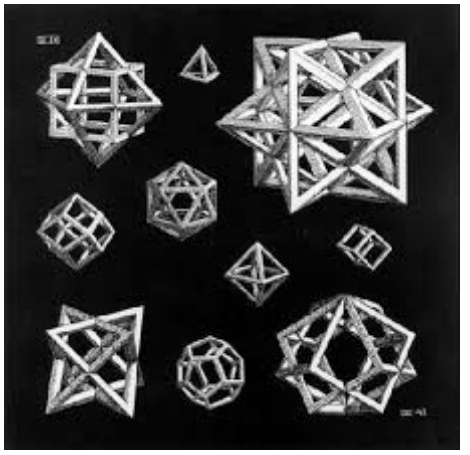
- Easy to use tools
  - May simply assist and not do whole job
  - Stress human factors in tool design and how people work in industry, not how academics think they should work
- Online course (Leveson/Thomas)
- Cost (ROI) data
  - Need to show that safety increases productivity and efficiency (bottom line)
  - More than some anecdotes is needed

Integration into system engineering

- Needs to start in early concept development
- System analysis tools (tradeoff studies, analyzing high level goals and constraints, …) [top left of V-model]
- STPA or maybe something else?
- Model-based system engineering
  - Stop trying to force STPA on architectural models
  - Hazard analysis on architectural models ends up being equivalent to FTA, FMEA, and other component failure analyses
  - Need multiple models (functional, architectural, …)
  - How use results of functional model analysis to create better architectural models and designs?

Software Architecture/Design Implications

- People are stuck on one (or maybe two) design approaches.
- In CS, this is object-oriented design
- Maybe control systems need a different type of “control-oriented design”
- Could potentially simplify
  - Design and development
  - Verification and assurance
  - Traceability
  - Etc.



Use in Operations

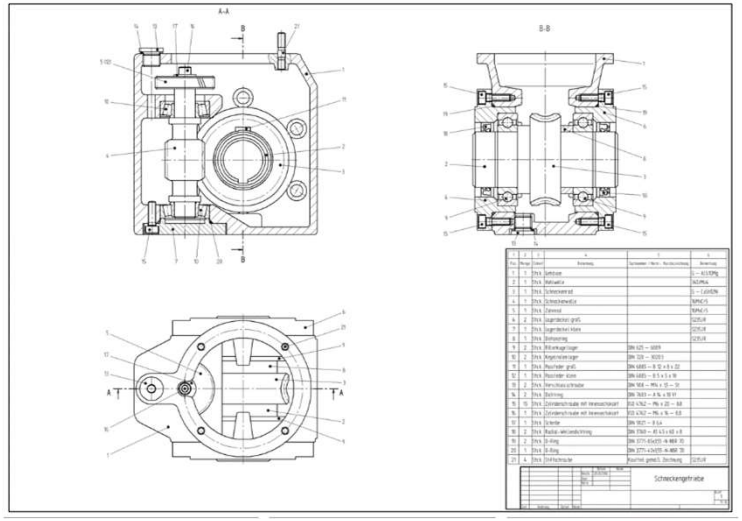
- Leading indicators
- Design of the safety management system
- Design of a risk management system
- Use in operational decision making (including tools)

Workplace Safety

- What people are doing now is obsolete and a waste of time and effort
- Nate Peper's MIT thesis
- Robotics but more than just this

Production Engineering

- Supply chain management
- Design of production/manufacturing systems

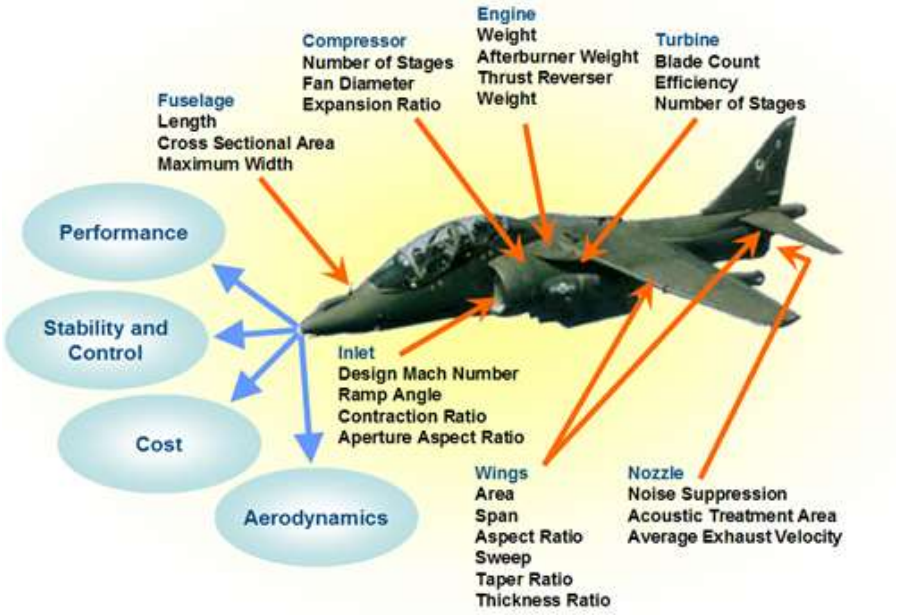


Organizational Analysis

- How to analyze and design an organizational/social system?
- How do culture and basic management principles fit into this new systems-theoretic framework?

System analysis for other emergent properties

- STAMP should apply to all emergent properties
  - Have used for quality in engineering, producibility, …
  - What others? Are special analysis tools needed?
- STPA? (are changes needed?)



Conclusions

- Amazing how fast this is spreading and the success people are having with it
- Lots of potential for improvements, new uses, and problems for researchers to work on.
- Don’t be afraid to tackle important but hard problems and to challenge what everyone knows is “true.”
- Our goal might be to show how companies and government can achieve their goals through environmentally protective and socially responsible actions.
  - Redesign to incentivize people to do the “right” things

BACKUP

Paradigm Change

- Does not imply what previously done is wrong and new approach correct
- Einstein:

  "Progress in science (moving from one paradigm to another) is like climbing a mountain"

As move further up, can see farther than on lower points

Paradigm Change (2)

New perspective does not invalidate the old one, but extends and enriches our appreciation of the valleys below

Value of new paradigm often depends on ability to accommodate successes and empirical observations made in old paradigm.

New paradigms offer a broader, rich perspective for interpreting previous answers.

STPA: System-Theoretic Process Analysis

- A top-down, system engineering analysis technique
- Identifies safety (or X) constraints (system and component requirements)
- Identifies scenarios leading to violation of constraints (requirements); use results to design or redesign system to be safer
- Can be used on technical design and organizational design
- Supports a safety-driven design process where
  - Analysis influences and shapes early design decisions
  - Analysis iterated and refined as design evolves
- Easily integrates into system engineering and MBSE tools

Safety Control Structure for FMIS