

Safety Analysis in Early Concept Development and Requirements Generation

Nancy Leveson

MIT



Bottom Line Up Front (BLUF)

- Complexity is reaching a new level (tipping point)
  - Old approaches becoming less effective
  - New causes of mishaps appearing (especially related to use of software and autonomy)
- Traditional analysis approaches do not provide the information necessary to prevent losses in these systems
- Need a paradigm change

  Change focus

  ~~Increase component reliability (analytic decomposition)~~

  Enforce safe behavior (dynamic control using systems theory)

BLUF (2)

- Allows creation of new analysis and engineering approaches
  - More powerful and inclusive
  - Orders of magnitude less expensive
  - Work on very complex systems (top-down system engineering)
  - Design safety and security and other properties in from the beginning
  - Compliant with MIL-STD-882E and other military standards, commercial standards being developed (autos, aircraft, defense)

- New paradigm works better than old techniques:
  - Empirical evaluations and controlled studies show it finds more causal scenarios (the “unknown unknowns”)
  - Can be used before a detailed design exists to create safety, security, and other requirements

Our current tools are all 40-65 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

Traditional Approach to Safety

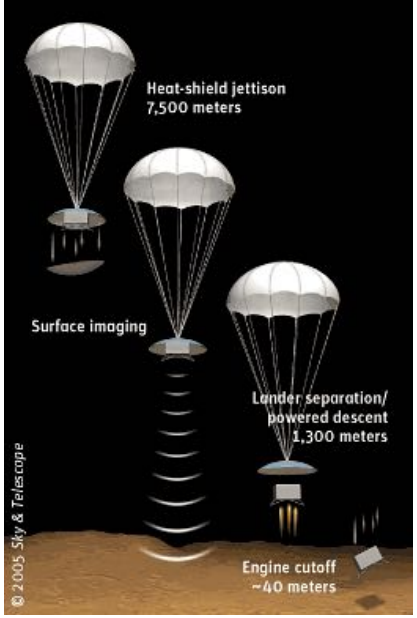
- Traditionally view safety as a failure problem
  - Chain of directly related failure events leads to loss
  - Try to prevent component failures or establish barriers between events
- Limitations
  - Systems are becoming more complex
    - Accidents often result from interactions among components
    - Cannot anticipate all potential interactions
  - Omits or oversimplifies important factors
    - Human error
    - New technology (including software)
    - Culture and management
    - Evolution and adaptation

Accidents are <u>not</u> just the result of random failure

What Failed Here?

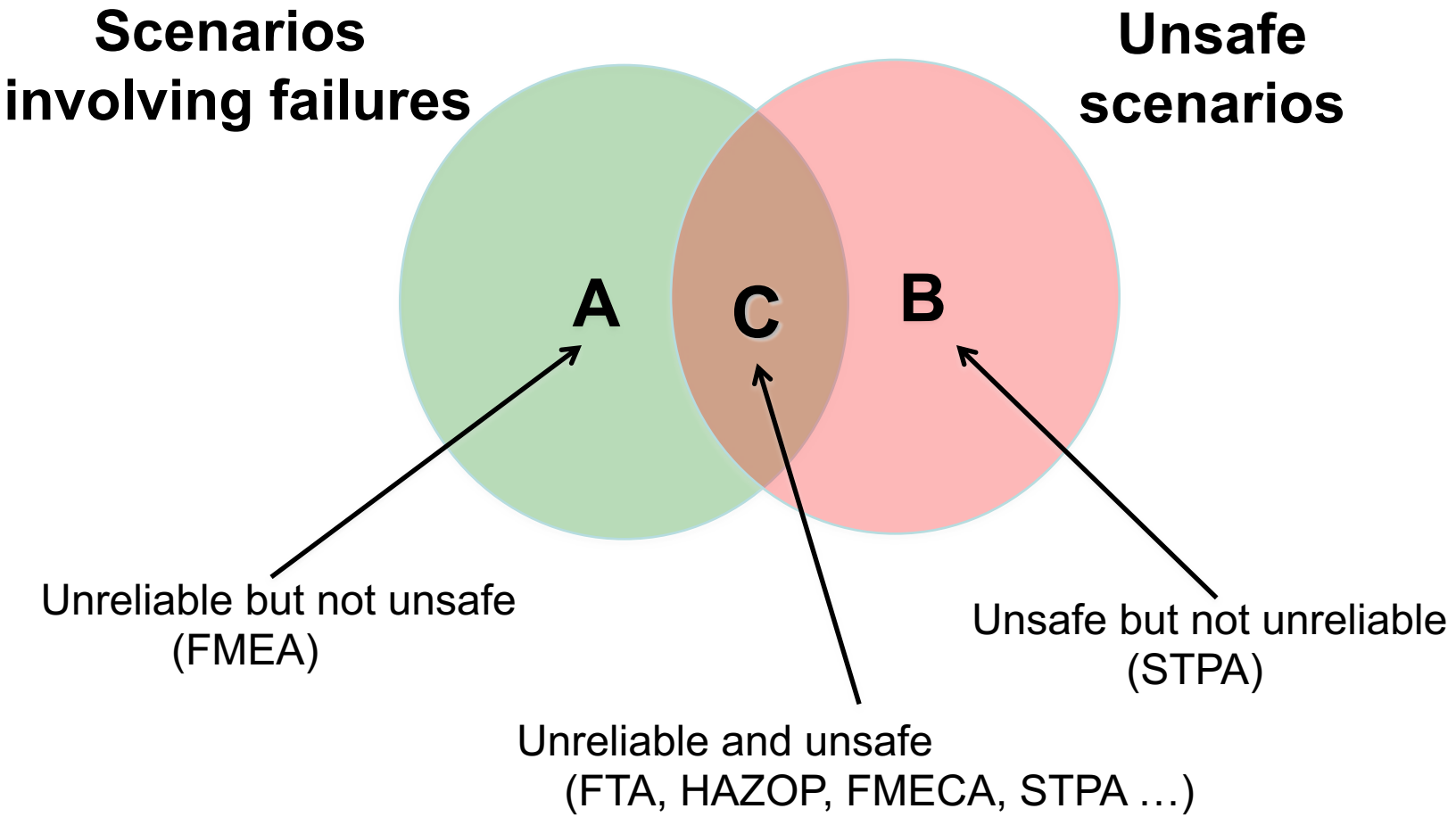
- Navy aircraft were ferrying missiles from one location to another.
- One pilot executed a planned test by aiming at aircraft in front and firing a dummy missile.
- Nobody involved knew that the software was designed to substitute a different missile if the one that was commanded to be fired was not in a good position.
- In this case, there was an antenna between the dummy missile and the target so the software decided to fire a live missile located in a different (better) position instead.

Accident with No Component Failures



- Mars Polar Lander
  - Have to slow down spacecraft to land safely
  - Use Martian atmosphere, parachute, descent engines (controlled by software)
  - Software knows landed because of sensitive sensors on landing legs. Cuts off engines when determines have landed.
  - But “noise” (false signals) by sensors generated when landing legs extended. Not in software requirements.
  - Software not supposed to be operating at that time but software engineers decided to start early to even out the load on processor
  - Software thought spacecraft had landed and shut down descent engines while still 40 meters above surface

Confusing Safety and Reliability



Preventing Component or Functional Failures is Not Enough

General Definition of "Safety"

- Accident = Loss: Any undesired and unplanned event that results in a loss
  - e.g., loss of human life or injury, property damage, environmental pollution, mission loss, negative business impact (damage to reputation, etc.), product launch delay, legal entanglements, etc.
  - Includes inadvertent and intentional losses (security)
- System goals vs. constraints (limits on how can achieve the goals)
- Safety: Absence of losses

Definition of Hazard and Hazard Analysis

Hazard/vulnerability:

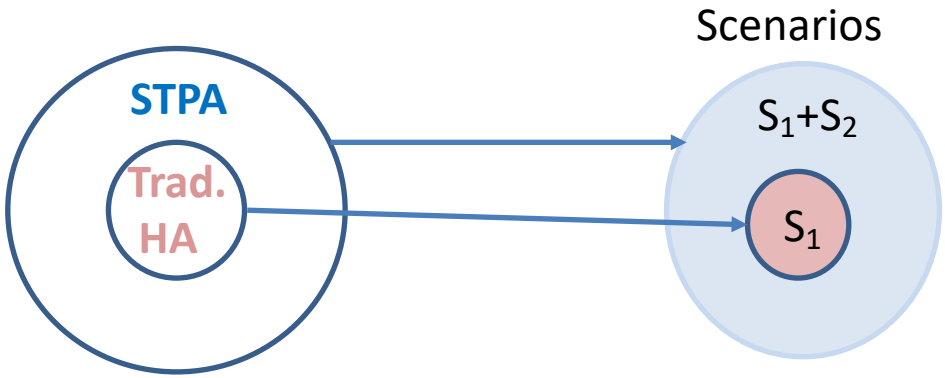
A system state or set of conditions that, together with some worst-case environmental conditions, will lead to a loss

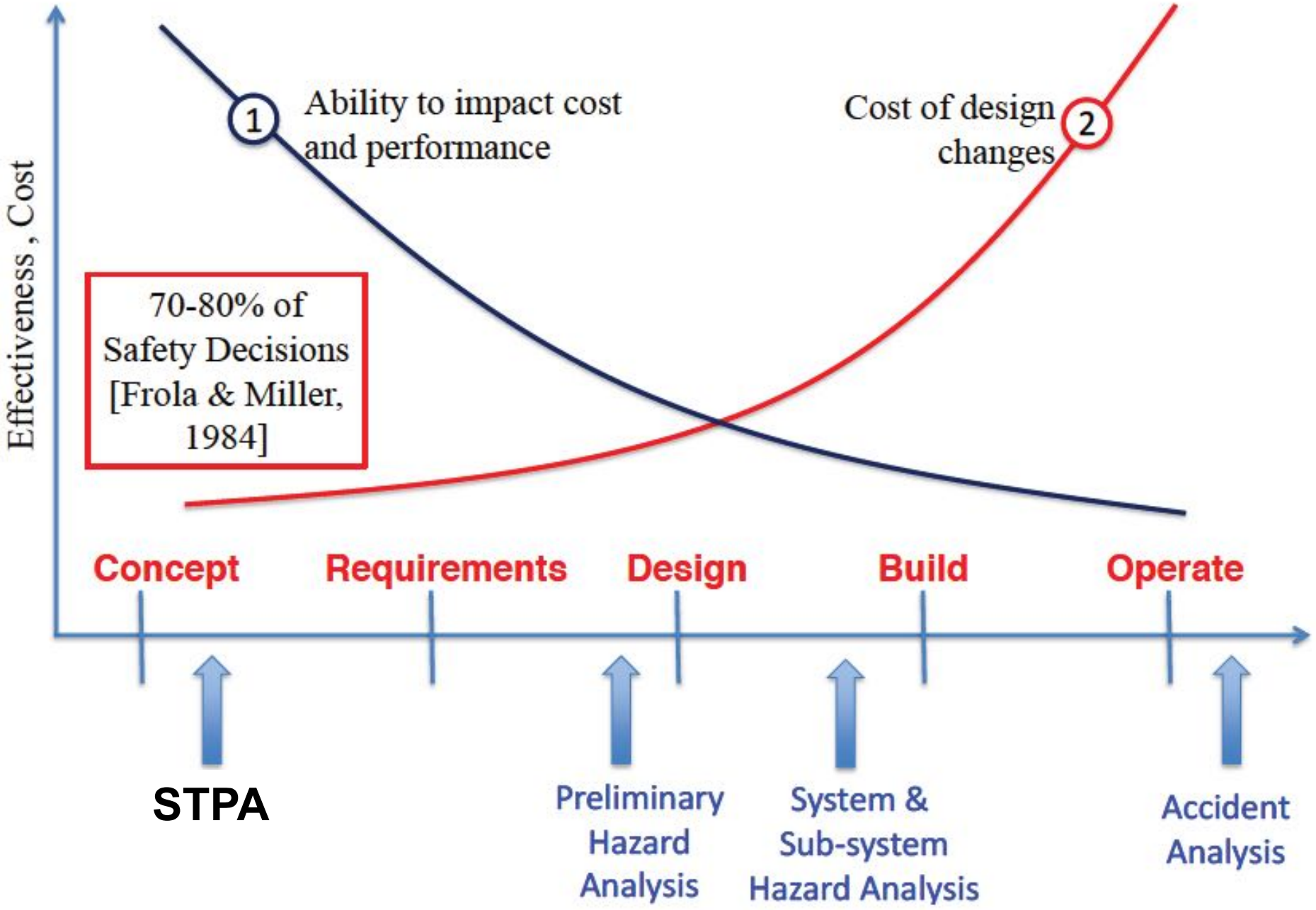
Hazard Analysis:

Identifying operational scenarios that can lead to a hazard/vulnerability

Safety Engineering:

Eliminating or controlling hazard scenarios in the system design and operations




Effectiveness , Cost
1
Ability to impact cost and performance
Cost of design changes
2
70-80% of Safety Decisions [Frola & Miller, 1984]
Concept
Requirements
Design
Build
Operate
STPA
Preliminary Hazard Analysis
System & Sub-system Hazard Analysis
Accident Analysis

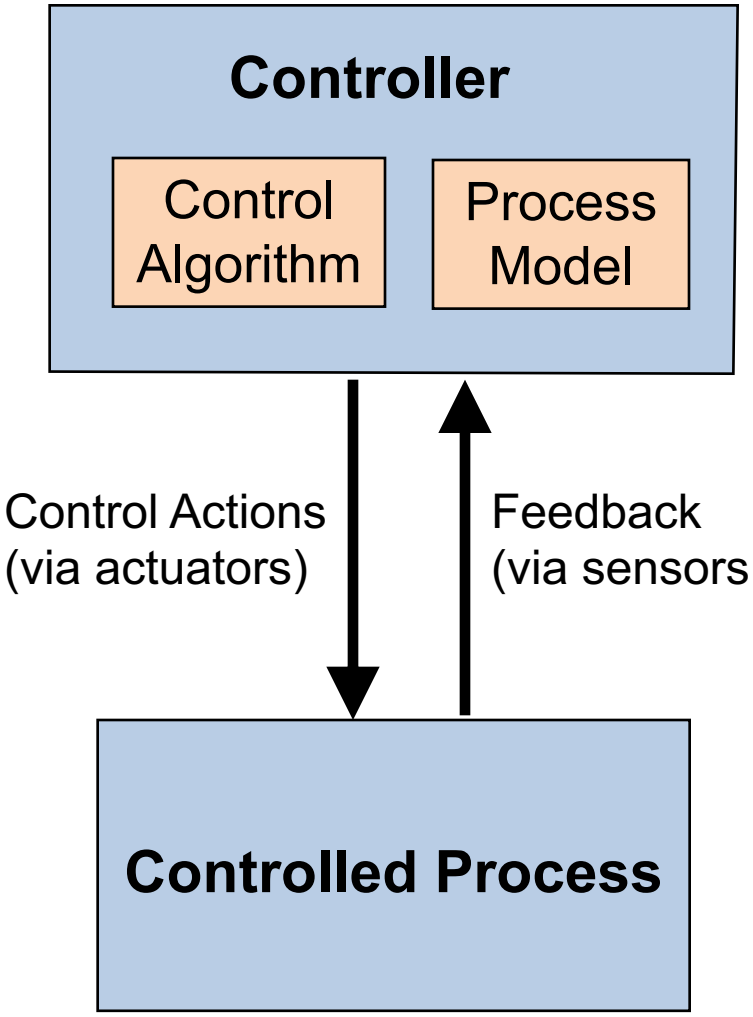
STPA: A New Type of MBSE Modeling and Analysis Technique

- Analysis performed on functional model, not in head
- Not an architectural model or physical or logical diagrams
- Allows a different type of analysis: functional vs. physical or logical behavior
- Perform functional analysis on the control structure
  - Allows generating functional requirements before any architecture or design is created.
  - More than just safety or even cyber security: works for any emergent system property

Some Uses Beyond Traditional System Safety

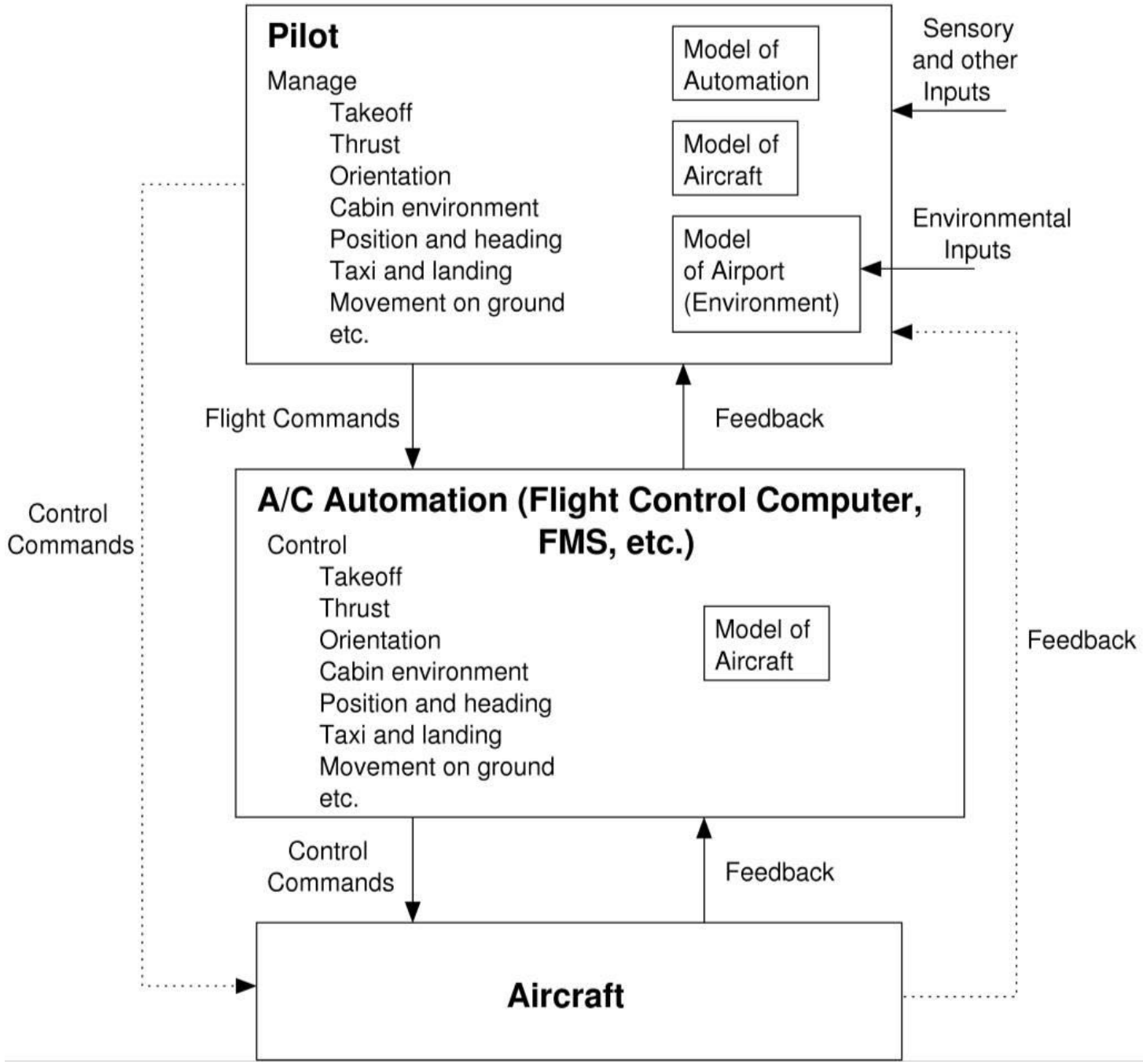
- Airline operations (leading indicators of increasing risk)
- Design of safety management systems
- Cybersecurity
- Quality
- Producibility
- Nuclear security, nonproliferation
- Production engineering
- System Engineering process optimization
- Organizational culture
- Workplace safety
- Banking and finance
- Criminal law

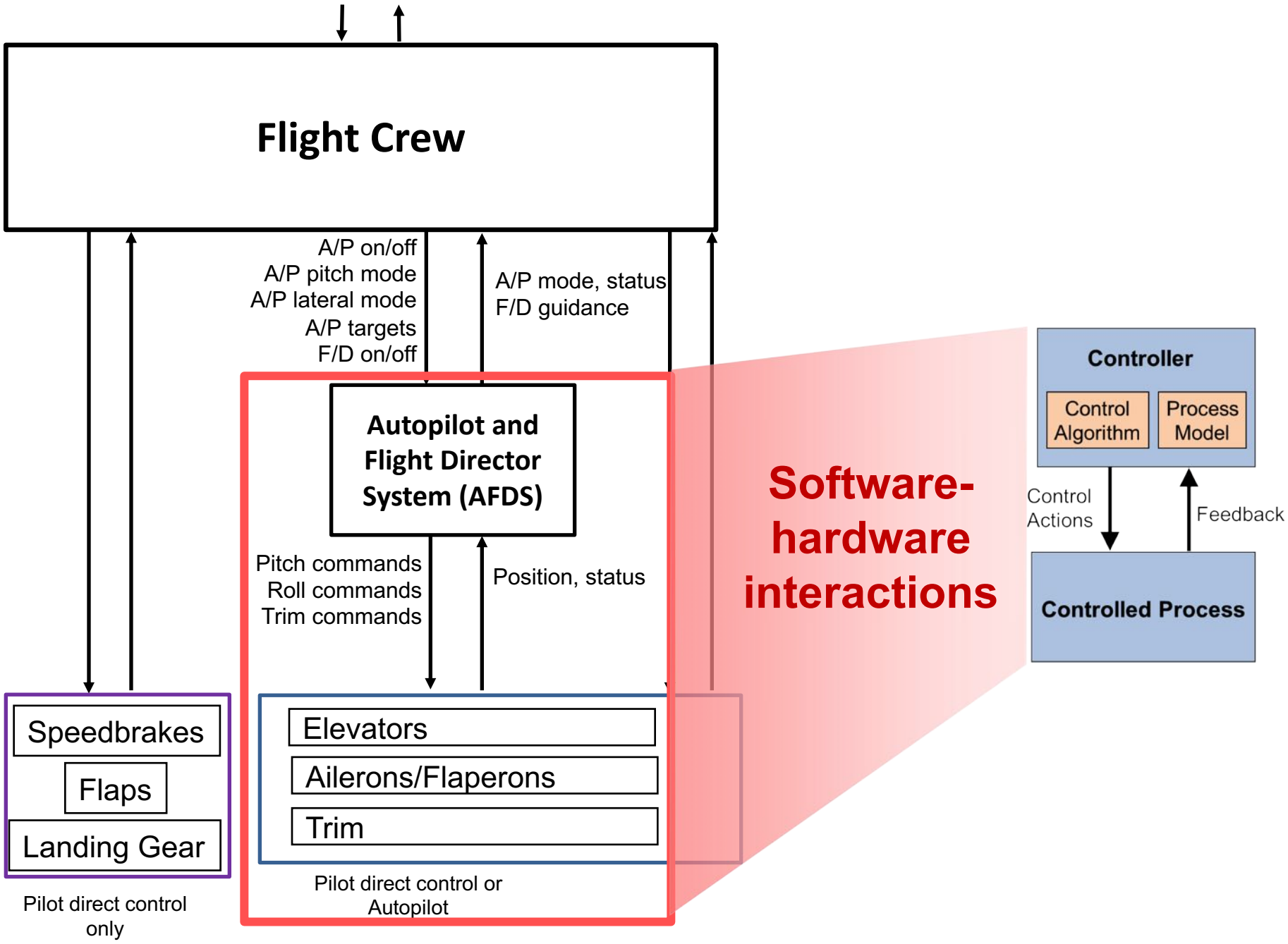
Models Constructed from Feedback Control Loops



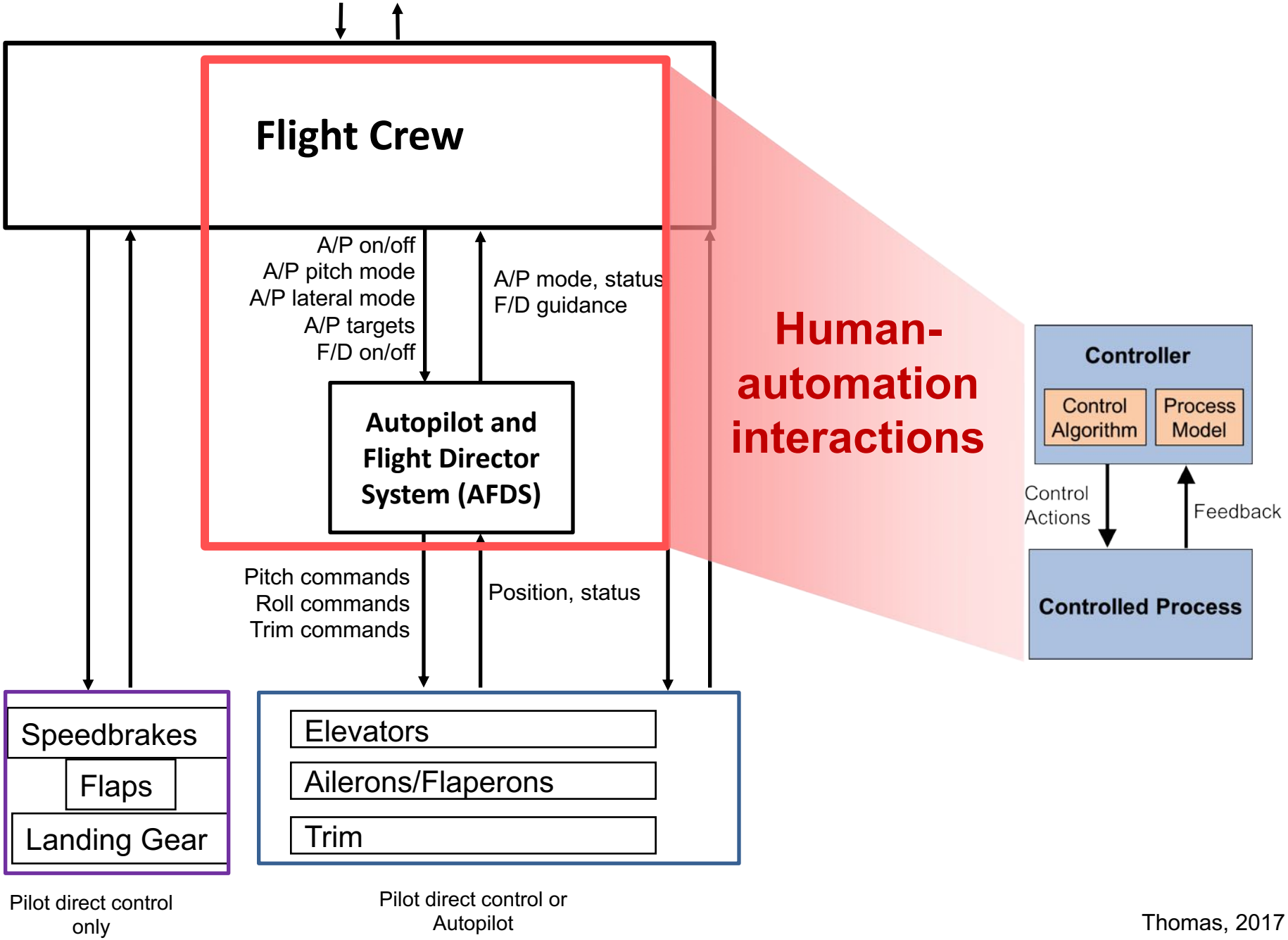
- Controllers use a **process model** to determine control actions
- Software/human related accidents often occur when the process model is incorrect
- Captures software errors, human errors, flawed requirements …

Treat safety as a control problem, not a failure problem

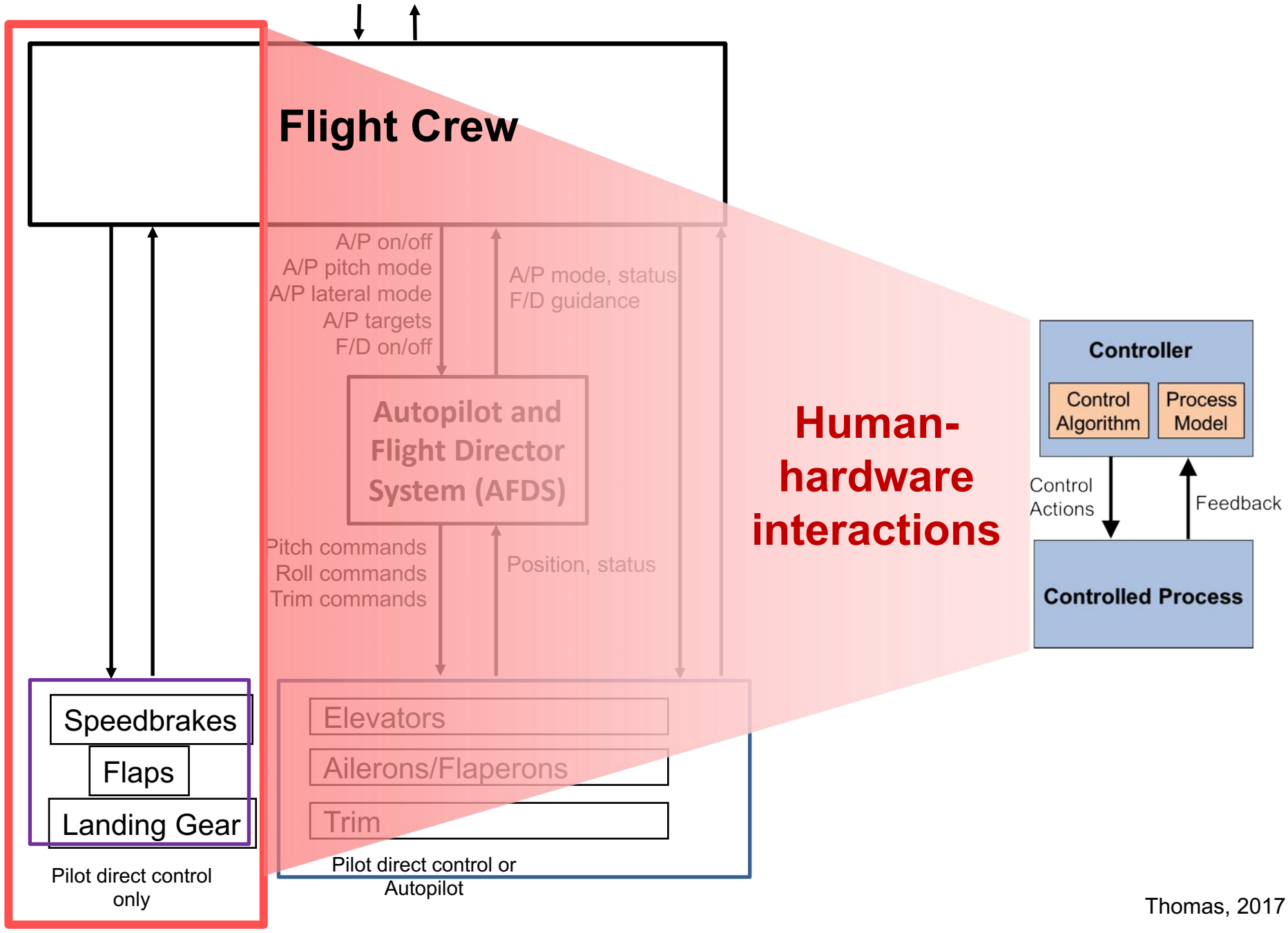

Pilot
Manage
Takeoff
Thrust
Orientation
Cabin environment
Position and heading
Taxi and landing
Movement on ground
etc.
Model of Automation
Model of Aircraft
Model of Airport (Environment)
Sensory and other Inputs
Environmental Inputs
Flight Commands
Feedback
A/C Automation (Flight Control Computer, FMS, etc.)
Control
Takeoff
Thrust
Orientation
Cabin environment
Position and heading
Taxi and landing
Movement on ground
etc.
Model of Aircraft
Control Commands
Feedback
Control Commands
Feedback
Aircraft


Flight Crew
A/P on/off
A/P pitch mode
A/P lateral mode
A/P targets
F/D on/off
A/P mode, status
F/D guidance
Autopilot and Flight Director System (AFDS)
Pitch commands
Roll commands
Trim commands
Position, status
Speedbrakes
Flaps
Landing Gear
Pilot direct control only
Elevators
Ailerons/Flaperons
Trim
Pilot direct control or Autopilot
Software-hardware interactions
Controller
Control Algorithm
Process Model
Control Actions
Feedback
Controlled Process

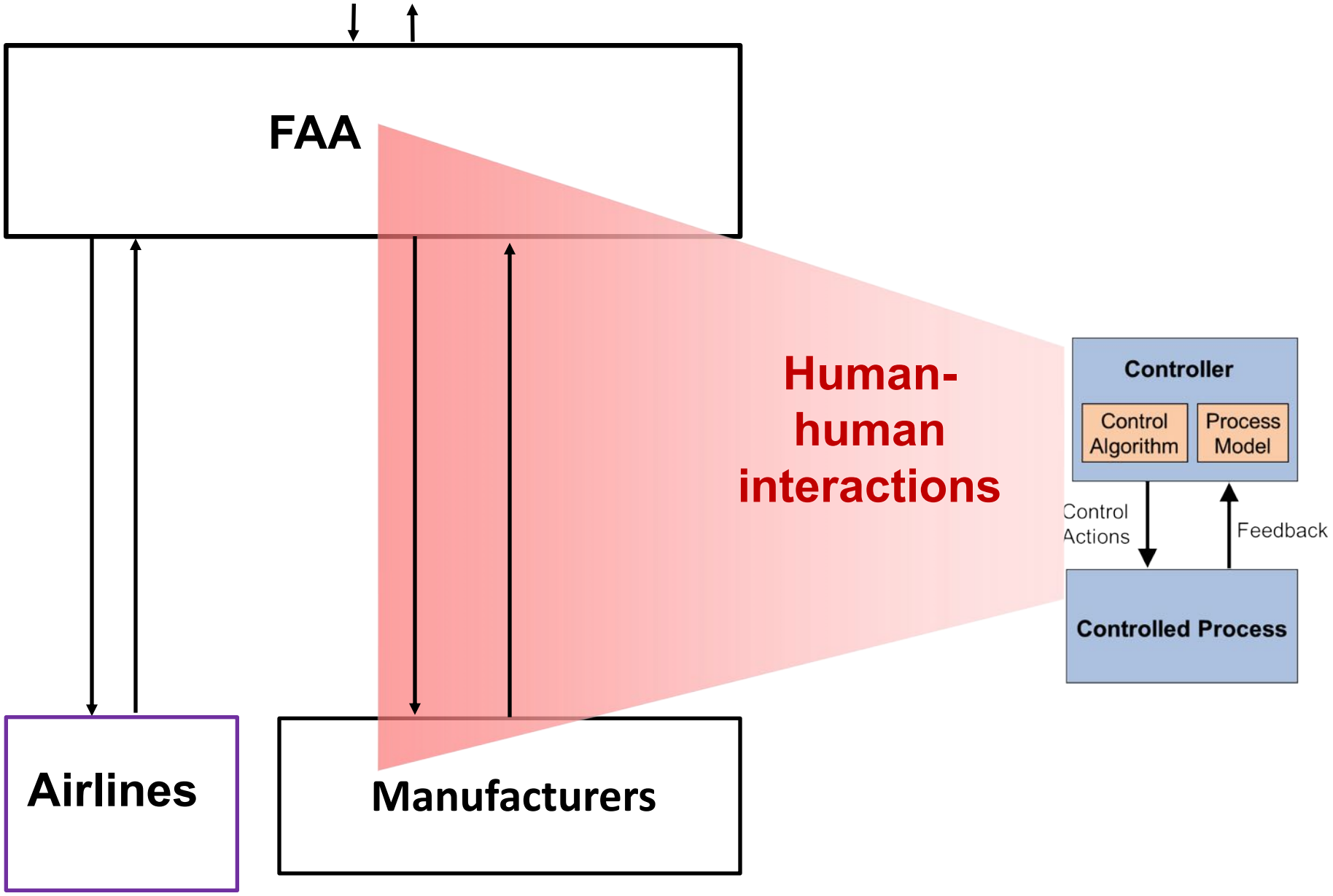
Thomas, 2017


Flight Crew
A/P on/off
A/P pitch mode
A/P lateral mode
A/P targets
F/D on/off
A/P mode, status
F/D guidance
Autopilot and Flight Director System (AFDS)
Pitch commands
Roll commands
Trim commands
Position, status
Speedbrakes
Flaps
Landing Gear
Elevators
Ailerons/Flaperons
Trim
Pilot direct control only
Pilot direct control or Autopilot
Human-automation interactions
Controller
Control Algorithm
Process Model
Control Actions
Feedback
Controlled Process

Thomas, 2017


Flight Crew
A/P on/off
A/P pitch mode
A/P lateral mode
A/P targets
F/D on/off
A/P mode, status
F/D guidance
Autopilot and Flight Director System (AFDS)
Pitch commands
Roll commands
Trim commands
Position, status
Speedbrakes
Flaps
Landing Gear
Pilot direct control only
Elevators
Ailerons/Flaperons
Trim
Pilot direct control or Autopilot
Human-hardware interactions
Controller
Control Algorithm
Process Model
Control Actions
Feedback
Controlled Process

Thomas, 2017


FAA
Human-human interactions
Controller
Control Algorithm
Process Model
Control Actions
Feedback
Controlled Process
Airlines
Manufacturers

Thomas, 2017

Establish Analysis Goals (Stakeholders)

- **Identify losses to be considered**

  L1. Death or serious injury to aircraft passengers or people in the area of the aircraft

  L2. “Unacceptable” damage to the aircraft or objects outside the aircraft

  L3: Financial losses resulting from delayed operations

  L4: Reduced profit due to damage to aircraft or airline reputation

- **Identify System-Level Hazards**

  H1: Insufficient thrust to maintain controlled flight

  H2: Loss of airframe integrity

  H3: Controlled flight into terrain

  H4: An aircraft on the ground comes too close to moving or stationary objects or inadvertently leaves the taxiway

  H5: etc.

Deceleration Hazards (H4)

H4-1: Inadequate aircraft deceleration upon landing, rejected takeoff, or taxiing

H4-2: Deceleration after the V1 point during takeoff

H4-3: Aircraft motion when the aircraft is parked

H4-4: Unintentional aircraft directional control (differential braking)

H4-5: Aircraft maneuvers out of safe regions (taxiways, runways, terminal gates, ramps, etc.)

H4-6: Main gear wheel rotation is not stopped when (continues after) the landing gear is retracted

High-Level (System) Requirements/Constraints

SC1: Forward motion must be retarded within TBD seconds of a braking command upon landing, rejected takeoff, or taxiing (H4-1).

SC2: The aircraft must not decelerate after V1 (H4-2).

SC3: Uncommanded movement must not occur when the aircraft is parked (H4-3).

SC4: Differential braking must not lead to loss of or unintended aircraft directional control (H4-4)

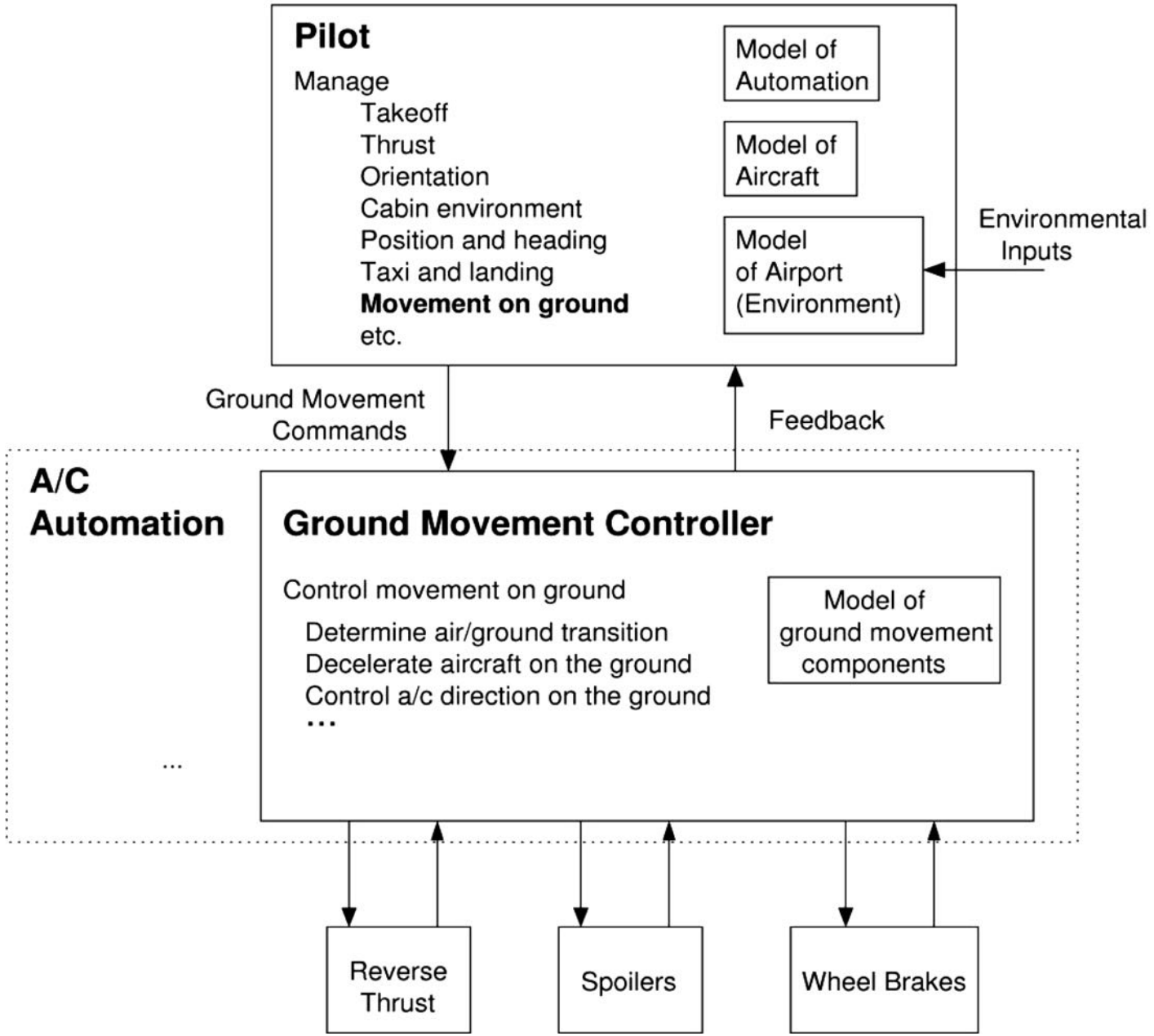
SC5: Aircraft must not unintentionally maneuver out of safe regions (taxiways, runways, terminal gates and ramps, etc.) (H4-5)

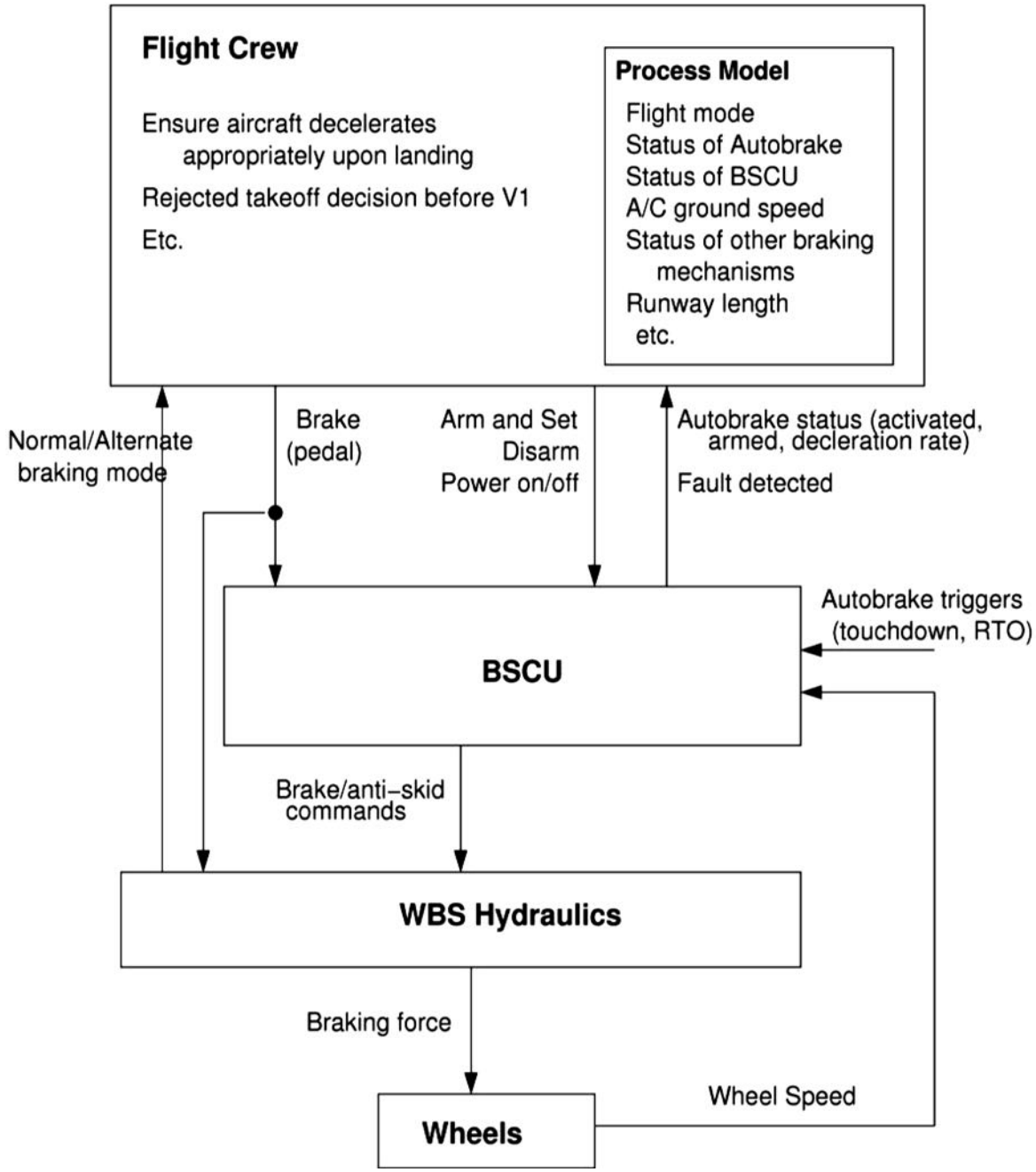
SC6: Main gear rotation must stop when the gear is retracted (H4-6)

STPA analysis will refine these into detailed requirements/constraints

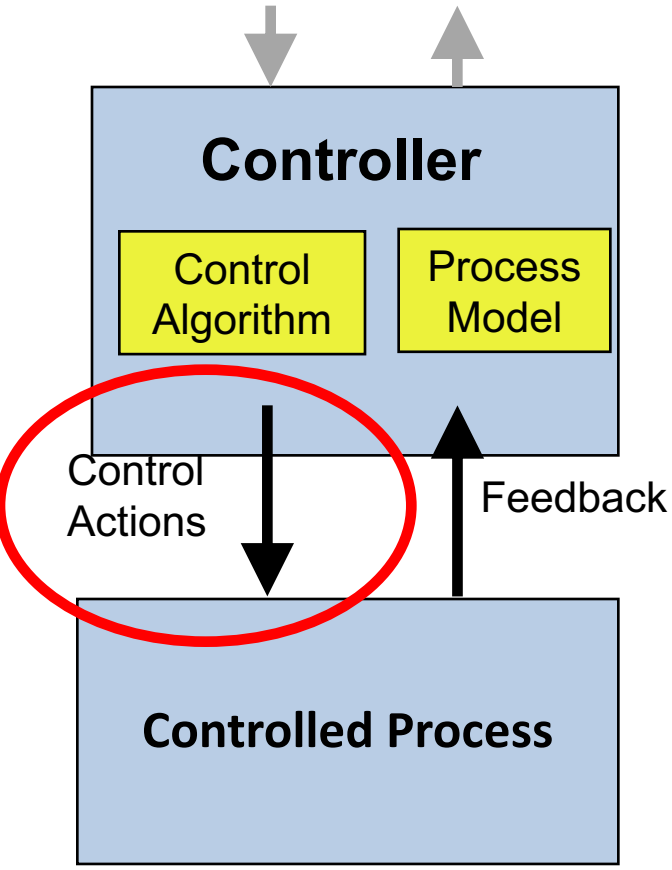
- **On system**
- **On components**

Construct a Functional Control Structure




Flight Crew
Ensure aircraft decelerates appropriately upon landing
Rejected takeoff decision before V1
Etc.
Process Model
Flight mode
Status of Autobrake
Status of BSCU
A/C ground speed
Status of other braking mechanisms
Runway length
etc.
Normal/Alternate braking mode
Brake (pedal)
Arm and Set
Disarm
Power on/off
Autobrake status (activated, armed, decleration rate)
Fault detected
Autobrake triggers (touchdown, RTO)
BSCU
Brake/anti-skid commands
WBS Hydraulics
Braking force
Wheel Speed
Wheels

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, or in wrong order
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?

Unsafe Control Actions for Crew

Control Action By Flight Crew:	Not providing causes hazard	Providing causes hazard	Too soon, too late, out of sequence	Stopped too soon, applied too long
CREW.1 Manual braking via brake pedals	CREW.1a1 Crew does not provide manual braking during landing, RTO, or taxiing when Autobrake is not providing braking (or insufficient braking), leading to overshoot [H4-1, H4-5]	CREW.1b1 Manual braking provided with insufficient pedal pressure, resulting inadequate deceleration during landing [H4-1, H4-5	CREW.1c1 Manual braking applied before touchdown causes wheel lockup, loss of control, tire burst [H4-1, H4-5]	CREW.1d1 Manual braking command is stopped before safe taxi speed (TBD) is reached, resulting in overspeed or overshoot [H4-1, H4-5]

Unsafe Control Actions by Autobraking

Control Action by BSCU	Not providing causes hazard	Providing causes hazard	Too soon, too late, out of sequence	Stopped too soon, applied too long
BSCU.1 Brake command	BSCU.1a1 Brake command not provided during RTO (to V1), resulting in inability to stop within available runway length [H4-1, H4-5]	BSCU.1b1 Braking commanded excessively during landing roll, resulting in rapid deceleration, loss of control, occupant injury [H4-1, H4-5]	BSCU.1c1 Braking commanded before touchdown, resulting in tire burst, loss of control, injury, other damage [H4-1, H4-5]	BSCU.1d1 Brake command stops during landing roll before taxi speed attained, causing reduced deceleration [H4-1, H4-5]

STPA-Generated Safety Requirements/Constraints

Unsafe Control Action	Description	Rationale
FC-R1	Crew must not provide manual braking before touchdown [CREW.1c1]	Could cause wheel lockup, loss of control, or tire burst
FC-R2	Crew must not stop manual braking more than TBD seconds before safe taxi speed reached [CREW.1d1]	Could result in overspeed or runway overshoot
FC-R3	The crew must not power off the BSCU during autobraking [CREW.4b1]	Autobraking will be disarmed
BSCU-R1	A brake command must always be provided during RTO [BSCU.1a1]	Could result in not stopping within the available runway length
BSCU-R2	Braking must never be commanded before touchdown [BSCU.1c1]	Could result in tire burst, loss of control, injury, or other damage
BSCU-R3	Wheels must be locked after takeoff and before landing gear retraction [BSCU.1a4]	Could result in reduced handling margins from wheel rotation in flight

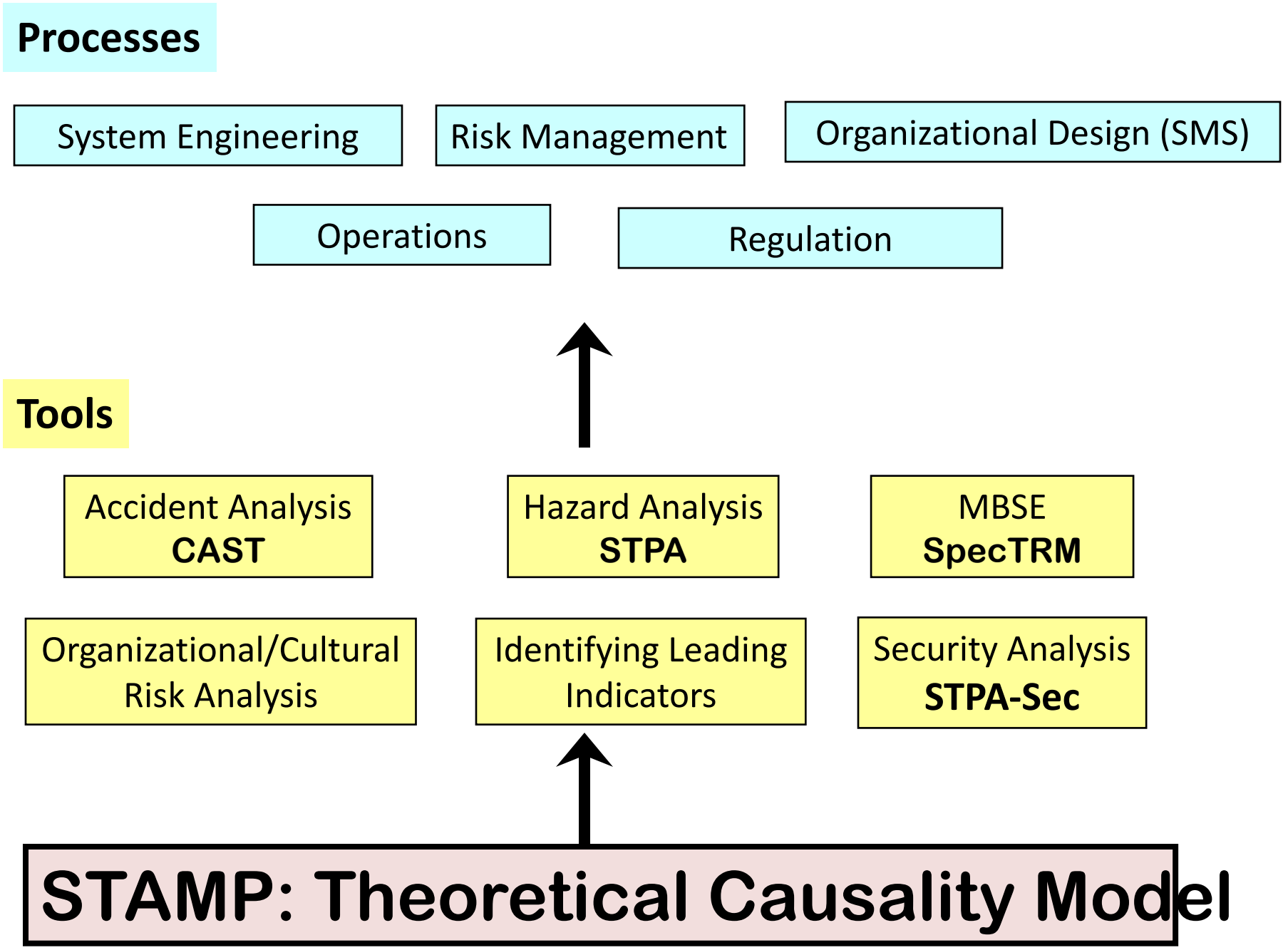
Generate Potential Causal Scenarios

BSCU.1a2: Brake command not provided during landing roll, resulting in insufficient deceleration and potential overshoot

Scenario 1: Autobrake believes the desired deceleration rate has already been achieved or exceeded (incorrect process model). The reasons Autobrake may have this process model flaw include:

- If wheel speed feedback influences the deceleration rate determined by the Autobrake controller, inadequate wheel speed feedback may cause this scenario. Rapid pulses in the feedback (e.g. wet runway, brakes pulsed by anti-skid) could make the actual aircraft speed difficult to detect and an incorrect aircraft speed might be assumed.
- Inadequate external speed/deceleration feedback could explain the incorrect Autobrake process model (e.g. inertial reference drift, calibration issues, sensor failure, etc.).

Possible Requirement for S1: Provide additional feedback to Autobrake to detect aircraft deceleration rate in the event of wheel slipping (e.g. fusion of multiple sensors)


Processes
System Engineering
Risk Management
Organizational Design (SMS)
Operations
Regulation
Tools
Accident Analysis
CAST
Hazard Analysis
STPA
MBSE
SpecTRM
Organizational/Cultural
Risk Analysis
Identifying Leading
Indicators
Security Analysis
STPA-Sec
STAMP: Theoretical Causality Model

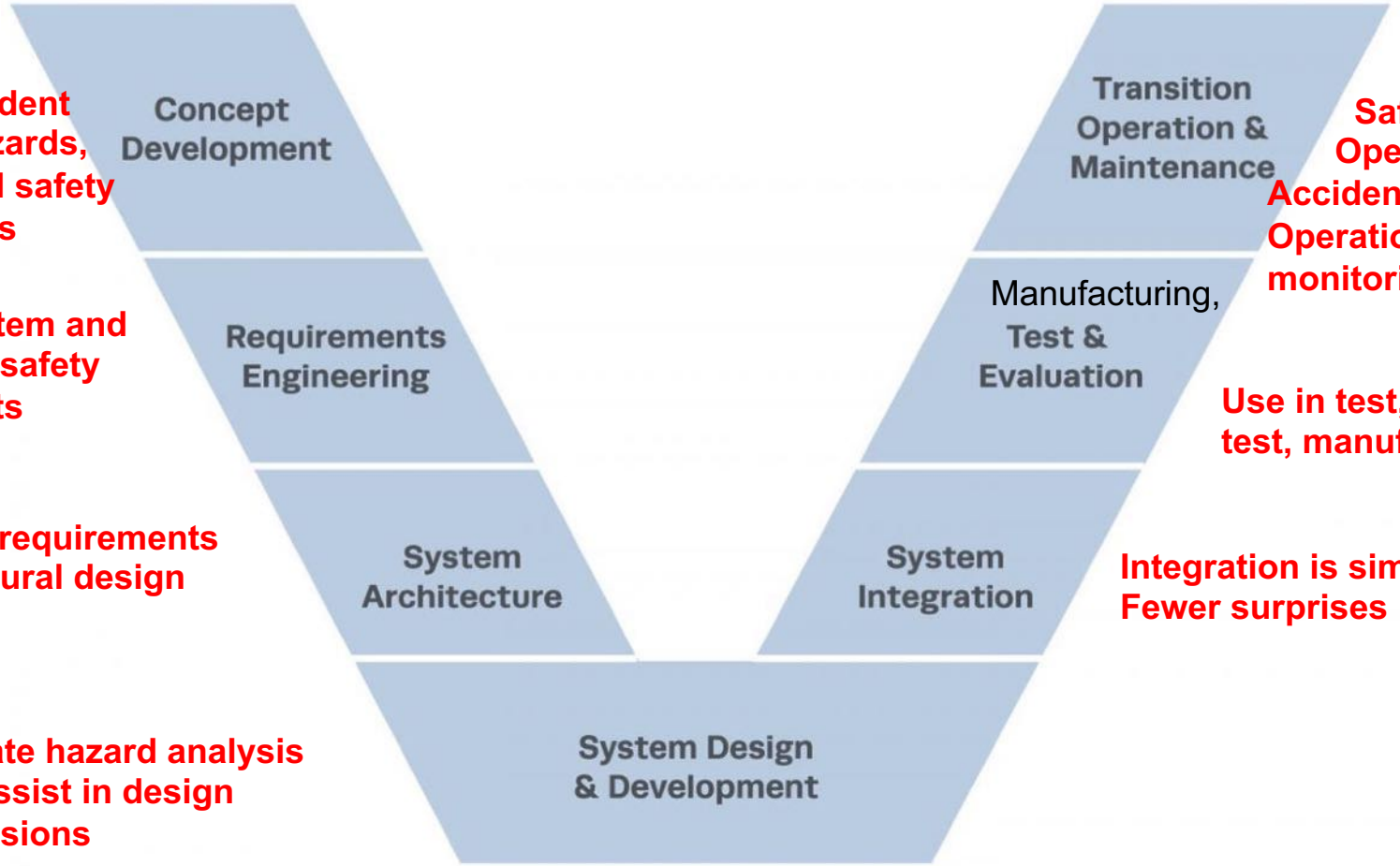
A Systems Approach to Safety

- Emphasizes building in safety rather than measuring it or adding it on to a nearly completed design
- Looks at system as a whole, not just components (a top-down holistic approach)
- Takes a larger view of causes than just failures
  - Based on system theory, not reliability theory
  - Accidents today are not just caused by component failures
- Goal is to use modeling and analysis to design and operate the system to be safe, not to predict the likelihood of a loss.
- Same analysis results can be used for cyber security

System Engineering Benefits

- Finds faulty underlying assumptions in concept development before flow downstream as anomalies (where more costly to change)
- Finds incomplete information, basis for further discussion with customer
- Provides quality, efficiency, security, and safety requirements/constraints before architecture and preliminary design begins
- Gives deeper insight into system vulnerabilities, particularly for cyber and human operator behavior.

STPA can be used throughout product development and operations



Identify accident (losses), hazards, system-level safety requirements

Identify system and component safety requirements

Use safety requirements In architectural design

Iterate hazard analysis to assist in design decisions

Integration is simplified Fewer surprises

Use in test, flight test, manufacturing

Safety in Operations Accident analysis, Operational safety monitoring

Is it Practical?

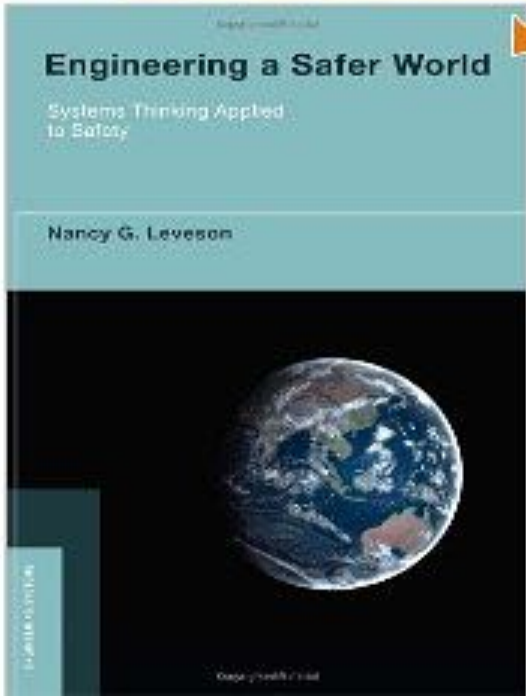
- STPA has been or is being used in a large variety of industries
  - Automobiles
  - Aircraft and Spacecraft
  - Air Traffic Control
  - UAVs (RPAs)
  - Defense systems
  - Medical Devices and Hospital Safety
  - Chemical plants
  - Oil and Gas
  - Nuclear and Electric Power
  - Robotic Manufacturing / Workplace Safety
  - Finance
  - etc.

Is it Effective?

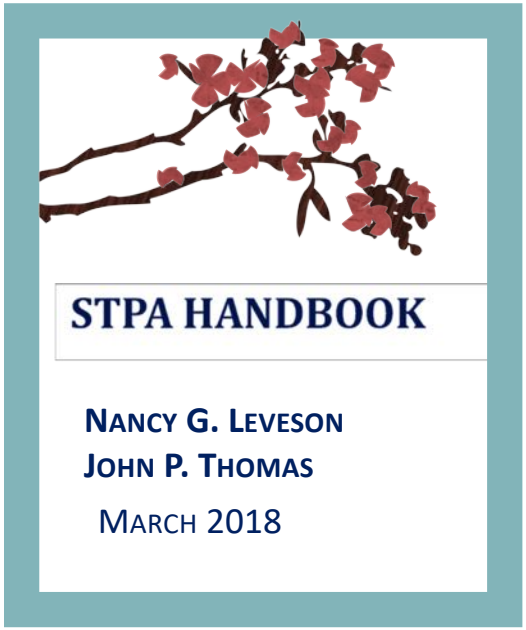
- Most of these systems are very complex (e.g., the new U.S. missile defense system)
- In all cases where a comparison was made (to FTA, HAZOP, FMEA, ETA, etc.)
  - STPA found the same hazard causes as the old methods
  - Plus it found more causes than traditional methods
  - In some evaluations, found 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
- ROI data (limited but mind blowing)

MIT STAMP/STPA Workshop

- 327 people from 32 countries registered
- Industry, academia, government
- Just about every safety-critical industry represented



MIT Press, 2012



http://psas.scripts.mit.edu

Questions?

Safety Control Structure for FMIS