



Common-Cause Analysis of Fault Trees with AI-Based Systems



John Shortle, Lance Sherry

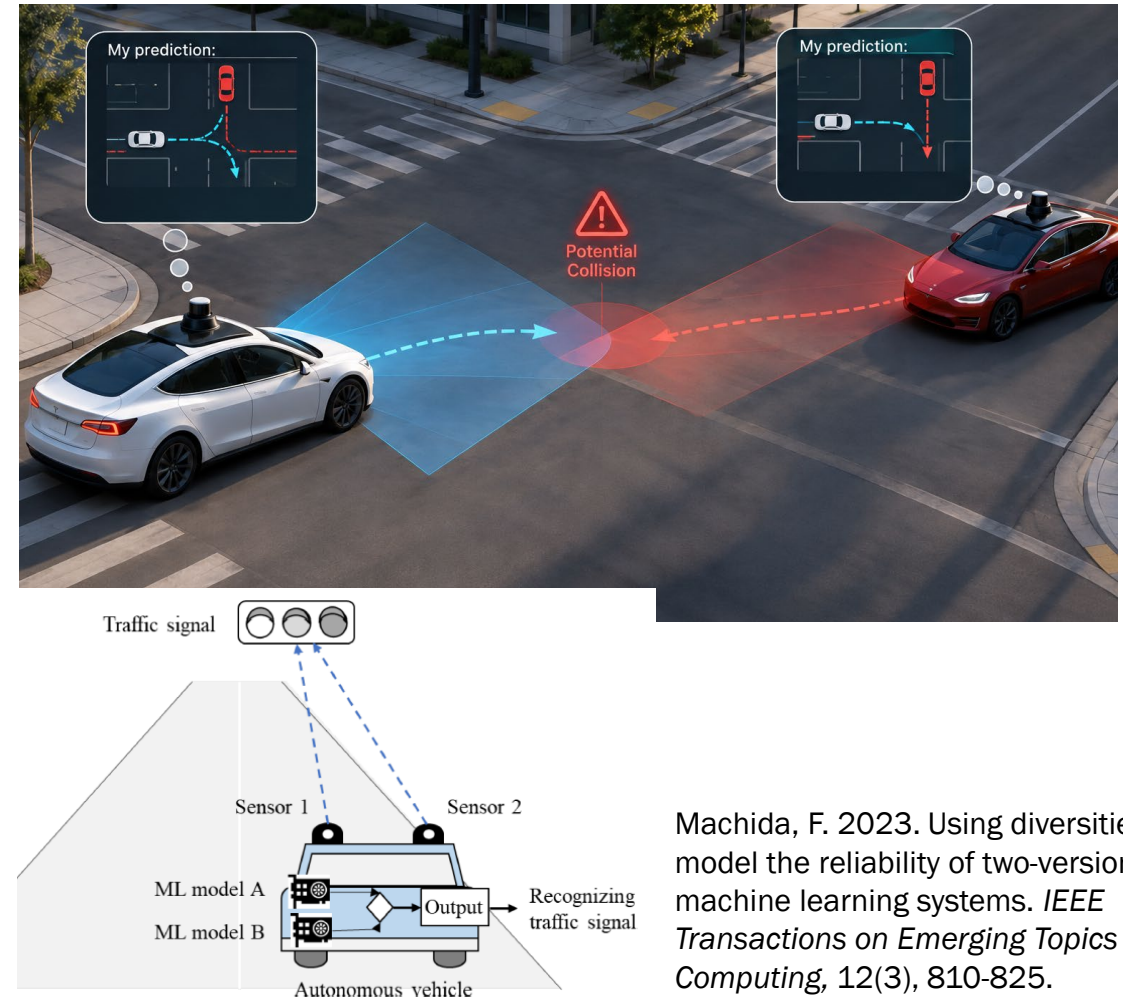
September 22, 2026

Risk Assessment with Correlation

- Common assumption in risk analysis is that component failures are independent
- Assumption might be made deliberately (analytic development easier), sometimes implicitly (assumption made without noting it)
- Systems with AI-based components have potential for correlated failures

Sources for Correlated Errors in AI Systems

- Sensor failures (camera, LiDAR, GPS)
- Environmental conditions (rain, fog, snow)
- Common out-of-distribution conditions
- Common training data and/or model architecture
- Common algorithms
- Cyberattacks
- Common sensor inputs

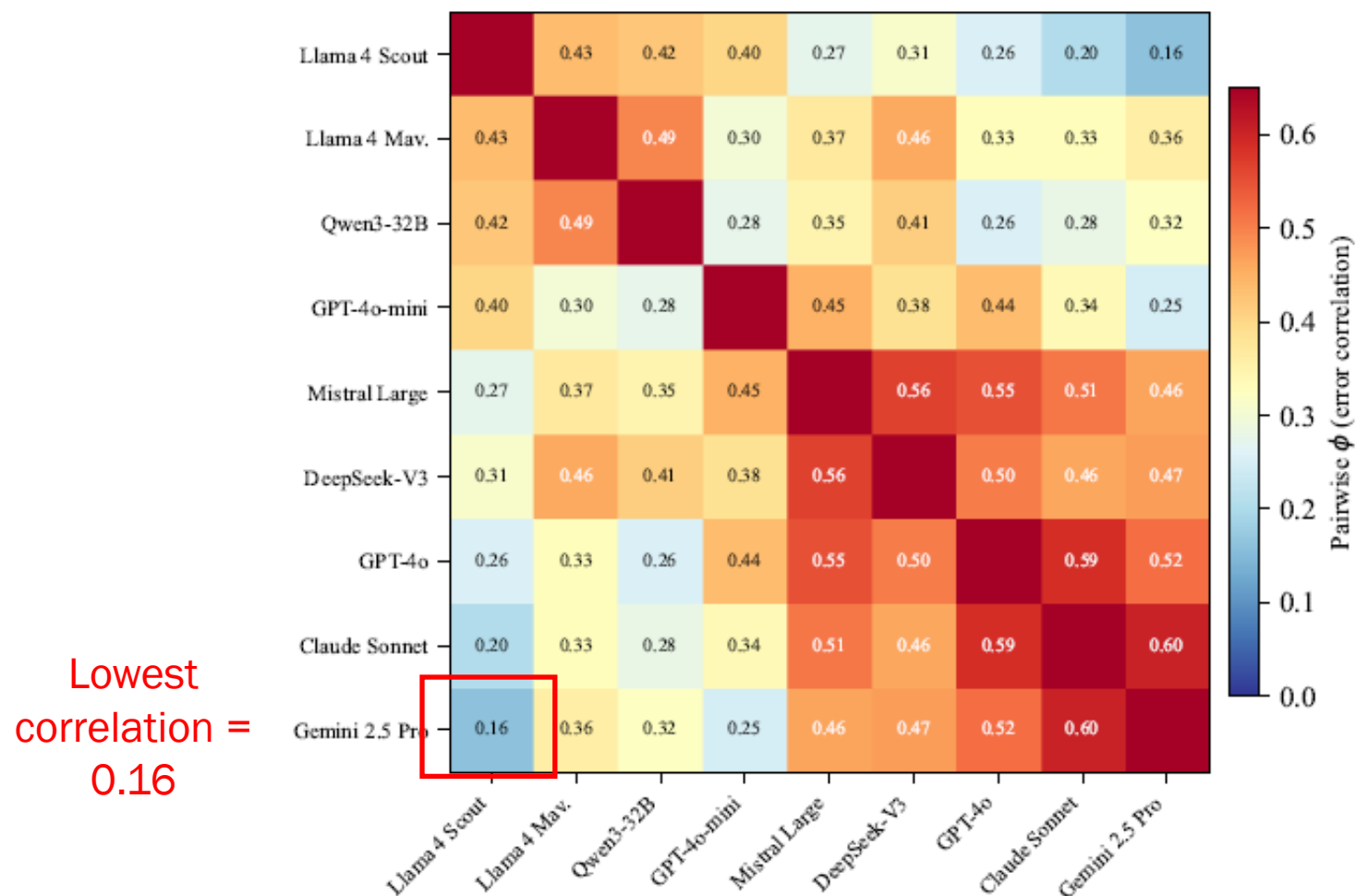


Machida, F. 2023. Using diversities to model the reliability of two-version machine learning systems. *IEEE Transactions on Emerging Topics in Computing*, 12(3), 810-825.

Correlated Errors in AI Systems

- Diversity in training data does not necessarily eliminate correlation (Kim et al., 2025):
 - Evaluated over 350 LLMs. On one benchmark, when two models were both wrong, they gave the same wrong answer 60% of the time (versus 33% for random wrong answer)
- Marginal value of adding “independent” judges is low (Kohli, 2026)
 - Nine frontier LLMs from seven companies behave as though they provide only about two independent votes
 - 75% of nominal independence lost through correlated errors
 - Panel does not outperform best judge

Error Correlations in AI Models

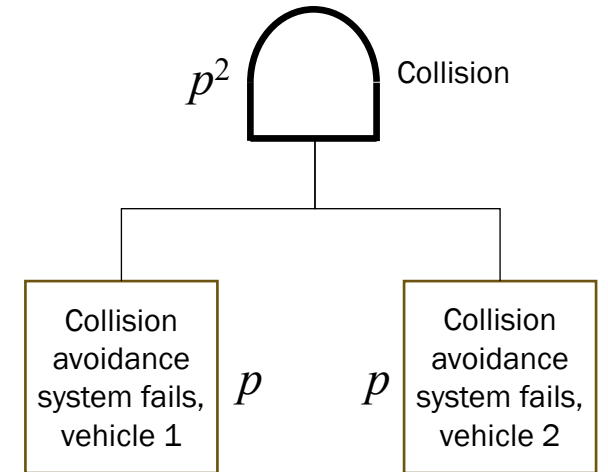


Correlation computed from error vector of each model ($e_{ij} = 1$ if model i makes error on question j ; $= 0$ otherwise)

Safety Critical AI-Based Systems



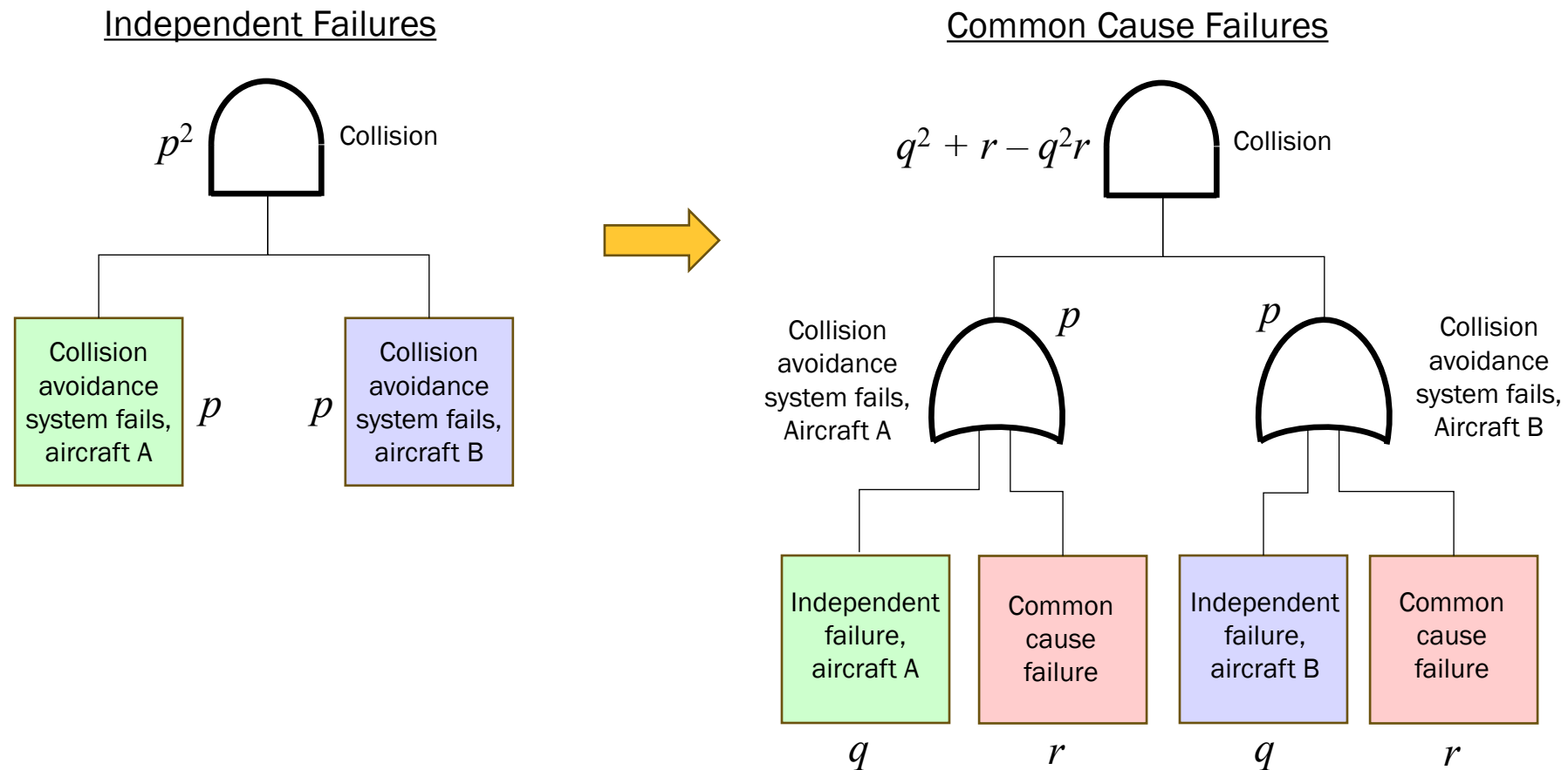
Independent Failures



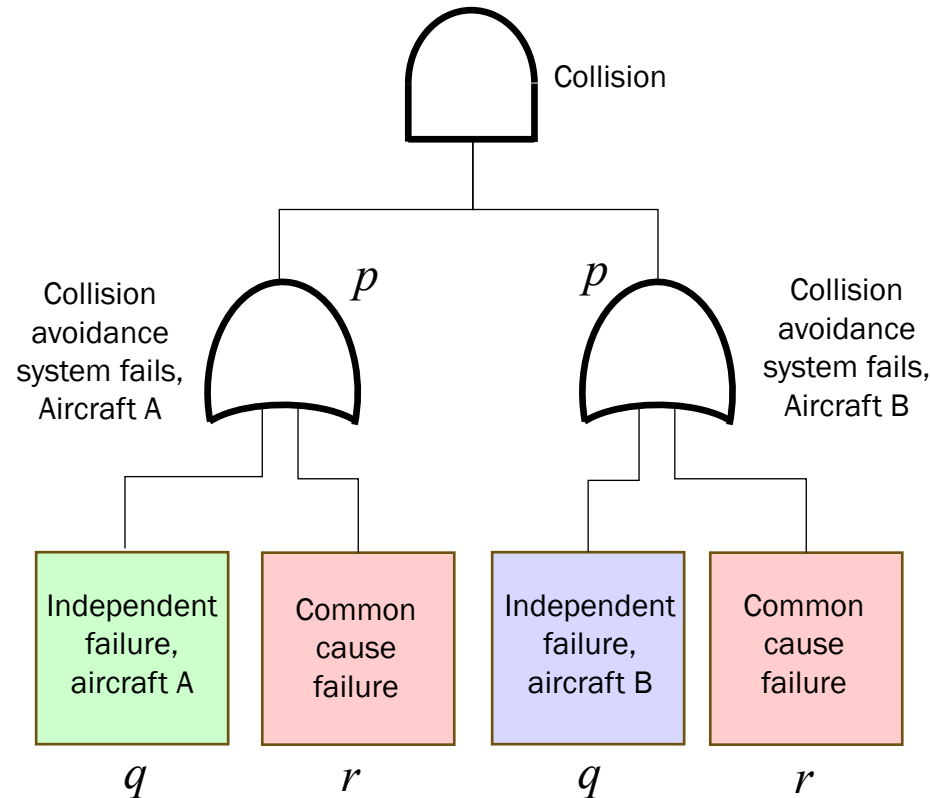
Collision occurs if both systems fail

E.g., if sub-system failure probability = 10^{-4} , then collision probability = 10^{-8}

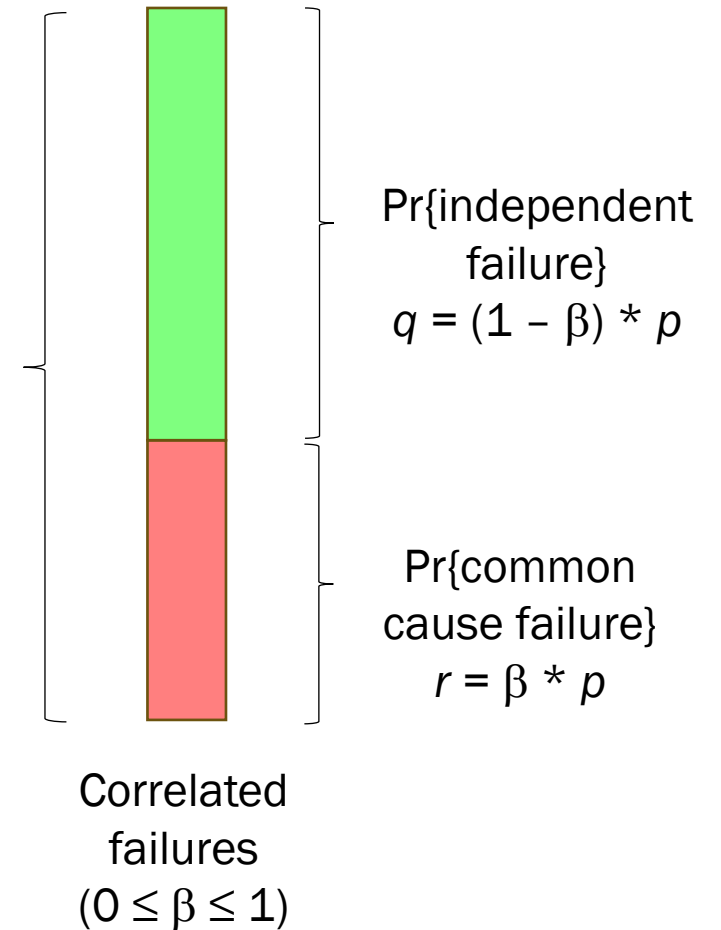
Common-Cause Failure (CCF) Model



CCF β -Factor Model



Probability of collision avoidance system failure = p

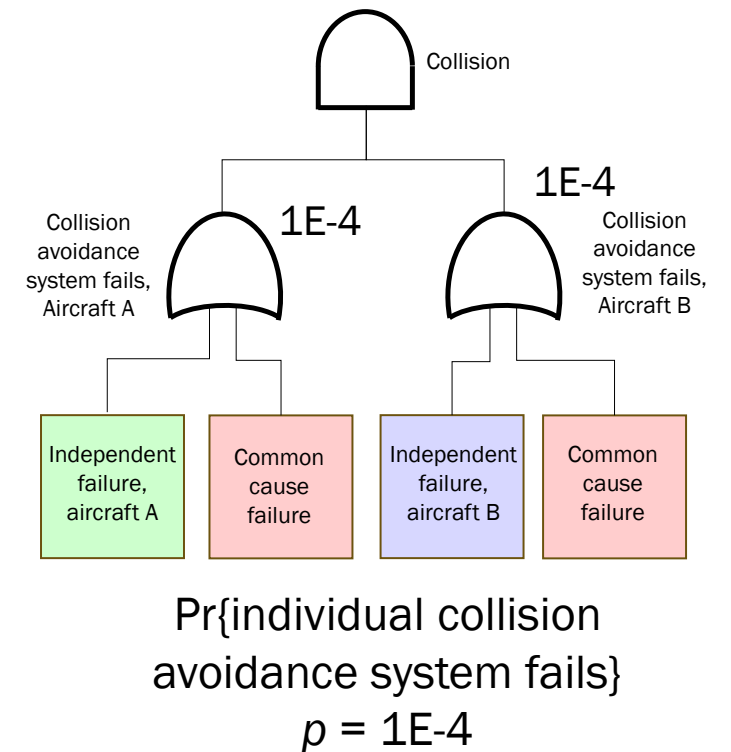
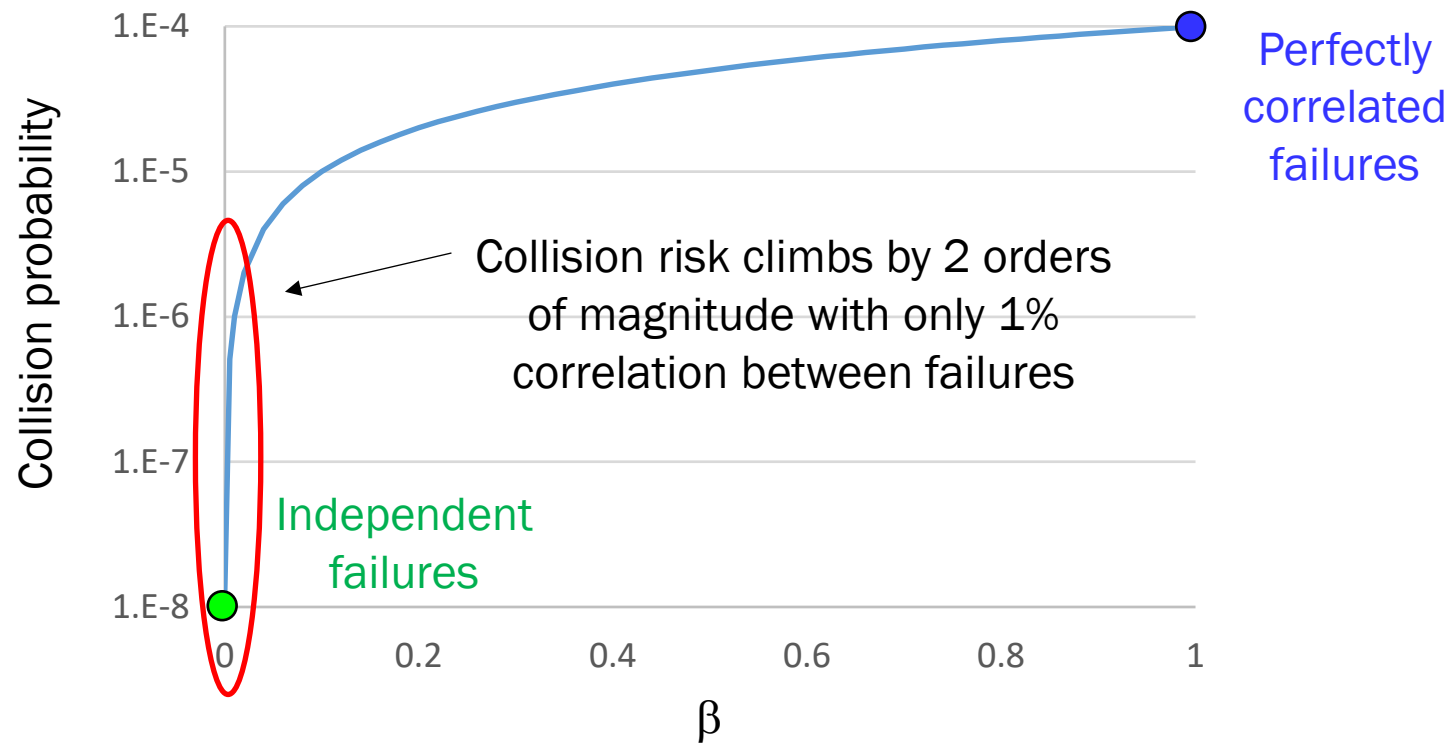


β = relative contribution of CCF to subsystem failure probability

$$\beta = \frac{r}{q + r}$$

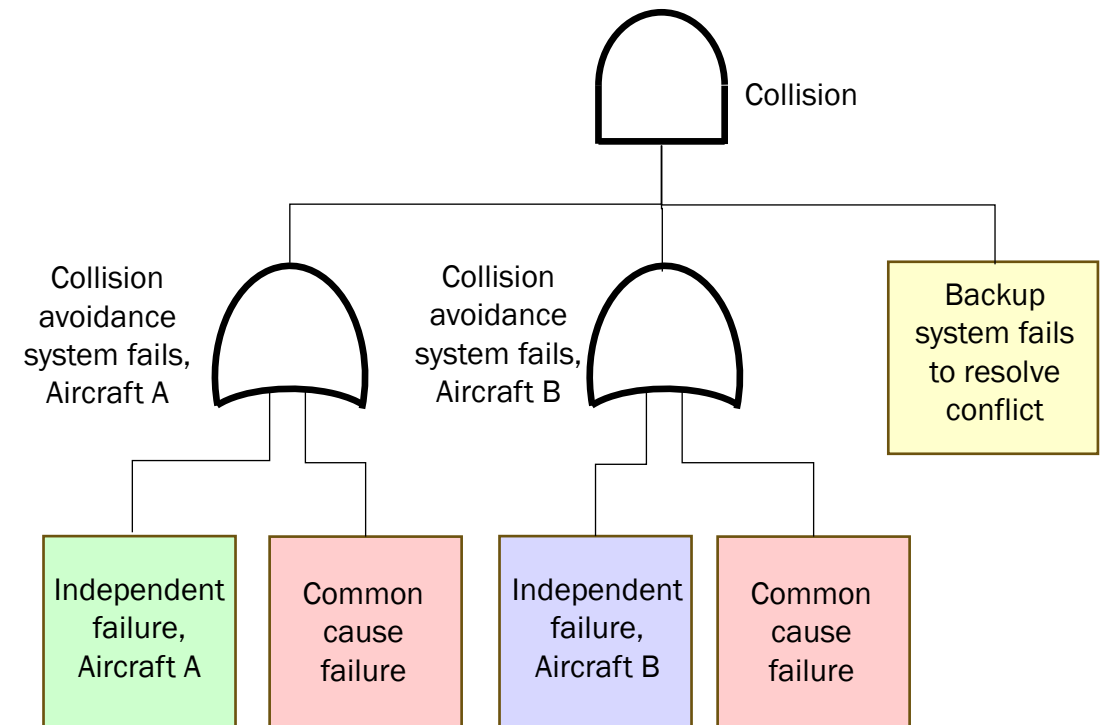
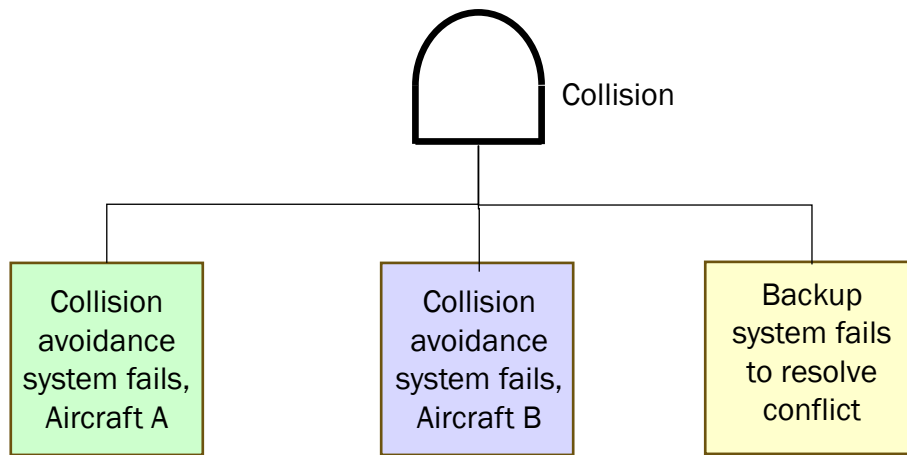
- Assumption: Failure probability of sub-system is fixed ($= p$)
- Varying β adjusts the relative contribution of the CCF to the subsystem failure, but does not change the sub-system failure probability itself

Potential Impact of CCFs

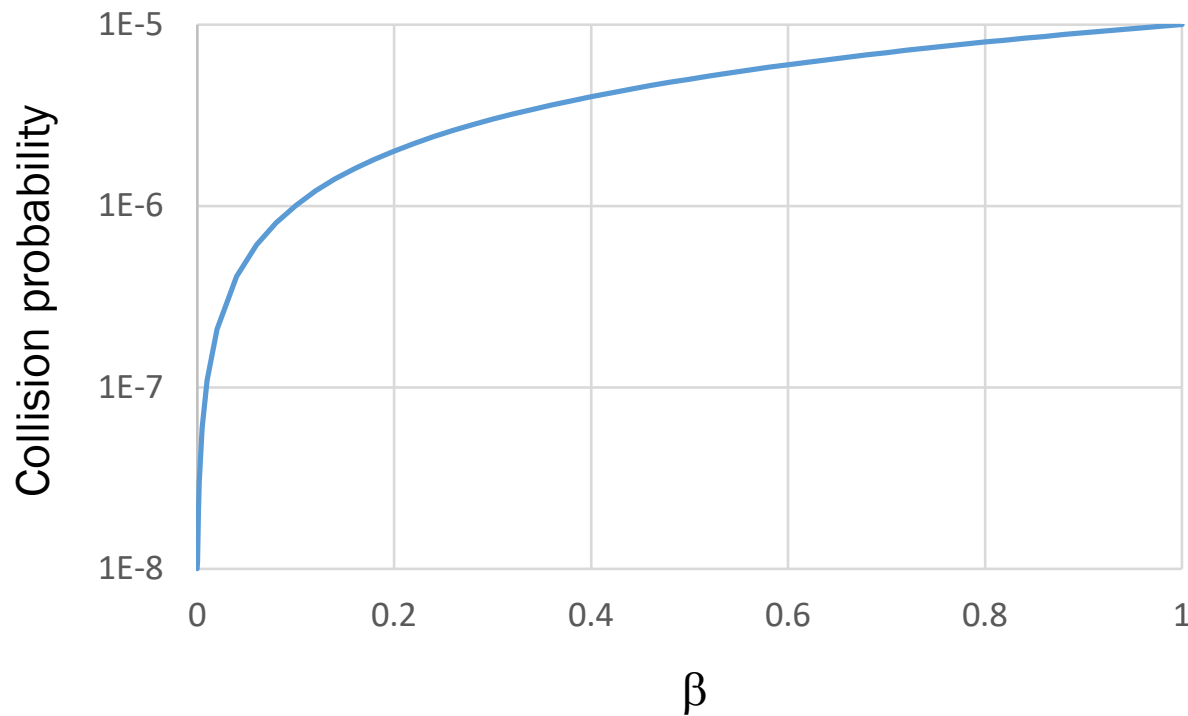


For rare events, β is approximately the correlation between the two failures

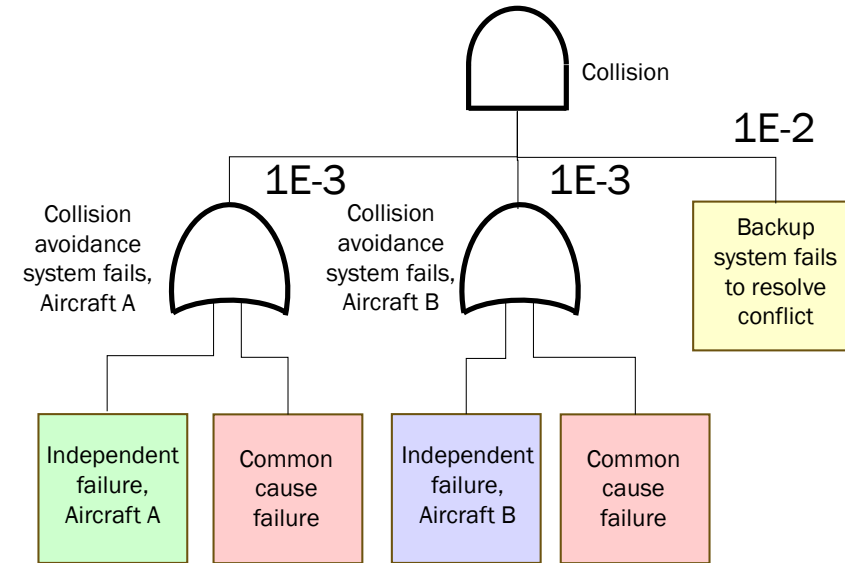
Collision Avoidance with Backup



Effect of Backup Systems



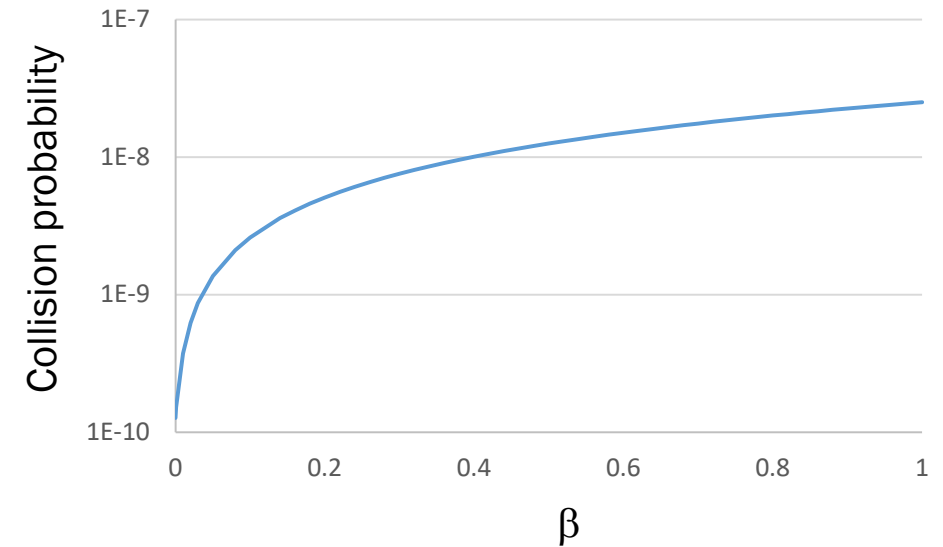
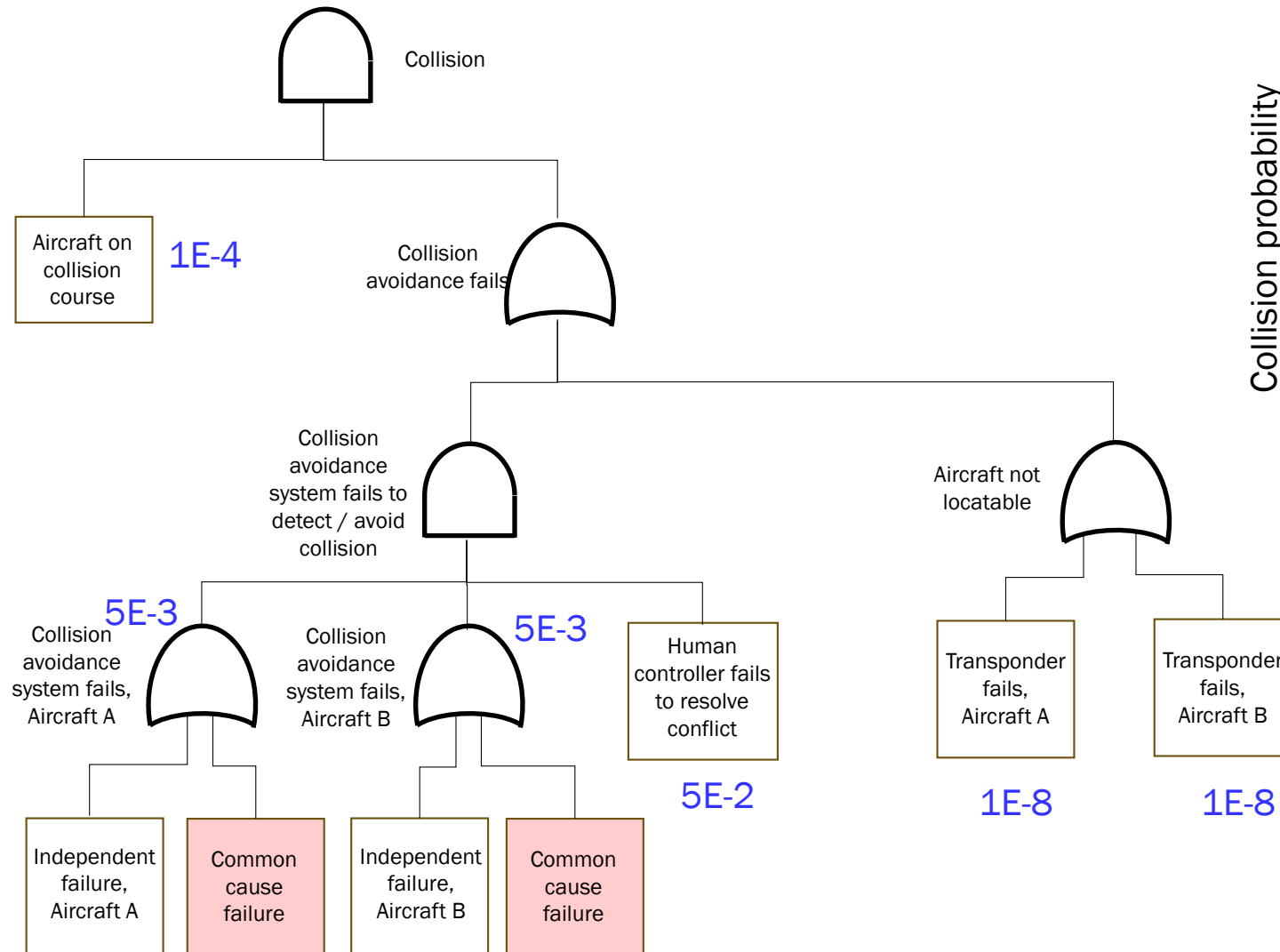
Backup system reduces effect of CCF, but still a potentially large increase in risk for small increase in β



$\Pr\{\text{individual collision avoidance system fails}\} = 1E-3$

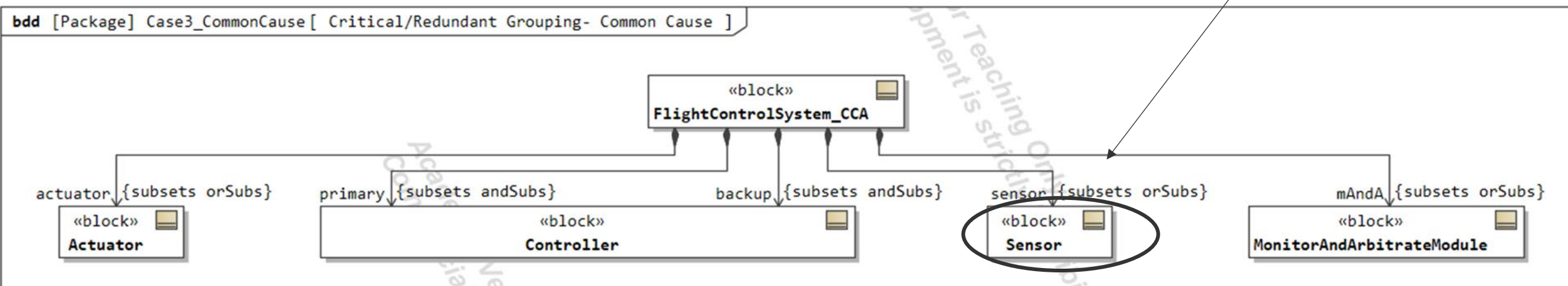
$\Pr\{\text{human controller fails}\} = 1E-2$

Example with Additional Components



CCF: SysML Implementation

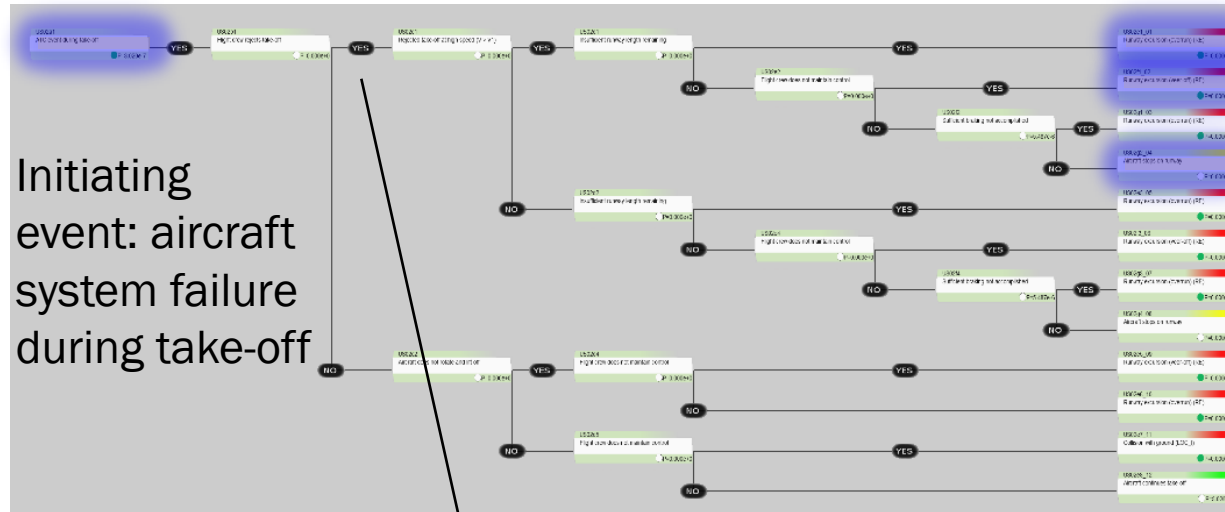
Assuming *dependent* sensors,
top-level probability is higher



#	Name	☒ leafNodeFailure... : Real	☒ failureProbabili... : Real	☒ orSubs : FaultTreePattern	☒ andSubs : FaultTreePattern
1	☒ ☒ FCS_CCA	1.0E-12	☒ 0.5968	☒ actuator : Structure::Case ☒ mAndA : Structure::Case ☒ sensor : Structure::Case3	☒ primary : Structure::Case ☒ backup : Structure::Case?
2	☒ actuator	0.1	0.1		
3	☒ mAndA	0.2	0.2		
4	☒ sensor	0.3	0.3		
5	☒ primary	0.4	0.4		
6	☒ backup	0.5	0.5		

Aviation Safety Model

Event sequence diagrams



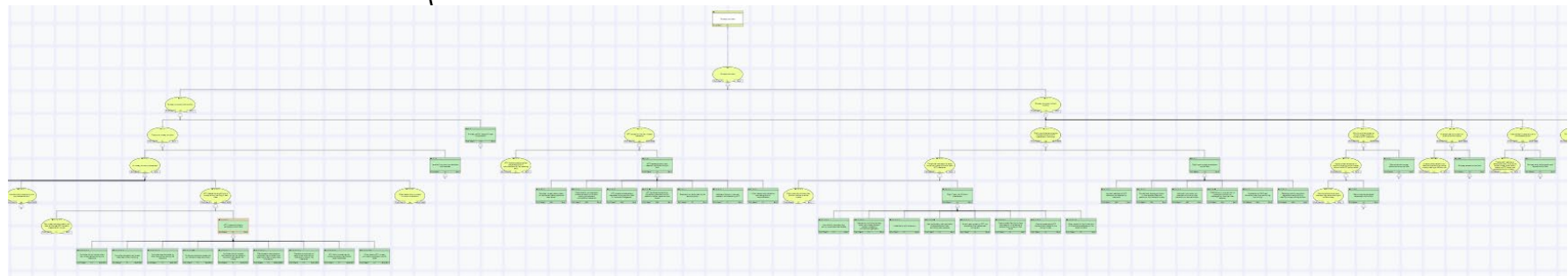
End events

Over-run

Veer-off

Aircraft stops on runway

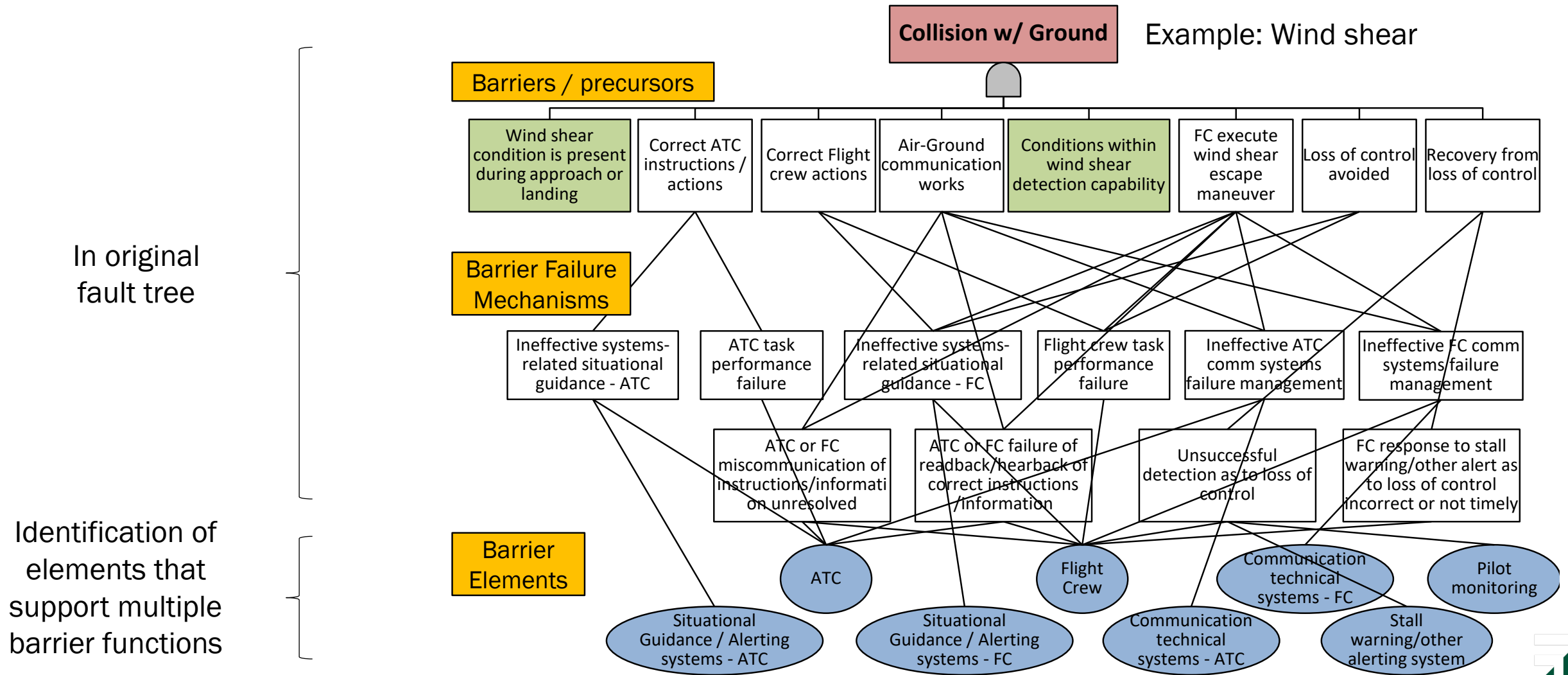
Supporting fault trees



Integrated Safety
Assessment Model
(ISAM, FAA)

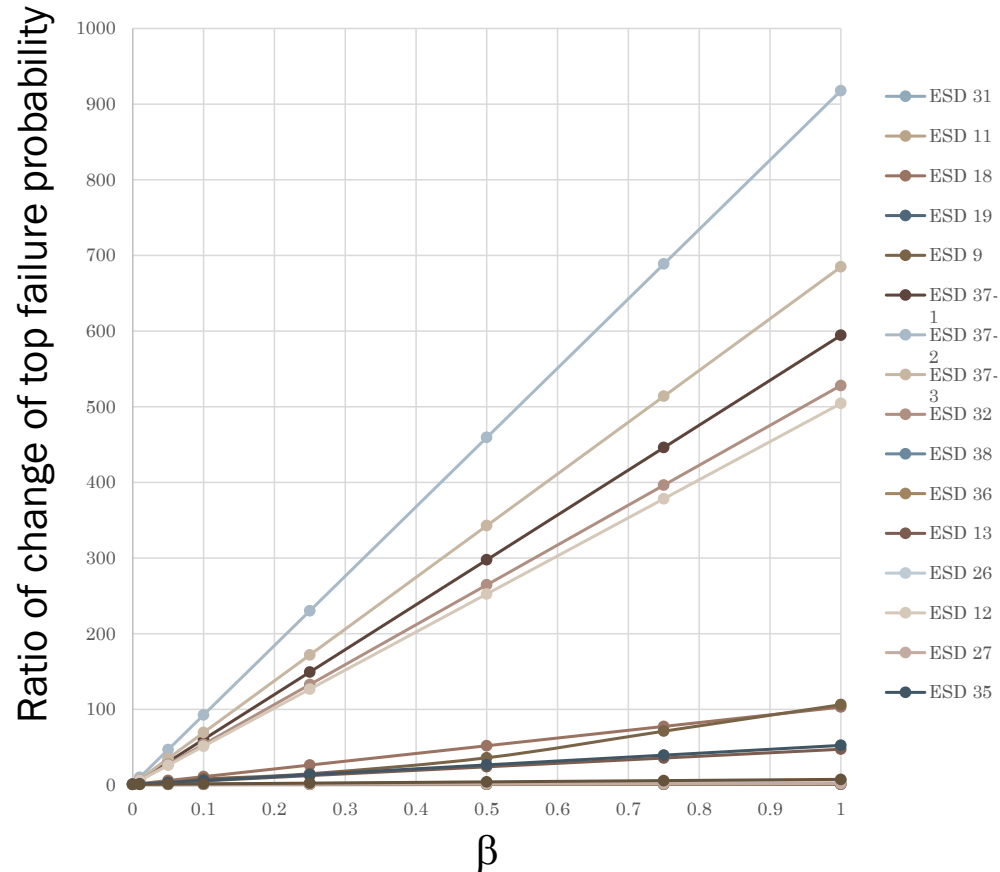
Initiating
event: aircraft
system failure
during take-off

Identification of Common Elements

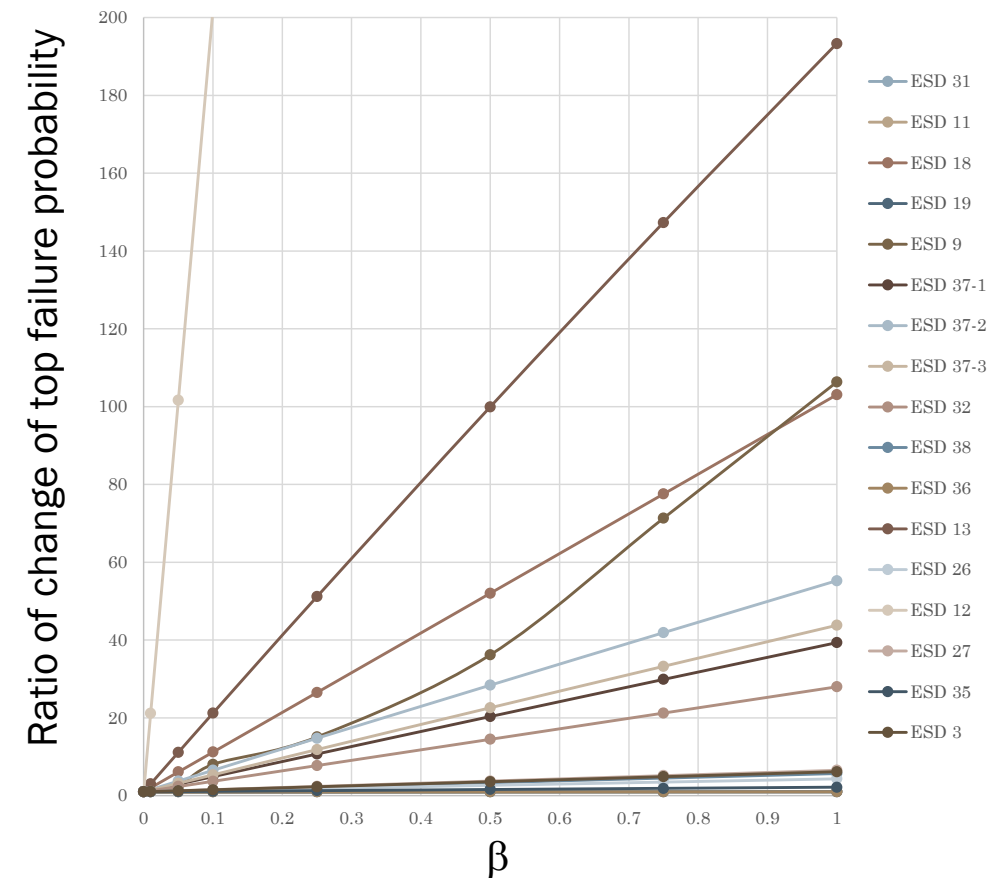


Effects of CCFs Related to Different Causes

CCFs related to human causes (air traffic control, flight crew, ground crew, etc.)



CCFs related to non-human causes (equipment failures)



Note: Scales on y-axis are different

CCF Importance Metrics

- Birnbaum importance metric. Measures change in system risk to change in CCF probability

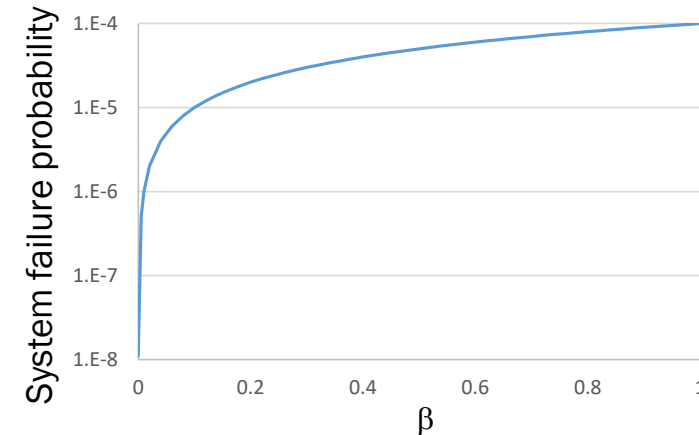
$$BI = \frac{\partial R(p_1, p_2, \dots, p_k, p_{CCF})}{\partial p_{CCF}}$$

- Criticality importance metric. Measures proportional change in system risk to proportional change in CCF probability

$$CI = \frac{p_{CCF}}{R(p_1, p_2, \dots, p_k, p_{CCF})} \cdot \frac{\partial R(p_1, p_2, \dots, p_k, p_{CCF})}{\partial p_{CCF}}$$

- Proposed metric. Measures order-of-magnitude change in system risk from baseline for small increase in β

$$OM = \frac{1}{R(p_1, p_2, \dots, p_k, \beta=0)} \cdot \frac{\partial R(p_1, p_2, \dots, p_k, \beta)}{\partial \beta}$$



OM = 1E4 $\rightarrow$ 2 orders of magnitude change in risk when $\beta = 1E-2$

- Pan, Z. & Nonaka, Y. (1995), "Importance analysis for the systems with common cause failures," Reliability Engineering & System Safety, 50(3), 297–300.
- Mi, J., Li, Y.-F., Beer, M., Broggi, M. & Cheng, Y. (2020), "Importance measure of probabilistic common cause failures under system hybrid uncertainty based on Bayesian network," *Eksploracja i Niezawodność – Maintenance and Reliability*, 22(1), 112–120.
- Zheng, J, H. Okamura, T. Dohi. 2018. Reliability importance of components in a real-time computing system with standby redundancy schemes, *International Journal of Mathematical, Engineering and Management Sciences*, 3(2), 64–89.

Summary

- AI-based systems have correlations that may be hard to eliminate
- The impact on safety critical systems can be large, even with small correlations
- Similar conclusions for simple architectures as well as large-scale fault trees (several hundred nodes)
- Criticality metrics provides assessment of vulnerability of system to CCF risks
- Ongoing work: SysML-based implementation



Questions?

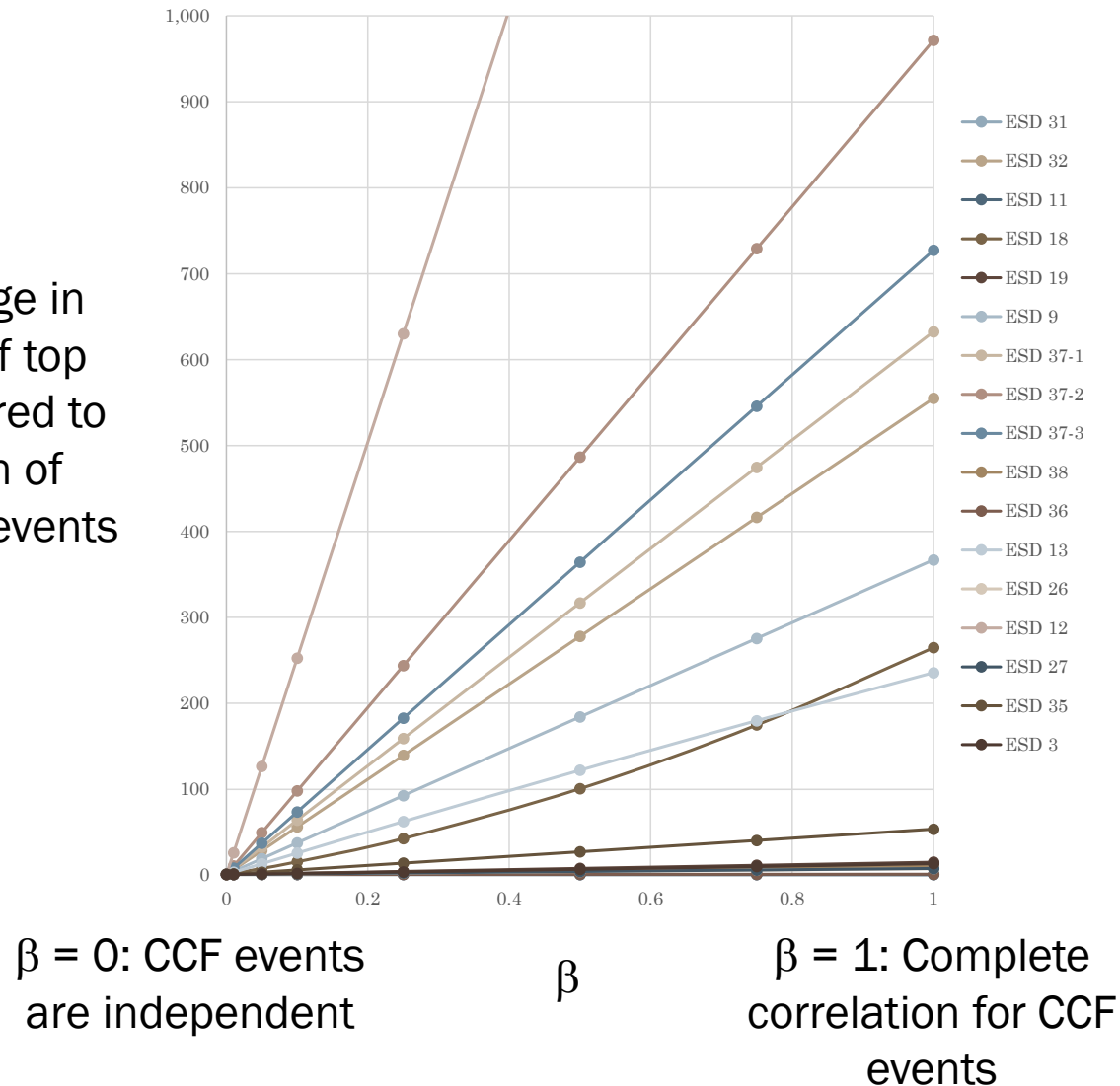


Event Sequence Diagrams Analyzed

ESD #	Initiating Event	Extreme Event
3	Aircraft directional control by flight crew inappropriate during take-off	Runway excursion
9	Single engine failure during take-off	Runway excursion
11	Fire onboard aircraft	Collision with ground
12	Flight crew member spatially disoriented	Collision with ground
13	Flight control system failure	Collision with ground
18	Single engine failure in flight	Collision with ground
19	Unstable approach	Collision with ground
26	Aircraft handling by flight crew inappropriate during landing roll	Runway excursion (veer-off)
27	Aircraft directional control related systems failure during landing roll	Runway excursion (veer-off)
31	Aircraft are positioned on collision course	Collision in mid-air
32	Runway incursion involving a conflict	Collision on runway
35	Potential conflict with terrain or obstacle	Controlled flight into terrain or obstacle
36	Conflict on taxiway or apron	Collision on taxiway or apron
37_1	Wake vortex encounter during departure	In-flight break-up
37_2	Wake vortex encounter during enroute	In-flight break-up
37_3	Wake vortex encounter during arrival	In-flight break-up
38	Loss of control inflight (LOC-I) due to inappropriate aircraft handling	Collision with ground

Potential Impact of CCFs

Factor change in probability of top event compared to assumption of independent events

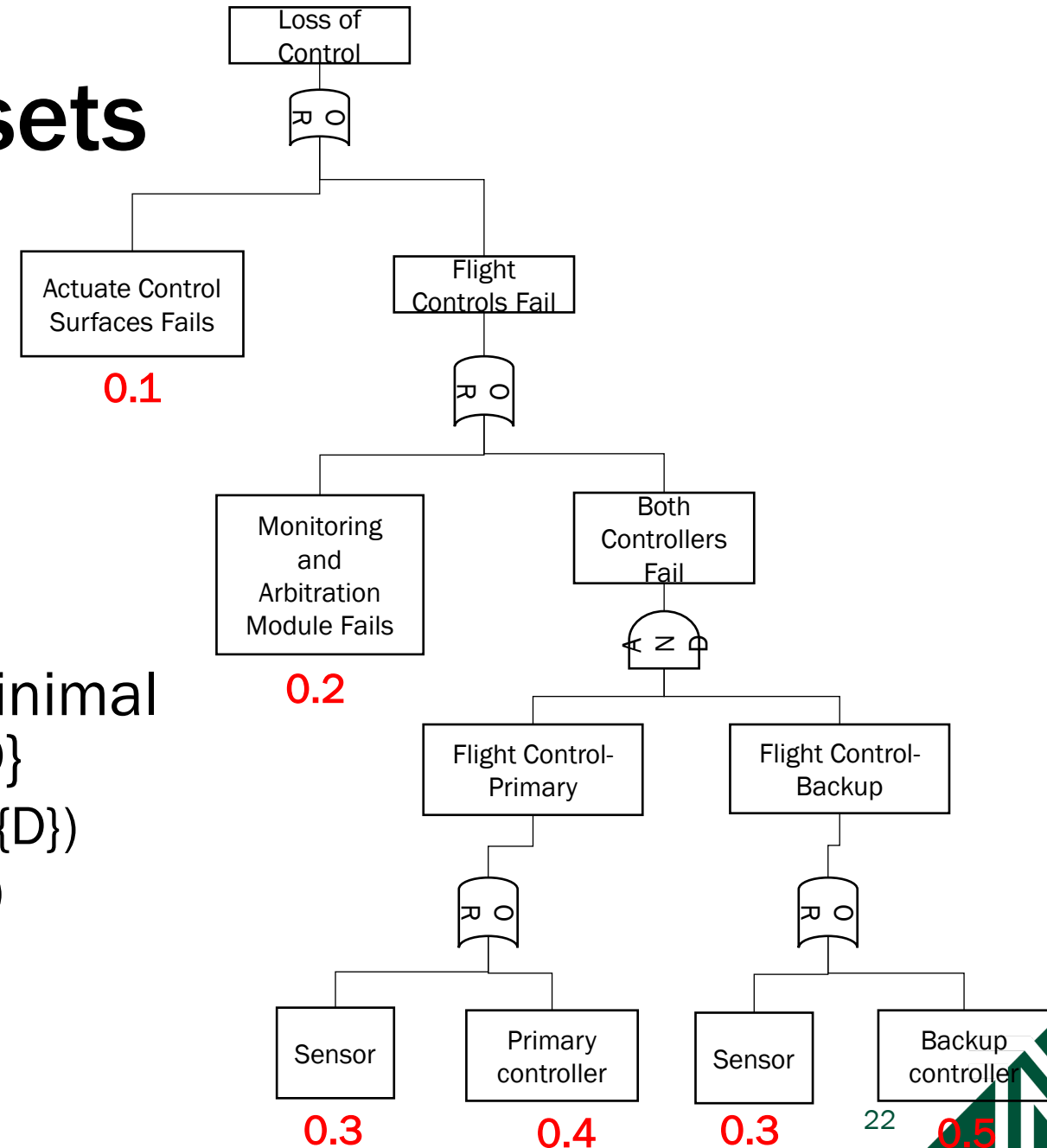


Evaluation of aviation safety event sequence diagrams (ESDs) with common cause failures

Nikdel, S., S. Noh, J. Shortle. 2021. Common cause failure analysis for aviation safety assessment models. *Proceedings of the 40th Digital Avionics Systems Conference*. San Antonio, TX.

Example: Minimal Cutsets

- Minimal cutsets of tree:
 - {A: actuator control surface fails}
 - {B: monitoring / arbitration fails}
 - {C: sensor fails}
 - {D: primary controller fails, backup controller fails}
- Probability of top event = $\Pr\{\text{any minimal cutset occurs}\} = \Pr\{A \text{ or } B \text{ or } C \text{ or } D\}$
 - $= 1 - (1 - \Pr\{A\})(1 - \Pr\{B\})(1 - \Pr\{C\})(1 - \Pr\{D\})$
 - $= 1 - (1 - 0.1)(1 - 0.2)(1 - 0.3)(1 - 0.4 * 0.5)$
 - $= .5968$



CCF β -Factor Model

Problem

- Fix p , failure probability of sub-system
- Choose β , relative contribution of CCF to sub-system failure

$$\beta = \frac{r}{q + r}$$

- Determine system failure probability as a function of p , β

Solution

- Solve for q from first equation

$$q = \frac{(1 - \beta) r}{\beta}$$

- Plug into OR relationship: $p = q + r - qr$

$$p = \frac{(1 - \beta)}{\beta} r + r - \frac{(1 - \beta)}{\beta} r^2$$

- Solve quadratic in r

$$r = \frac{1 - \sqrt{1 - 4\beta(1 - \beta)p}}{2(1 - \beta)}$$

- Plug in to get q

$$q = \frac{1 - \sqrt{1 - 4\beta(1 - \beta)p}}{2\beta}$$

- Plug q and r into $\Pr\{\text{system failure}\} = q^2 + r - q^2r$

