A BAYESIAN SOLUTION TO INCOMPLETENESS
IN
PROBABILISTIC RISK ASSESSMENT

14th International Probabilistic Safety Assessment & Management Conference

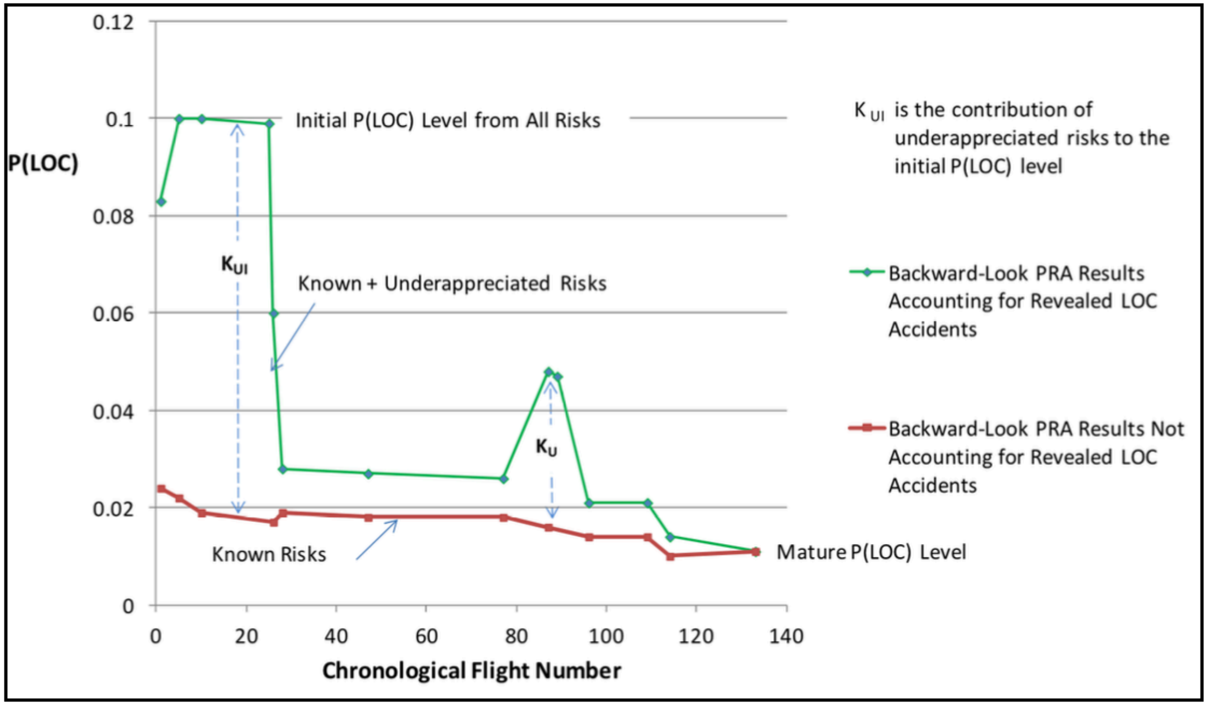
PSAM-14

September 17-21, 2018
Los Angeles, United States

Chris Everett, Information Systems Laboratories, New York, NY
Homayoon Dezfuli, NASA, Washington, DC

The Issue: PRA Incompleteness

- **The issue of incompleteness is a persistent challenge in PRA, where probability of failure is systematically underestimated**
- **PRA logic models typically represent only known accident causes, which can be just a small fraction of the total set of causes, especially for new systems**
- <u>Example:</u> Comparison between risk from known failure causes and total risk for the Space Shuttle*



*NASA/SP-2014-612, *NASA System Safety Handbook Vol. 2*, November 2014.

Diagnosis: PRA Answers the Wrong Question

- The question of interest is, *"What is the probability of failure?"*
- The question that PRA answers is, *"What is the probability of failure from known, modeled causes?"*
- So, instead of pretending that PRA *directly* answers the question of interest, we can *treat PRA results as evidence* that can be brought to bear on it
- How? ***Bayes' Theorem***

$$\underbrace{f_T(P_T|f_K(P_K))} \propto \underbrace{f_T(P_T)} \cdot \underbrace{L(\overbrace{f_K(P_K)}^{\text{PRA result}}|P_T)}$$

Posterior belief about the total probability of failure (P_T) given the PRA result ($f_K(P_K)$)

Prior belief about the total probability of failure (P_T)

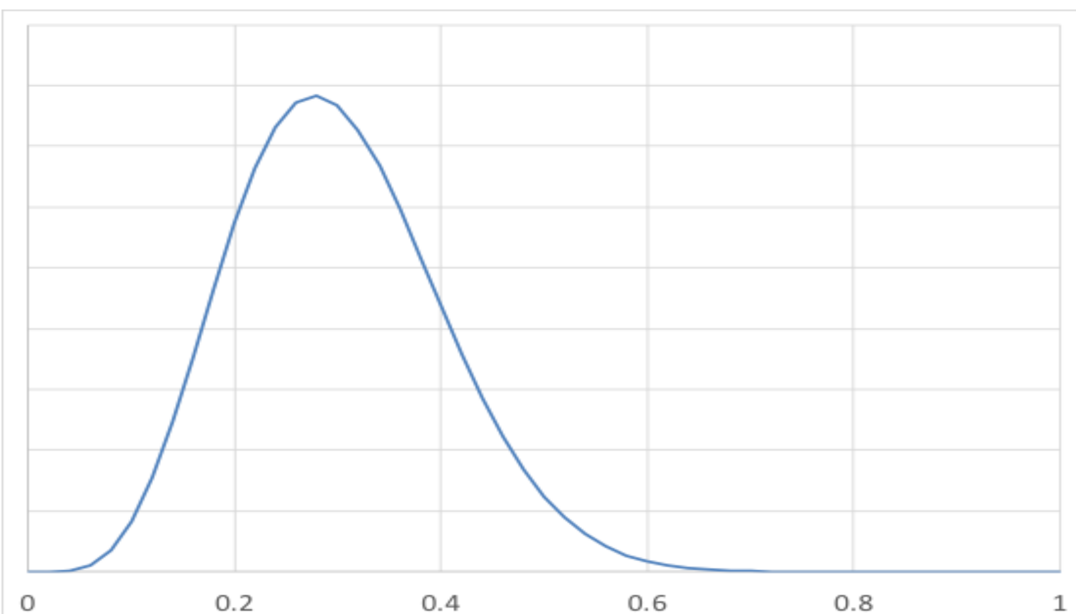
Likelihood that a given total probability of failure (P_T) would produce the observed PRA result ($f_K(P_K)$)

Quantifying Analysis Completeness

- **The driving factor behind the likelihood is *PRA incompleteness***
- **So, we introduce an *analysis completeness factor* C_A:**

$$C_A = P_K/P_T \quad \textit{analysis completeness factor}$$

- **We don't know C_A precisely, so we characterize it by a probability density function, $f_C(C_A)$**



Analysis Completeness Factor, C_A

$f_C(C_A)$ = beta(6.0, 14)

($\mu = 0.3$, $\sigma = 0.1$)

Roughly consistent with guidelines in the NASA System Safety Handbook Vol. 2 for new systems developed under moderate to significant time pressure, e.g., Space Shuttle

Constructing the Likelihood Function

- **The strategy used was to first develop the likelihood function $L(P_K|P_T, C_A)$, from which $L(f_K(P_K)|P_T, C_A)$ can then be constructed by treating $f_K(P_K)$ as the result of a large number *n* of individual samples P_{Ki}, each drawn from $f_K(P_K)$:**

$$L(f_K(P_K)|P_T, C_A) = L(P_{K1} \wedge P_{K2} \wedge \ldots \wedge P_{Kn}|P_T, C_A) = \prod_{i=1}^{n} [L(P_{Ki}|P_T, C_A)]$$

- **We impose the boundary condition:**

$$L(P_K|P_T, C_A) = f_K(P_K) \text{ when } C_A = 1$$

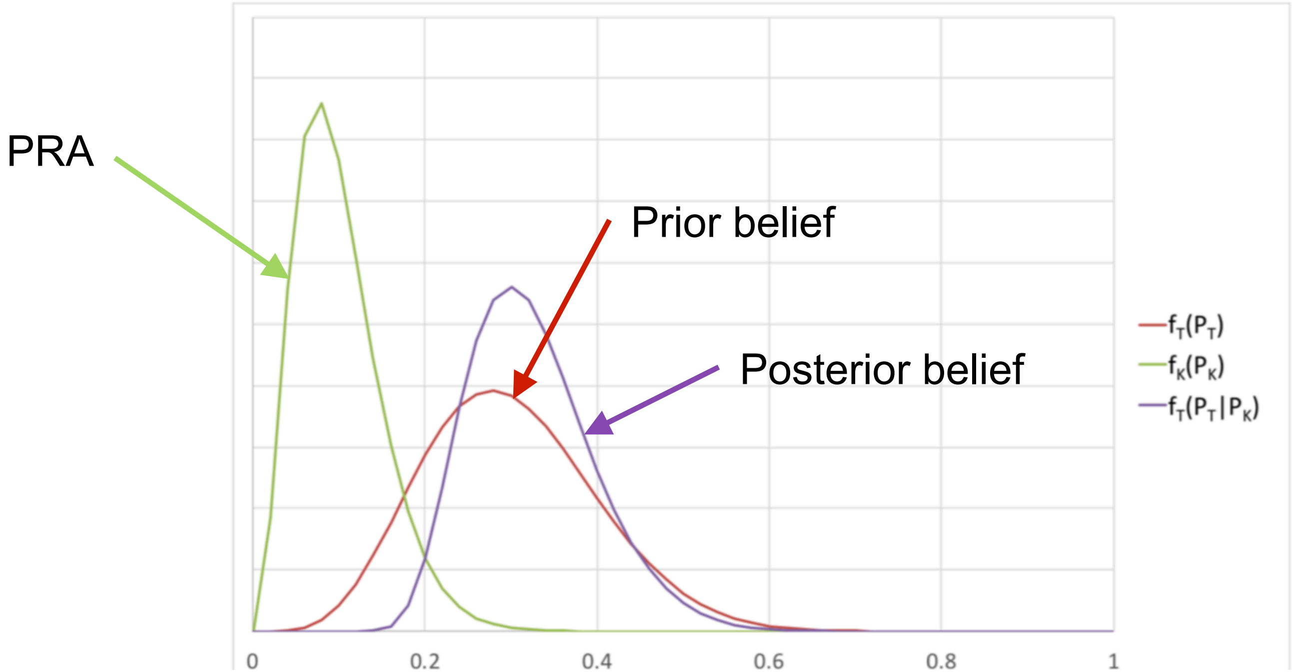
  - In other words, when we trust the PRA "completely" we believe its results

- **Integrating over $f_C(C_A)$ yields:**

$$\boxed{L(f_K(P_K)|P_T) = \int_0^{\infty} \{\prod_{i=1}^{n} [f_K(P_{Ki} - (E[P_K] - C_A \cdot P_T))] \cdot f_C(C_A)\} \cdot dC_A}$$

Proof of Concept

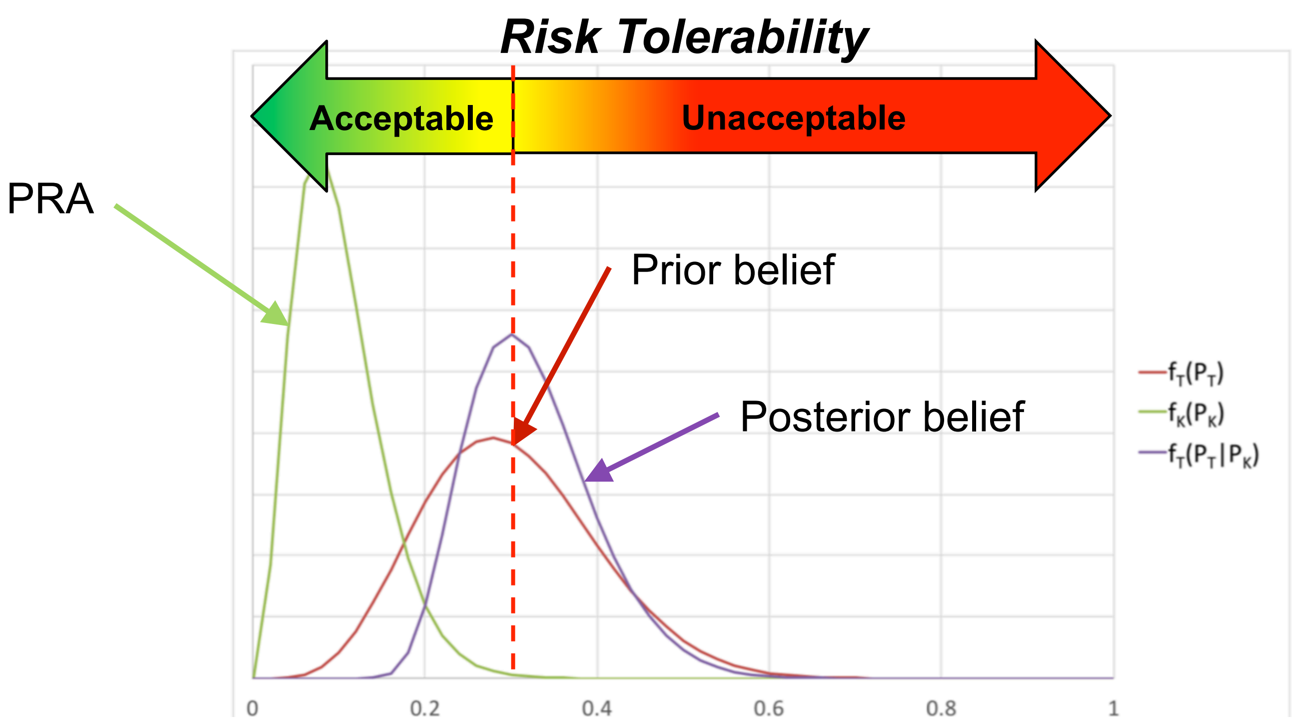
- **Given:**
  - Prior belief: $f_T(P_T)$ = beta(6, 14) (μ = 0.3, σ = 0.1)
  - Analysis completeness: $f_C(C_A)$ = beta(6.0, 14) (μ = 0.3, σ = 0.1)
  - PRA result: $f_K(P_K)$ = beta(3.5, 32) (μ = 0.1, σ = 0.05)
- **Posterior belief: $f_T(P_T|f_K(P_K)) \approx$ beta(11,23) (μ = 0.32, σ = 0.079)**



Updating of $f_T(P_T)$ with $f_K(P_K)$

Why Does It Matter?

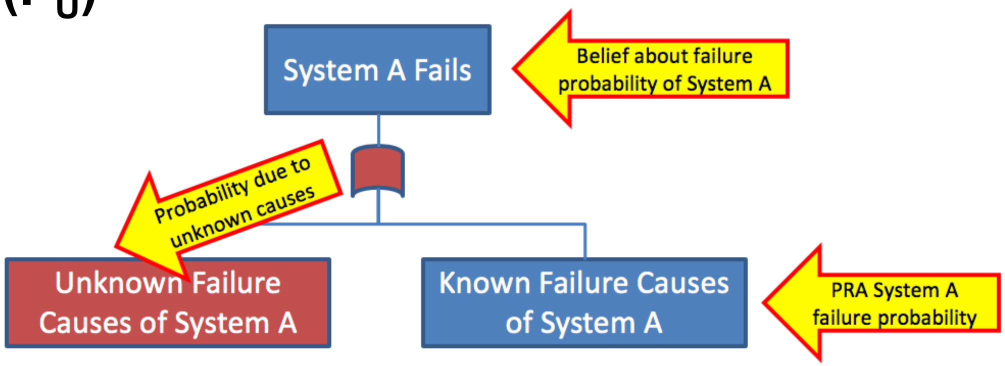
- **One reason: Risk Acceptance**
  - Belief that PRA characterizes the total probability of failure, $f_T(P_T)$, can lead to poor risk acceptance decisions
    - PRA suggests that the risk is ***acceptable*** with very high confidence
    - Bayesian analysis shows that the risk is likely to be ***unacceptable***



Updating of $f_T(P_T)$ with $f_K(P_K)$

Quantification of Unknown Failure Causes

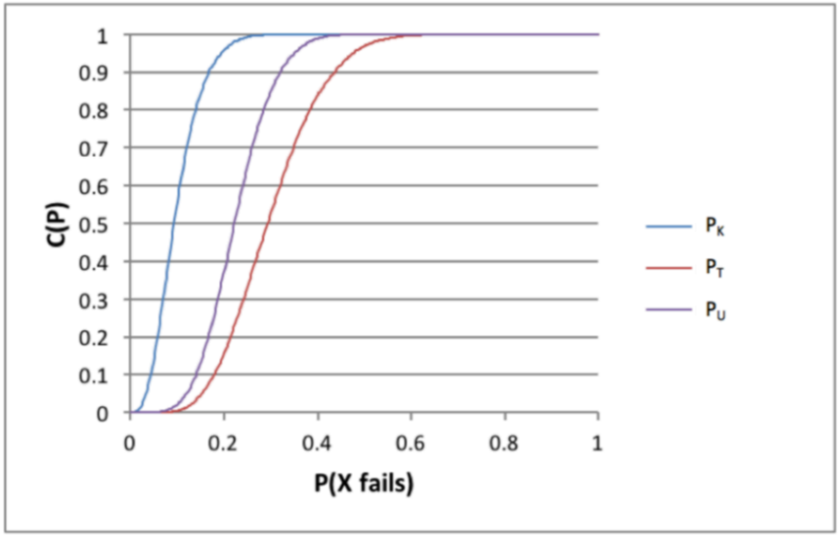
- **Given the PRA results and our posterior belief $f_T(P_T|f_K(P_K))$, we can infer the probability of failure due to unknown/unmodeled causes, $f_U(P_U)$**



- **This is a non-trivial problem in the general case (de-convolution), but for P_U, P_K correlated is trivial:**

$$P_U = (P_T - P_K)/(1 - P_K)$$

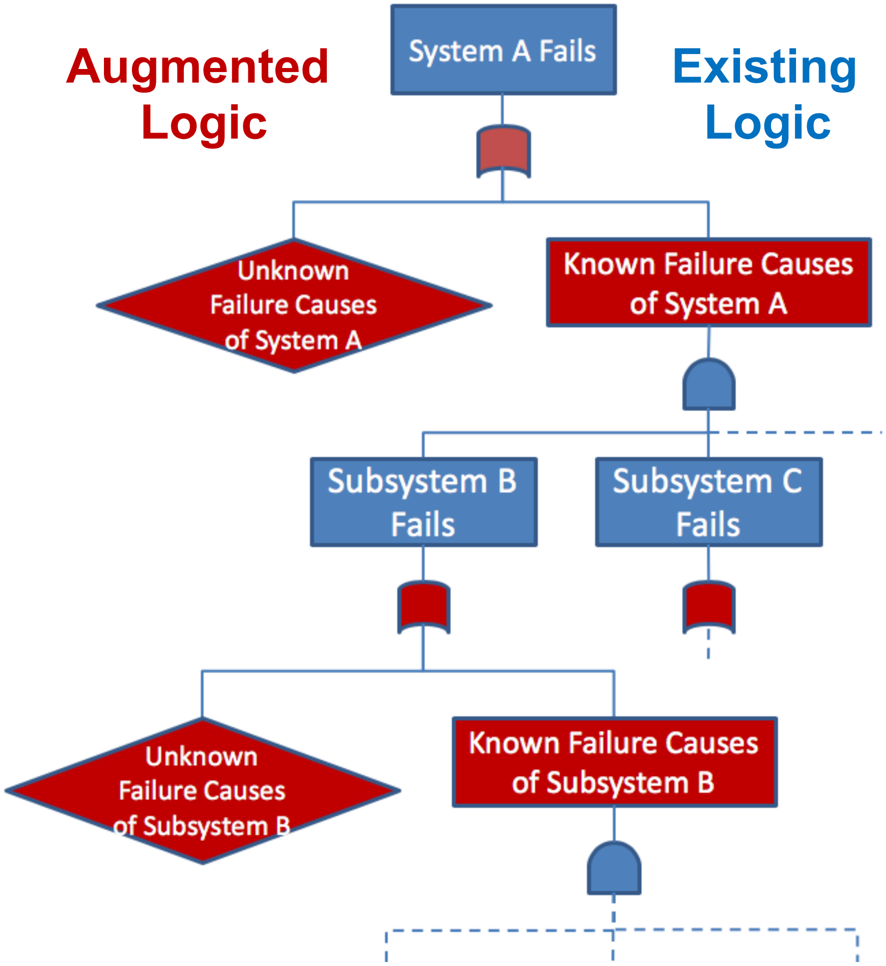
$$P_U|P_K = F_U^{-1}(F_K(P_K))$$



P_K, P_T, and P_U for Correlated P_K, P_U

The Vision – Universal Incorporation of Unknown Failure Causes into PRA

- **The issue of incompleteness is not limited to the analysis of top events and/or end states**
- **It applies to *every* causally decomposed event in a PRA model**



- **The result is a PRA that:**
  - Is inherently complete at every level of decomposition
  - Fully represents belief about the event probabilities in the model
  - Allocates unknown failure cause probabilities into the system
  - Provides vectors for information that is traditionally excluded from PRA
  - Supports analysis use cases that traditional PRA does not address

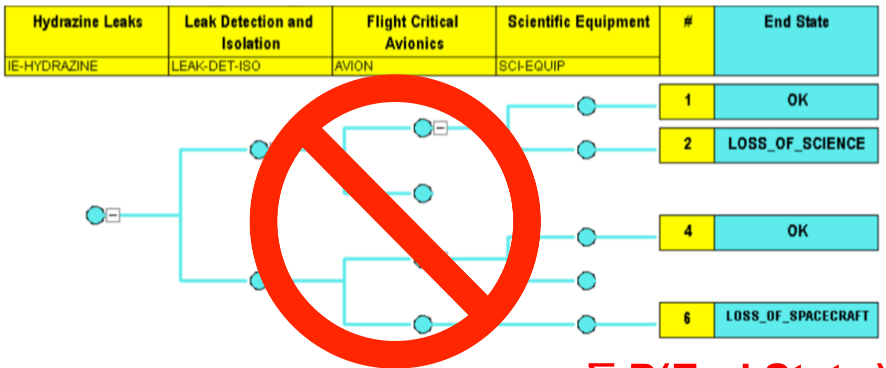
What's not to like?

Provides Vectors for Risk-Related Information

- **The inclusion of prior belief enables historical information, expert opinion, similarity analyses, etc., *at any and all levels of decomposition* to be incorporated into the analysis**
- **Belief about analysis completeness, $f_C(C_A)$, can be developed further developed as a function of indicators of completeness, e.g.,:**
  - **Analysis credibility**: NASA-STD-7009A, "Standard for Models & Simulations," includes an M&S Credibility Assessment, which bears on $f_C(C_A)$
  - **Technology readiness level (TRL)**: TRL is basically a proxy for completeness
    - Low TRL correlates to high probability of unknown failure causes
- **The inclusion of unknown failure causes enables testing and operational history to be incorporated into the analysis**
  - In particular, ***operational successes*** strongly affect $f_U(P_U)$ despite having a negligible effect on $f_K(P_K)$
  - Allows PRA to be used in a general **value-of-information** (VOI) capacity

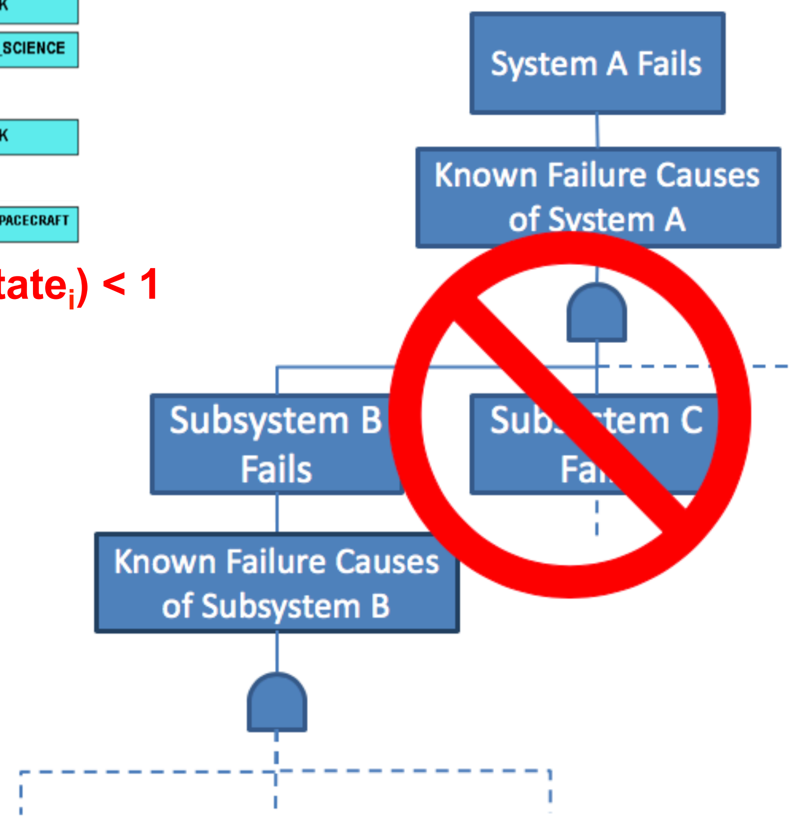
A Constraint on the Method

- It is doubtful that an incomplete event tree would be considered acceptable



$\sum P(\text{End State}_i) < 1$

$P(\text{System A Fails}) < P_T$



- In other words, event tree analysis (ETA) implicitly enforces ***completeness of event consequences***
- Similarly, fault tree analysis (FTA) could, or *should*, enforce ***completeness of event causes***
- Both are needed for ***completeness of the analysis***

Conclusion

- **This work *points a way* towards the explicit incorporation of unknown failure causes into PRA**
- **The benefits are manifold:**
  - It results in a *"complete" risk model* that captures the full scope of belief concerning system failure probability, at every level of logical decomposition
  - It results in an analysis that is appropriate for *risk acceptance decision-making* in a way that "synthetic-only" PRA is not
  - It *allocates "UU risk*" throughout the logic model, informing risk management decisions such as margin determination by indicating what parts of the system may be more likely than others to be harboring vulnerabilities
  - *It accommodates test and operational experience, particularly successes*
- **Traditional *PRA is recovered* under the assumptions that:**
  - The *priors are uniform* (justifiable as non-informative agnosticism)
  - The *causal decompositions are complete* (when is this implicit assumption justifiable?)
- **It answers the right question, *"What is the probability of failure?"***