Letters to the Editor

Software Safety vs Software Reliability

While looking back through Vol. 56, No. 1 (Summer 2020) of *Journal of System Safety*, I finally took the time to read Nathaniel Ozarin's article "Lessons Learned in a Complex Software Safety Program."

The article is quite interesting and thought provoking, comparing what actually occurs while implementing a system safety program to the idealized descriptions found in documents such as MIL-STD-882, JSSSEH and AOP-52.

While I found the article interesting and informative, I noted that the author consistently characterizes the "software safety problem" as a "reliability" problem, focused on finding and preventing "failures" and ensuring high "reliability." Clearly, the problem of software safety goes far beyond reliability or the prevention of failures — in fact, those two concepts as generally understood might not be associated with software at all. Software doesn't "fail," thus, software is 100% "reliable." Hardware fails. Software might be designed in such a way as to fail to take those failures into account. People "fail" by making mistakes. Designs might "fail" by not achieving the desired results. But software does what software does — accepts inputs, crunches logic and issues commands.

I noticed that Mr. Ozarin's background is as a reliability engineer and that he was named Reliability Engineer of the Year by the IEEE Reliability Society. While I quibble with Ozarin's characterization of software safety as being centered upon failures and reliability to function "correctly," I suggest that in the context of his article, the issue is more one of semantics than anything else. I would also like to suggest that "failure," as he described it in the article, might be a general concept such as "failure to function safely" or being able to reliably ensure safety.

Of course, both of these are actually system safety considerations, not "reliability" in the strict meaning of the terms. "Failure to function safely" boils down to a failure to identify and mitigate hazards and risks. Reliability doesn't seem to be applicable to software — unless perhaps it means that the design (or coding) of the software is different than intended. In any case, software only consists of logic; it can be expected to do what the code demands — unless the associated hardware doesn't support the software, has a failure or has a reliability problem.

While I enjoyed the article and think it contains much wisdom for those embarking on a software safety program, I also think clear descriptions of the terms "software failure" and "software reliability" are in order. These would help the reader understand that traditional reliability tools, analyses, measurements and probability estimation techniques don't apply to the field of software system safety.

— Charles Hoes
Hoes Engineering

We Want to Hear from You!

Journal of System Safety is seeking papers and articles on many topics where system safety makes a critical contribution, including:

- Explosive Safety
- Nuclear Safety
- Hazardous Material Management
- Chemical Safety
- Biotech Safety
- Safety Management Issues
- Human Error
- Software Safety
- Safety-Critical Processes
- Lessons Learned
- Industry Book Reviews

Please send summaries or abstracts or letters to the editor to Chuck Muniak, Technical Editor, at journal@system-safety.org.

Some Thoughts on the Probabilistic Criteria for Ensuring Safe Airplane-System Designs

We (Ted Yellman and Thomas Murray) have been employed in the risk sciences for a total of 86 years, including 62 years in reliability engineering and safety engineering positions at The Boeing Company. For many of those years, Yellman was the designated "Risk-Analysis Focal" (person) for Boeing's 707, 727, 737 and 757 airplane models.

For several decades, the United States government has published the same criteria, created by the U.S. Federal Aviation Administration (FAA), intended to ensure that the systems on large (transport-category) aircraft have been designed to be safe [Refs. 1 and 2]. But we believe that the criteria have failed to prevent certain aircraft accidents, and we think that the reasons for that should be better understood. We hope that this discussion will contribute to a better understanding by examining the part potentially played in those accidents by the FAA's criteria that are defined probabilistically.

The FAA's Probabilistic Airplane-System Design-Safety Criteria

To understand the FAA criteria, one must first understand what "catastrophic accident," "failure," and "failure condition" mean.

Catastrophic Accidents

The FAA's criteria do not define "catastrophic accident." In practice, this term can be interpreted to mean an accident that causes at least five deaths and/or a financial loss of at least a million dollars. Precision is rarely required to define a "catastrophic accident" because the great majority of accidents considered "catastrophic" are typically "crashes," which are almost always "catastrophic" by any standard.

Failures

The FAA defines "failure" as "a loss of function, or a malfunction, of a system or a part thereof" [Ref. 2]. Such an event is more precisely called a "system functional failure." In the case of the FAA's "failure," it is more specifically a "system functional failure condition," meaning a state of a system that results in it having lost its ability to function in some way that it should, or in its having acquired an ability to function in some way that it shouldn't.

Failure Conditions

The FAA's probabilistic airplane system design safety criteria are based on a concept that it calls a "failure condition," which it defines as "the effects on the airplane and its occupants, both direct and consequential, caused or contributed to by one or more failures, considering relevant adverse operational or environmental conditions" [Ref. 2]. So the FAA's "failure condition" is similar to a system functional failure elevated to the airplane level, but it also sometimes includes its implications for airplane behavior in certain operational (e.g., flight phase) and/or environmental (e.g., weather) conditions.

Unfortunately, the FAA's definition of "failure condition" does not tell the whole story. It applies only to "airplane systems," which by FAA definition do not include propulsion systems, structural systems, or programmed content (software). And the FAA's "failure condition" definition seems to imply that the occurrence of any particular failure condition will always have the same "effects" if its "relevant operational or [and] environmental conditions" are "considered." However, at the moment that a particular functional failure arises, concurrent operational and environmental conditions can vary greatly. So neither such operational or environmental conditions can be properly accounted for by just "considering" which one(s) of them apply. Often any of many combinations of such conditions can apply, so properly accounting for them requires using probability theory. Finally, although the FAA considers the *demands* put on a flight crew by an airplane functional failure to be "consequences" of failure conditions, it does not consider the *responses* of flight crews to such demands to be included as "consequences." That is a critical point, as we will show.

We suggest this definition for the FAA's failure condition: "An airplane functional failure caused by one or more component physical failures not in a propulsion system or a structural system, sometimes combined with particular operational and/or environmental environments, causing an airplane to either (1) lose its ability to perform in some particular way that it should, or (2) acquire an ability to function in some particular way that it shouldn't." The FAA requires that airplane manufacturers identify a set of such failure conditions for each new airplane type and divide them into three categories: Catastrophic, Major and Minor. Each has its own definition and probabilistic requirements, which we will now discuss, but not in that order.

Minor Failure Conditions — The FAA defines "minor failure conditions" as "failure conditions which would not significantly reduce airplane safety, and which involve crew actions that are well within their capabilities" [Ref. 2]. It also explains that "minor failure

conditions may include, for example, a slight reduction in safety margins or functional capabilities, a slight increase in crew workload such as routine flight plan changes, or some inconvenience to occupants."

A more practical interpretation of "minor failure condition" is "a failure condition that an airplane manufacturer considers to be not sufficiently hazardous to require a probabilistic analysis."

Catastrophic Failure Conditions — The FAA defines "catastrophic failure conditions" as "failure conditions which would prevent continued safe flight and landing" [Ref. 2]. It defines "continued safe flight and landing" as "the capability for continued controlled flight and landing at a suitable airport, possibly using emergency procedures, but without requiring exceptional pilot skill or strength" [Ref. 2]. And "prevent continued safe flight and landing" is a euphemism for "cause a catastrophic accident."

A more practical definition of the FAA's "catastrophic failure condition" is "a failure condition that an airplane manufacturer assumes will always cause a catastrophic accident." The FAA requires that the probability of each catastrophic failure condition modeled as time-dependent (for example, a failure of some airplane critical continuously-operating flight-control system) be no greater than $1/10^9$, based on an experiment consisting of one random airplane flight hour. And it requires the probability of each catastrophic failure condition modeled as time-independent (for example, the loss of the ability to extend an airplane's landing gear) also to be limited to$1/10^9$, but based on an experiment consisting of one random time-independent system operation.

Major Failure Conditions — A reorganized definition of "major failure conditions" as defined in Ref. 2 is: "Failure conditions which would reduce the capability of the airplane or the crew to cope with adverse operating conditions to the extent there would be, for example, (1) a significant or a large reduction in safety margins or functional capabilities, (2) a significant increase in or higher crew workload, (3) physical distress such that the crew could not be relied on to perform its tasks accurately or completely, (4) a significant increase in conditions impairing crew efficiency, or (5) some discomfort to or adverse effects on occupants." But that definition fails to recognize that *those effects are not the effects of most concern from occurrences of major failure conditions! The effects of most concern are (or should be) catastrophic accidents!*

The remainder of this article is about how major failure conditions causing catastrophic accidents can be better controlled, because their lack of control in the FAA's safety criteria is a serious flaw.

The FAA requires the probability of each major failure condition (based on an experiment consisting of one random airplane flight hour) to be no greater than $1/10^5$. But as we will show, depending on how flight crews respond to those conditions, that limitation can be far too weak. First, however, we will try to put the three types of failure conditions and their frequency limits into better perspective by using a simple probabilistic model.

A General Probabilistic Model for Airplane Failure Conditions Causing Catastrophic Accidents

Figure 1 illustrates a probabilistic model that is based on the Event-Sequence Analysis concept [Refs. 3,4,5,6 & 7]. Unlike Fault-Tree Analysis, Event-Sequence Analysis consistently relates undesired events to each other in the various time orders in which each of them might arise, meaning in the time orders that the Event Assertions that define those events can become true. (An event assertion and alternatively an event statement is a natural-language declarative sentence, such as "The dog bites the mailman" [Ref. 8].) The time-based foundation of Event-Sequence Analysis makes it much more effective and realistic for estimating event probabilities that vary depending on what other events arose or didn't arise previously.

The value of the probability of *any* individual failure condition causing a catastrophic accident can be calculated very simply as the algebraic product of these two probabilities:

1. The probability that the failure condition arises, and
2. Given that the failure condition arises, the probability that the airplane flight crew fails to cope with that failure condition and that results in a catastrophic accident.

Failure conditions identified as "catastrophic" are treated as being so hazardous that when they occur, flight crews will never be able to save their airplanes from catastrophic accidents — and that worst-case assumption makes 100% the value of the second probability in Figure 1. But for a failure condition identified as "major," that second probability is rarely so straightforward because estimating a value for a human failure probability is often controversial and is seldom supported by good statistical data. In fact, the FAA criteria state (twice!) that "...quantitative assessments of the probabilities of crew errors are not considered feasible" [Ref. 2].

Although the FAA allows each major failure condition to be predicted to occur as frequently as once per 100,000 flight hours, probabilities for flight crews then

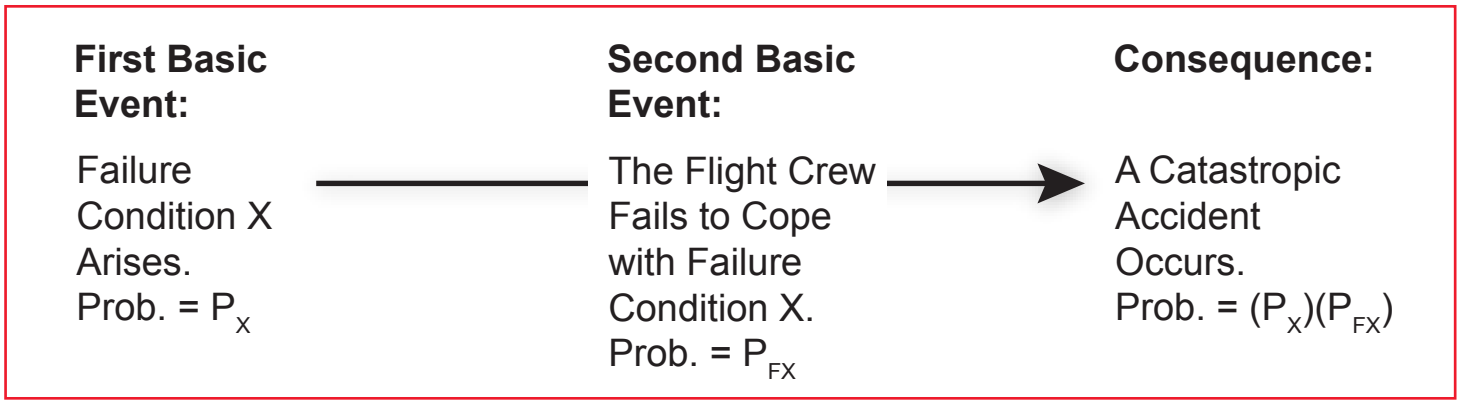

Figure 1 — An Event-Sequence Diagram that can be applied to any *Airplane Failure Condition causing a Catastrophic Accident. (The "first event" and "second event" refer to the time orders in which the basic events originate.) The probability of each basic event is conditional upon the basic events that have previously arisen.*

failing to cope with those major failure conditions are not even recognized. And even the *possibility* that major failure conditions can cause catastrophic accidents is not recognized! But they can, and if the value of a flight-crew-failure probability is larger than $1/10^4$ (that is, one in 10,000), the major failure condition that triggered that flight-crew failure is allowed to cause catastrophic accidents more frequently than catastrophic failure conditions are allowed to cause catastrophic accidents! But a flight-crew failure probability value of less than $1/10^4$ is often extremely optimistic. If it is 1/10, for example, that allows its preceding major failure condition to cause catastrophic accidents with a probability of $(1/10^5)(1/10) = 1/10^6$. And that is 1000 times as large as the $1/10^9$ probability value (based on an experiment consisting of one random airplane flight hour) that the FAA allows for each catastrophic failure condition!

Furthermore — and more subtly but at least as significantly — the FAA's lack of recognition of the possible catastrophic losses caused by major failure conditions (unintentionally of course) *almost invites airplane manufacturers to under-design some systems for safety (or at least it gives those manufacturers an excuse for doing that)*, by not requiring any systematic accounting for flight crews sometimes failing to cope with major failure conditions. It leaves that entirely up to airplane manufacturers should they choose to do so. And it can be tempting to those manufacturers to go no deeper than just showing that a major failure condition meets the allowed $1/10^5$ probability value, because doing only that satisfies the FAA's probabilistic criteria and, thus, makes it very likely that the FAA will find the risk from that major failure condition to be acceptable.

So Where Should We Go from Here?

We suggest that the FAA also create a limit on the probability values predicted for individual major failure conditions causing catastrophic accidents, just as it has done for catastrophic failure conditions. An obvious choice for such a limit is $1/10^9$, the same as the limit for an individual catastrophic failure condition causing a catastrophic accident. But if we can convince the aviation community that the current criteria for major failure condition safety risks need improvement, we prefer that limit to be chosen then, and chosen by the aviation community as a whole.

Conclusions

The FAA's probabilistic criteria are intended to serve as checks to ensure that airplane systems are designed to be safe, not to be a substitute for airplane manufacturers and their engineers taking it upon themselves to design airplane systems to be safe. Over the years, however, getting new airplane types through the FAA's process so they will be approved for commercial service has become increasingly dependent on showing that airplane systems meet the FAA's probabilistic criteria. And given how firmly that process has been established, that dependency is not likely to change. Therefore, to ensure that commercial airplane flight is safe, those criteria must be both effective and comprehensive. But for major failure conditions, they are not, because the FAA was apparently unwilling to even recognize that major failure conditions can cause catastrophic accidents! The current probabilistic criteria for controlling losses caused by major failure conditions completely neglect possible ineffective flight-crew responses to major failure conditions that can then cause catastrophic accidents. If those criteria are not improved, airplane manufacturers may again take advantage of the weakness of those criteria and under-design their airplane systems for safety.

Estimating probabilities of flight crew failures, of course, involves more uncertainty than, for example, estimating the probabilities of various poker hands. But that does not justify completely neglecting the likeli-

hoods of flight crews failing to save their airplanes from catastrophic accidents following major failure condition occurrences!

Good probabilistic analysis requires that all the events that can be causes of catastrophic airplane accidents be addressed. So we recommend that the FAA require airplane manufacturers to also account probabilistically for the flight-crew failures that can follow major failure condition occurrences. Such a requirement would force those manufacturers to formally and critically evaluate (and if required by the FAA, to defend) the adequacy of the flight-crew procedures and training that manufacturers provide to airlines to help pilots successfully deal with major failure conditions.

— Ted W. Yellman
Bellevue, Washington

— Thomas M. Murray
St. George, Utah

References

1. U.S. Code of Federal Regulations, Title 14, Section 25.1309. "Equipment, systems, and installations," November 8, 2007.
2. U.S. Federal Aviation Administration Advisory Circular 25.1309-1A. "System Design and Analysis," June 21, 1988.
3. Yellman, Ted W. "Event-Sequence Analysis," *Proceedings of the 1975 Annual Reliability and Maintainability Symposium*, Washington, D.C., January 28-30, 1975.
4. Yellman, Ted W. "Event-Sequence Analysis vs. the Fault Tree," *Proceedings of the 1981 Annual Reliability and Maintainability Symposium*, Philadelphia, Pennsylvania, January 27-29, 1981.
5. Yellman, Ted W. "The Event-Sequence Analysis Concept," *Tenth International System Safety Conference*, Dallas, Texas, July 1991.
6. Yellman, Ted W. "Applying the Event-Sequence Paradigm to Learn from an Accident," *Probabilistic Safety Assessment '96 Conference*, Park City, Utah, September 1996.
7. Yellman, Ted W. and Murray, Thomas M. "Mishap Event-Sequence Analysis," *Safety Across High-Consequence Industries Conference*, Saint Louis, Missouri, September 20-22, 2005.
8. Yellman, Ted W. "The Event: An Underexamined Risk Concept," *Risk Analysis*, Vol. 36, No. 6, 2016.