UNIVERSITY of York

System Safety Principles from 1999: Challenges for 2019?

Dr David Pumfrey

About me…

- 30 years working on safety critical systems
- 27 years working for the University of York
- Submitted my DPhil thesis 20 years (and two weeks) ago
  - "The Principled Design of Computer System Safety Analyses"
- Started teaching on the university's MSc in System Safety Engineering in 1998
  - A very unusual lectureship… I spend 90% of my time interacting with industry
    - I have personally worked with ~40 companies in 8 countries
    - Our group total is much higher
  - Gives a privileged overview of how different industries are meeting challenges of safety critical systems

… and why I'm reflecting on 20 years' progress

- 20 years ago

"This safety stuff is all very good – but how do we persuade our managers to spend on it?"

- Now

We're doing it, but it's costing too much

The safety effort isn't adding anything to the programme

We are

Nobody seems to be able to tell us which

ile

There are so many safety standards, they are all

c

There are all these different safety gurus. There's Sid Dekker's human factors approach, Nancy Leveson with STAMP, you lot from York with Safety Cases… you're all just in it to make money from us!

- What have we learned? What are *my* views on the "state of the art?

My 1999 thesis: 9 principles…

1: Safety analysis must have value as part of the engineering process

2: Method is more important than notation

3: Techniques should be as simple as possible

4: Techniques should guide without unnecessarily constraining

5: The role of the technique should be clear

6: Safety analysis starts at the system level

7: Projective analyses are key to software safety

8: Safety analyses must consider hardware and software

9: Techniques should use familiar concepts and models

and two analysis techniques: SHARD and LISA

An honest look back: How well did I do?

SHARD:

- A variant of HAZOP aimed at software
- Based on lots of structured case studies
- I score it about 6/9 for following my principles (pretty good)
- Has been widely applied, adapted, extended…

LISA

- Low-level analysis of hardware / software interactions based on the concept of *resources*
- Invented in a hurry to solve a specific problem!
- I score it about 2/9 for following my principles (terrible)
- It has hardly ever been attempted since my case studies

The view from 2019?

Good news:

- I (mostly) still believe what I wrote 20 years ago 🙂

Not so good news:

- There is a lot of “muddled thinking” 😐
- There is a lot that is over-simplified
- The scope was too narrow

Really bad news:

- I didn’t even stick to my own principles in the rest of the thesis 😳

Great news

- With a bit of organisation, those 9 principles point at what I believe to be key challenges for the safety critical industry today 😃

What sort of challenges do I see?

- We (safety community) have researched and written a lot about how others work
  - In "front line" safety critical jobs such as pilots, railway signallers…
- I'm really interested in how we work in safety
  - As "safety thinkers", whether analysts, consultants, engineers or managers
    - People whose role is explicitly to consider safety, then make critical decisions (or advise others…)
- There is "creative tension" between many of the ideas, e.g.
  - the principles that talk about safety *thinking* suggest that much may be gained by guiding / helping analysts to think in new ways
  - the principle that introduces safety *communication* emphasises the use of familiar notations and models

A bit of organisation: 4 “challenge areas”:

Challenge area 1: *Thinking* about safety

2: Method is more important than notation

3: Techniques should be as simple as possible

4: Techniques should guide without unnecessarily constraining

Challenge area 2: Safety as an *emergent system property*

6: Safety analysis starts at the system level

7: Projective analyses are key to software safety

8: Safety analyses must consider hardware and software

Challenge area 3: *Communicating* about safety

9: Techniques should use familiar concepts and models

Challenge area 4: Making safety work *effective*

1: Safety analysis must have value as part of the engineering process

5: The role of the technique should be clear

Challenge area 1: *Thinking* about safety

Note the implicit assumption of human safety activities

2: Method is more important than notation

- Help analysts to think creatively. Notations can be information dense, but are only as good as their effectiveness in communication
- Is “method” just procedure for applying analysis – or much broader (e.g. where to find “encoded expertise”)

3: Techniques should be as simple as possible

- Time / effort to learn and apply
- What is our “go-to” method?

4: Techniques should guide without unnecessarily constraining

- Structuring can help with “search space”, completeness, BUT risk of impeding creativity, slowing work, or “tick box thinking”
- Concept of confidence in safety cases

Challenge area 1: Fundamentals

- There are hundreds of different safety analysis methods
- Most are subtle variants on a limited number of themes
  - Asking questions
    - “what if...” (inductive); “how could...” (deductive), or “how likely...”, “how significant...” (quantitative)
    - at different level of detail / technologies / stages of lifecycle
- All based on fundamental underlying concepts:
  - System models
    - abstractions of real world but retaining sufficient detail to produce useful / meaningful results
  - Causal representations
    - “why things happen”
  - Systematic approach to “search space”
    - rules or guidance on how to achieve (acceptable) completeness

Challenge Area 1: Where are we now?

- A (heretical) question:
  - Do we need more analysis methods? Where? Why?

Observation: My own thinking was way too narrow back in 1999. Safety thinking is far more than just using good analysis methods.

- How do we develop (and teach) the “safety way of thinking”?
- How do we structure and codify that for automation?
- How do we update safety thinking to reflect new technologies (AI, IOT?)
- How do we ensure we reflect and respond to societal changes in perspective on safety?

Challenge area 2:
Safety as an *emergent system property*

Principles from 1999:

6: Safety analysis starts at the system level

- Get the requirements right
- What do we think of as the system?

7: Projective analyses are key to software safety

- Move away from the "bugs in code" model of software failure and think about desirable and undesirable behaviour in context
- The challenge of completeness – have we thought about sufficient possibilities?

8: Safety analyses must consider hardware and software

- A challenge to the formalists! Algorithmically correct software can fail when run on real hardware.
- Much too narrow! Add humans, environment, organisations…

Challenge area 2: Where are we now?

- The change in scope (we must always consider safety of a system in context) is the single biggest change since I started work
- This subject has more active research and publication than any of my other challenge areas
  - What is valid / important to consider is constantly extending
    - Human factors
    - Organisation and culture
    - Socio-technical safety
    - The influence of national culture, religion etc.
- Two key challenges:
  - Get some of the more “traditional” industries to accept that this is necessary!
  - Cope with the complexity

Challenge area 3: *Communicating* about safety

- Principle from 1999:

  9:Techniques should use familiar concepts and models

  - If safety analysis is to be embedded in project processes (e.g. driving design), other engineers must be able to understand the language

  Too narrow in two ways

  - Many safety activities should be part of other lifecycle activities (i.e. owned by design or operational staff) anyway … outside scope of this discussion
  - We need to communicate about safety with a very broad spectrum of people – technical, management, customers, regulators, public – and that requires far more than just using the right models

Some quotes to reflect on

“I am not qualified to manage these safety issues. I need you to tell me what to do.”

Director, medium-sized (~100 employee) company [in a board meeting], 2019

“The engineers at project level know what’s needed for safety. We ask for the right products and the right evidence, but then it goes to Commercial, and it all gets ‘value engineered’ out. We end up with contracts that don’t specify the safety properties we need, or don’t mandate the evidence. Then we end up with no confidence that the product is really safe, or we can’t make the safety case, so we end up having to go back and do lots of re-work. Or we just kind of bodge something together and hope.”

Project engineer, large international company, 2019

Communicating Safety: Some reasons it’s hard

- Unpopular messages
  - “The way you have always done things is not actually very safe”
  - “This design has some safety problems”
- Strange ways of thinking
  - “Safety people” often think and talk negatively (in terms of failures and problems); most engineering and project management is the language of success (how it will work, how we will do this….)
  - “Safety people” think backwards (from undesired outcomes back to causes);
- We have turned safety into a specialised discipline, with its own language
  - Some of it needs to be (specialist analyses etc)
  - The results and responses *should not be*

Challenge Area 3: Where are we now?

- In many parts of mass culture, safety now has negative associations
  - There is a culture of distrust in “experts”
  - “We can’t do that any more because of Health and Safety”
  - This carries over into more technical aspects of safety work
- Like justice – safety must be “seen to be done”
- We have got to develop positive safety communication strategies
  - Not just stating facts, but how they are presented
    - c.f. some of the excellent work being done to help combat misinformation about vaccination risks
  - Work on real understanding
- Still far more to do on “safety is everyone’s job”

Consider this:

150,000 work hours last year

0 Fatal Accidents

0 Serious Injury Accidents

0 Serious Road Traffic Collisions

Thank You for your patience

Communicating About Safety: Coda

Communicating between ourselves: “Celebrating Success”

- Safety activities are “open loop
  - No confirmation of correctness
    - We only find out we were wrong when something bad happens
  - Predictive safety tasks (analysis etc) are inherently subjective
    - Often no obvious “right answer”
    - Primary check on soundness of work is peer review
- We are good at giving negative feedback
- We are less good at positive feedback
  - “That’s a great design solution”, “Very thorough analysis”…

Challenge area 4: Making safety work *effective*

1: Safety analysis must have value as part of the engineering process

- Too many safety activities were being done to get a “tick in the box”
  - Or even just because someone had suggested them and no-one evaluated their worth
- Decision-guiding activities should have obvious value
- The value of evidence-providing activities should be clear IF a safety case strategy (argument) is developed in advance.

5: The role of the technique should be clear

- Some safety analyses didn’t appear to make sense (e.g. an analysis method intended to guide design decisions that needed a complete design to start with…)
- Many methods can be used in different ways. Any given application should have a clear purpose.
  - “Conclusions and recommendations is the most important section…”

Pause for thought: More Quotes

“What you are describing isn’t our process at all. We do [*analysis activity*] because we are working to [*safety standard*] and we need that tick in the box”

Engineering team manager, very large international company, 2019

“Lots of highly technical safety analysis is being done on this project, but it is all being done by suppliers and contractors. We don’t own any of it, and we don’t have the skills to do it in house. It isn’t seen as important to develop those skills.”

Project manager, large international company, 2019

“Lots of safety analysis was done to support [*system*] entering service. There is supposedly a complete safety case. We are trying to use it for in-service support and diagnostics, but it’s useless. A lot of the analysis doesn’t seem to match the system, and it just doesn’t make sense.”

Service delivery engineer, international aerospace company, 2019

Challenge area 4: Where are we now?

- Can consider effectiveness of safety work at 3 levels:
  - How do we (“safety thinkers”) understand our role and activities?
    - What really matters?
    - Who are our “customers”… now, and in the future?
  - How are safety tasks built into projects and companies?
    - Project management
    - Stated priorities?
    - Safety culture
    - What really matters?
  - How are these tasks organised and structured in documents and other instruments that mandate or guide work in projects and companies?
    - Safety standards etc.
    - Safety critical industries DO safety. Why does it often feel as if they are fighting the safety process?

Challenge area 4: A personal perspective

- Engineering (in general) is a “divide and conquer” activity
  - Take a big, complex problem, chop it up into smaller parts, until each part can be handled by a single person or small team
  - Specialists deal with their own discipline or technology
- Safety engineering has to be an *integrating* discipline
  - I worry that too many companies and projects are attempting “divide and conquer” safety engineering
  - Yes, we need <u>every</u> engineer and manager to think about safety, but NOT lots of “safety people” doing super-specialised tasks *without context*
    - Several companies seem to have relied on “safety gurus” – a lifetime of diverse experience – but do modern career paths create that same depth of expertise?

Some closing thoughts

- The world of system safety has moved on significantly since 1999
  - Mostly for the better!
- As “safety thinkers”, we need to pay more attention to the way we work
  - That doesn’t mean we all need to think about “all of everything” – safety is a huge discipline – but, whilst we work at our niche area we need to try to retain an awareness of the “big picture”
  - How do we think about safety?
  - How do we ensure the effectiveness of the work we do?
  - How do we communicate effectively about safety?