Safety Synthesis

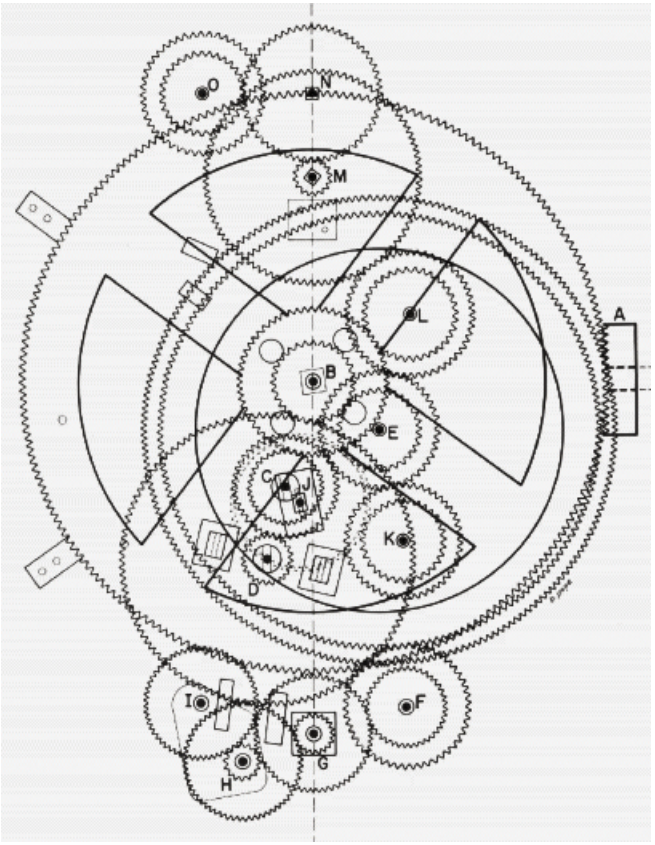
Understanding How Something Happens – When It Works And When It Doesn't

Erik Hollnagel

Professor, University of Southern Denmark
Chief Consultant Center for Quality, RSD (DK)

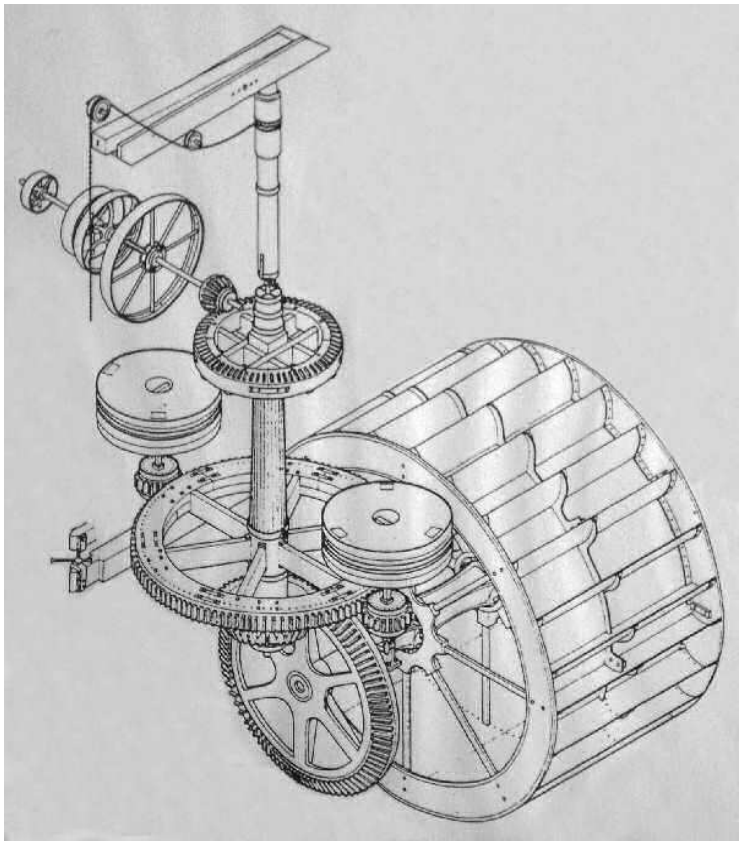
hollnagel.erik@gmail.com

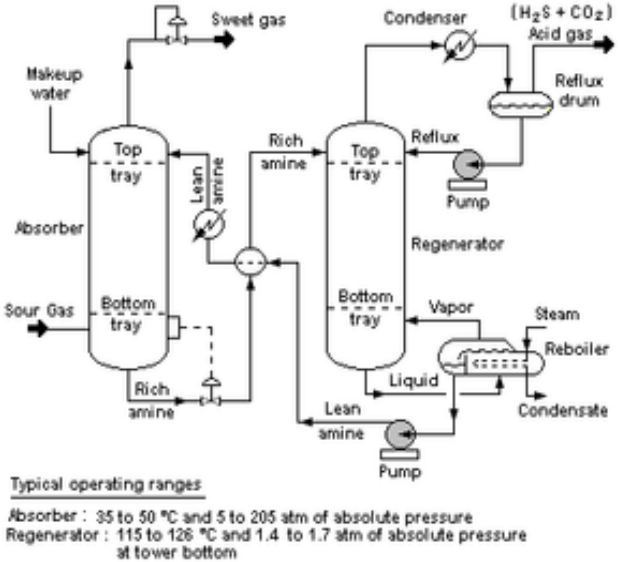
Understanding simple systems



Antikythera mechanism, (150-100 BC)

We can explain how things work in terms of cause-effect relations





We can therefore understand risks in the same way: as cause-effect chains starting from a component failure.

Principle of causation

Every cause has an effect

If we know what this is ...

Cause

Effect

Then we can look for this!

Observable (physical)

Not observable (metaphysical)

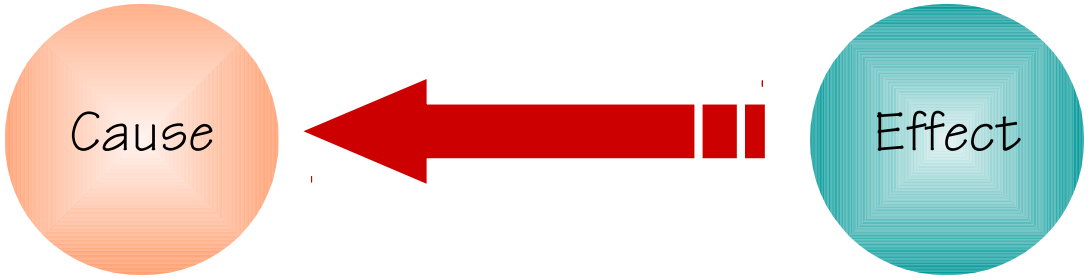
Observable physical)

A cause is prior in time to effect.
Cause and effect are contiguous in space and time.

David Hume 1711-1776

Reverse causation

Then we can find out what this is!



If we can see what this is ...

Every event (effect) has a prior cause

Humans are prone to reason in ways that are not logically valid.
(Affirming the consequent.)

If A, then B

B, therefore A

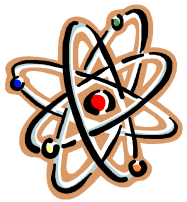
Sequentiality in a description is partly an artefact of time being one-dimensional.

Strengths of analytical thinking

Medicine
Diagnosis
Virus, germs

Genetics
Chromosomes
Genes

Physics & chemistry
Elementary particles
Theory of everything



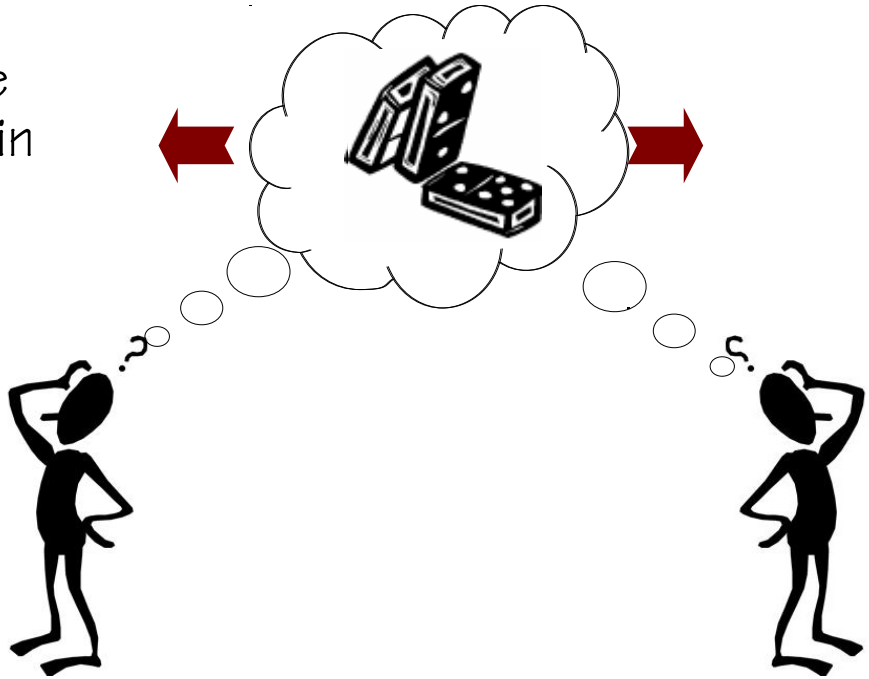
Criminal investigation (real / fictional)
Motives
Modus operandi

Geology
Earthquakes
Volcanoes

Simple, linear model (cause-effect chain)

Simple linear models
(cause-effect chains)

If accidents are the culmination of a chain of events ...

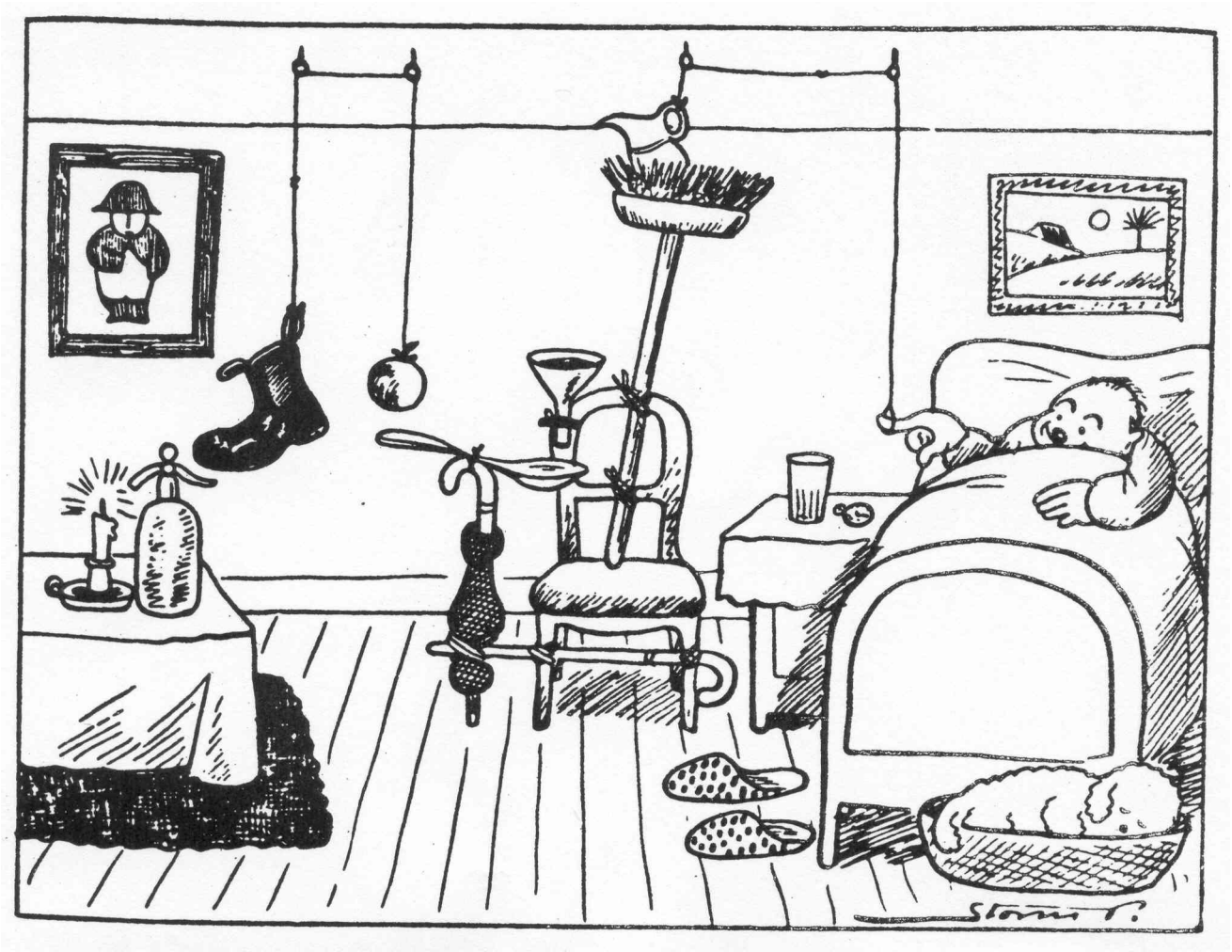


... then risks can be found as the probability of component failures

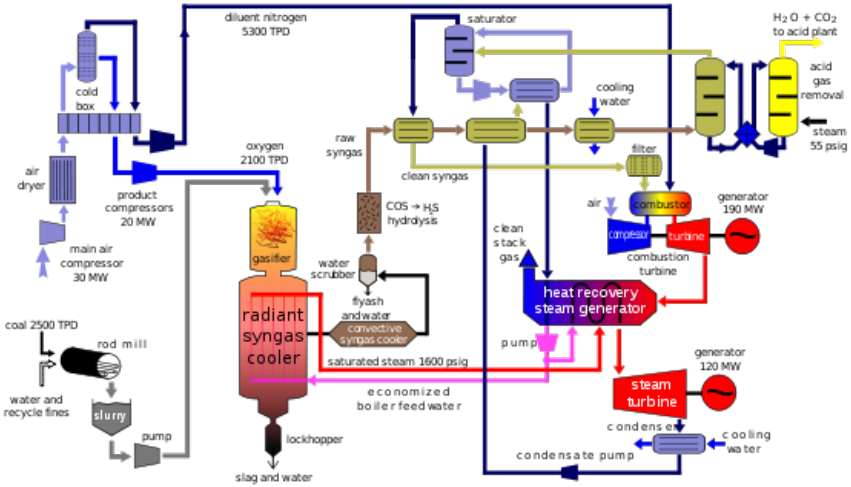
Find the component that failed by reasoning backwards from the final consequence.

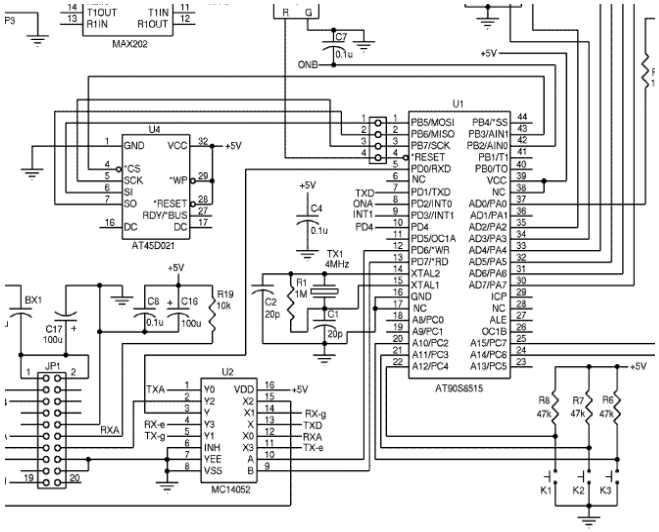
Find the probability that something "breaks", either alone or by simple, logical and fixed combinations.

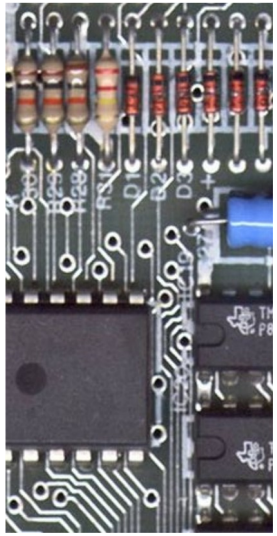
Cause-effect reasoning



Understanding not-so-simple systems



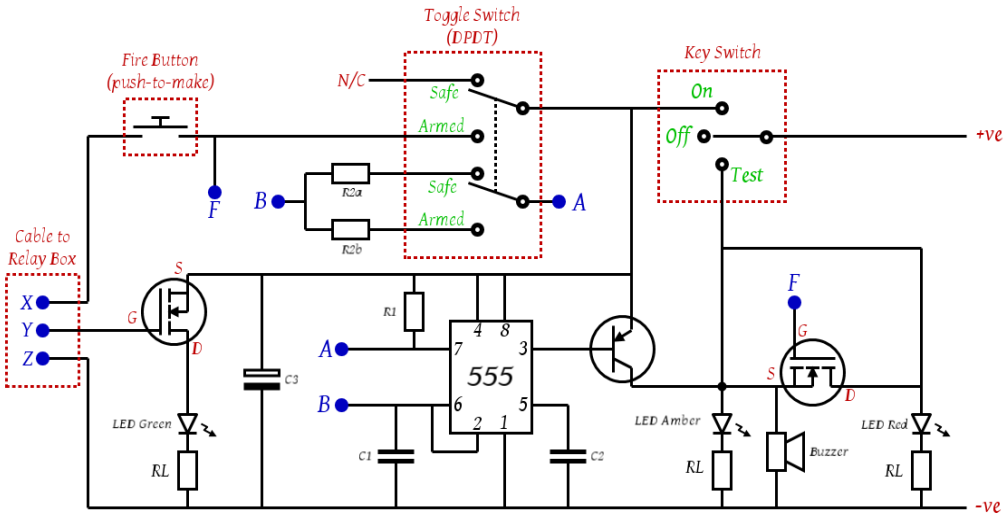




Reasoning in cause-effect relations is no longer adequate.

Difficult to imagine how events and conditions may combined.

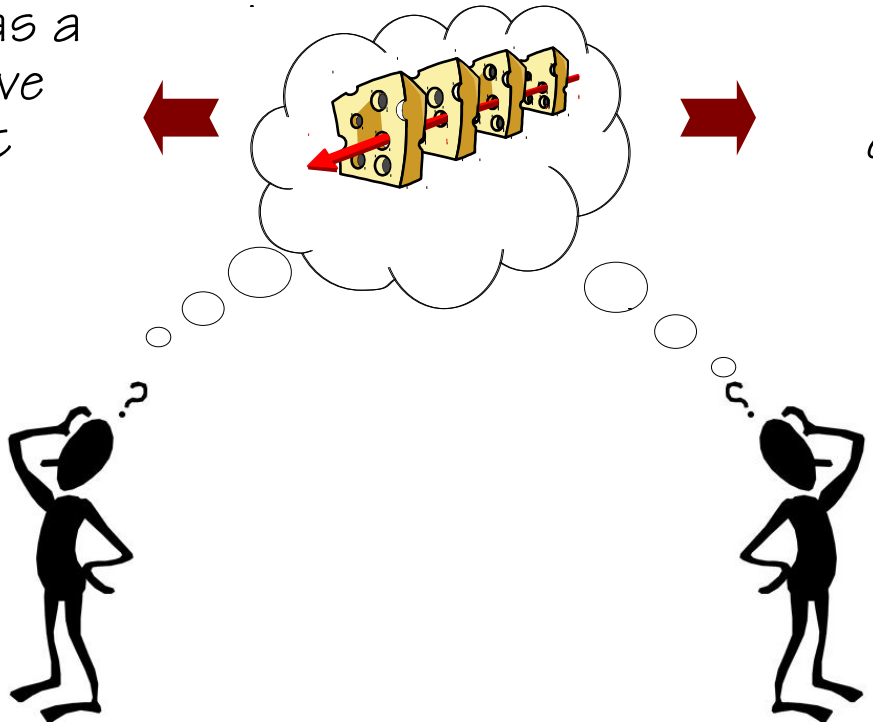
A growing number of risks therefore remain unknown.



Combinatorial (complex) linear model

Complex linear models

If accidents happen as a combination of active failures and latent conditions ...

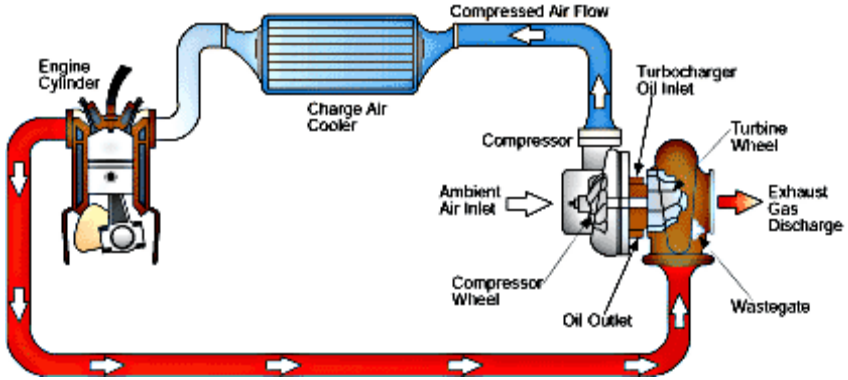


... then risks are the likelihood of weakened defences in combination with active failures

Look for how degraded barriers or defences combined with an active (human) failure.

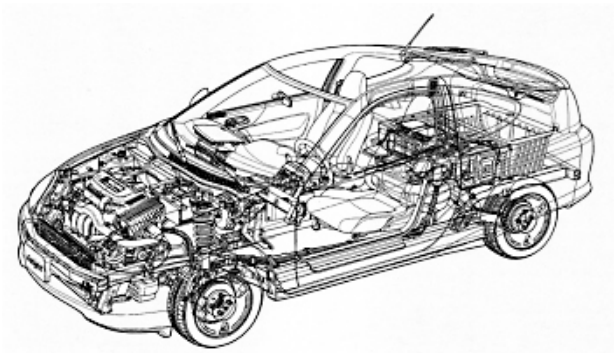
Combinations of single failures and latent conditions, leading to degradation of barriers and defences.

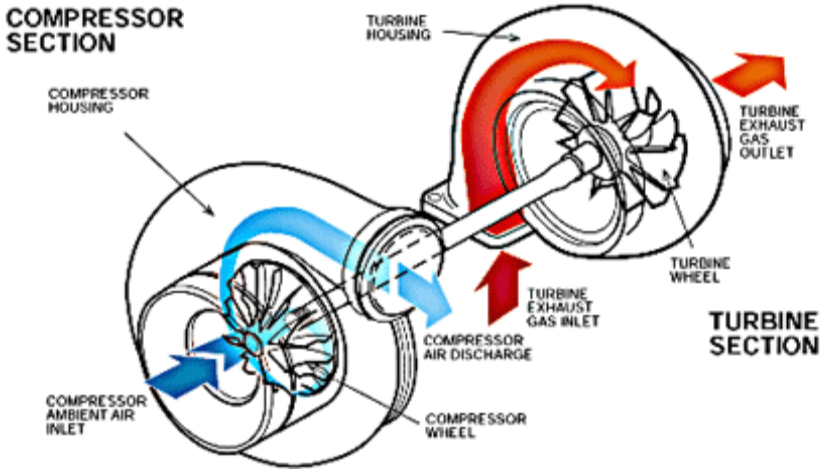
Explaining technical systems



When looking at a piece of equipment, such as a machine, it is natural to explain how it works by decomposing it into smaller parts.

We know how components work and how they are put together, because we have constructed the devices.





Explaining socio-technical systems

All systems unique

Must be described top-down in terms of functions and objectives.

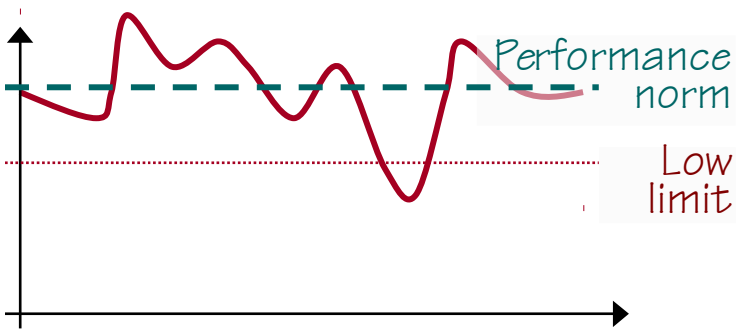
Decomposition does not work for socio-technical systems, because they are emergent.

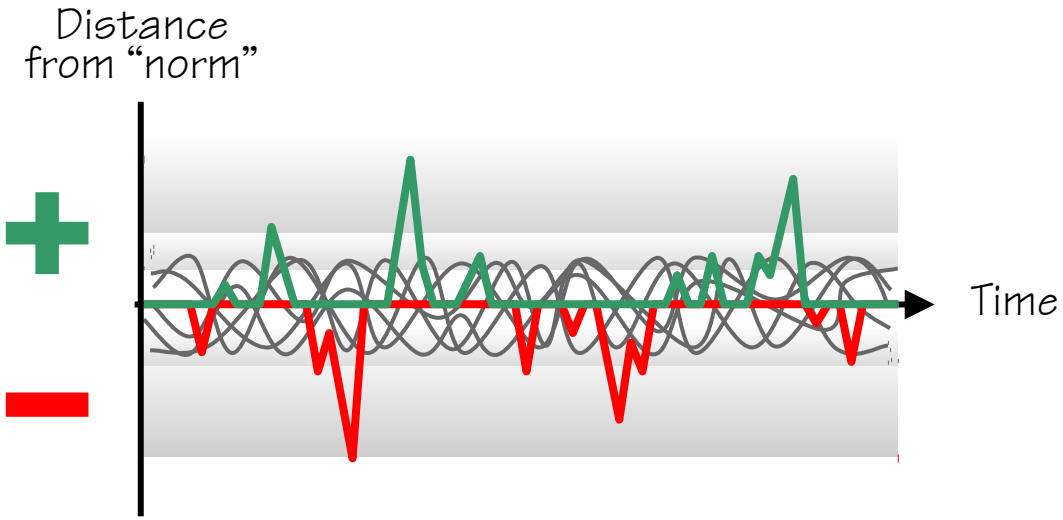
Risks and failures must therefore be described relative to functional wholes.

Complex relations between input (causes) and output (effects) give rise to unexpected and disproportionate consequences. Socio-technical systems are non-linear and event outcomes are intractable.

Socio-technical systems are not bimodal

Humans and social systems are not bimodal. Everyday performance is variable and this – rather than failures and 'errors' – is why accidents happen. Since performance shortfalls are not a simple (additive or proportional) result of the variability, more powerful, non-linear models are needed.





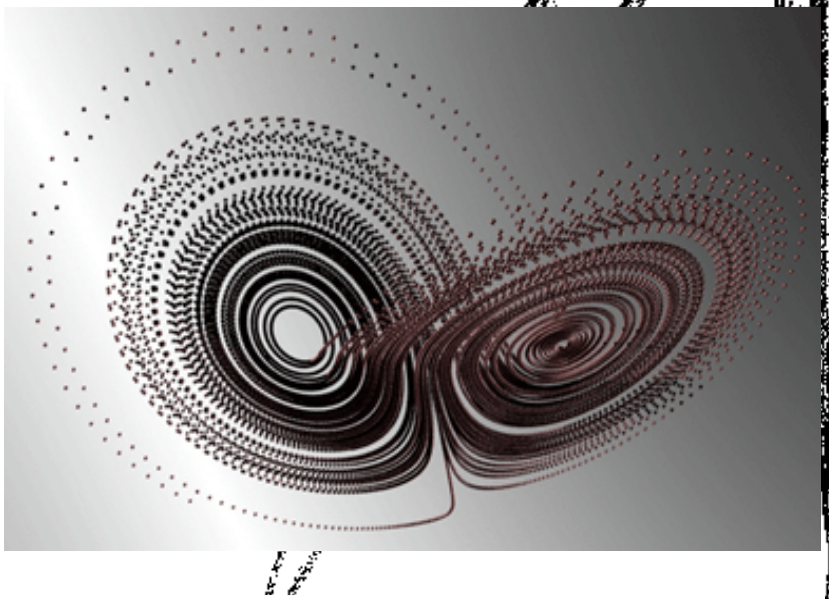
Performance variations can be have positive as well as negative outcomes!

But human factors has tended to look for negative aspects of performance - deviations or "errors"

Understanding complex systems

Systems have become too complex to understand in detail (chaotic, emergent).

Systems change so fast that complete descriptions – of the real system – are impossible.

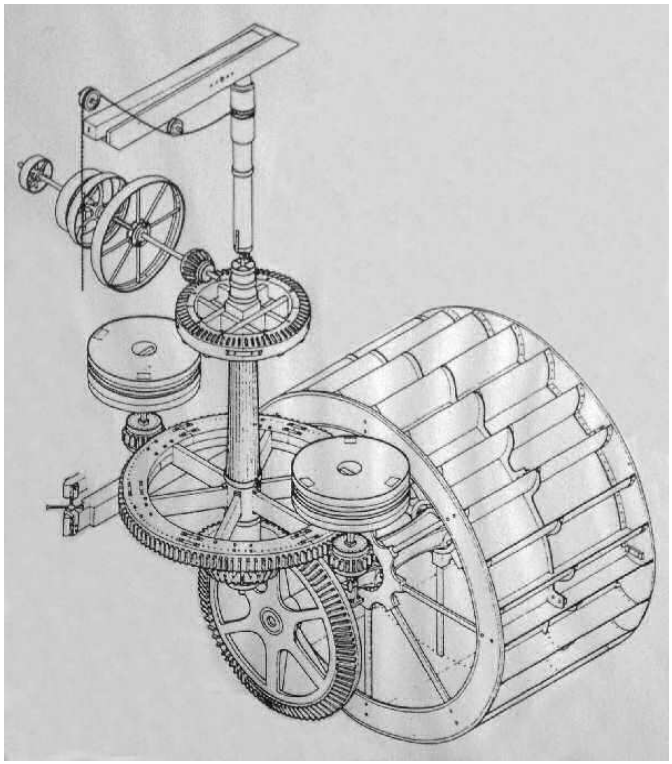


Working / operating environments are unstable and unpredictable. Actions / changes therefore often have unanticipated consequences.

Understanding how systems work

Understanding in terms of interconnected parts.

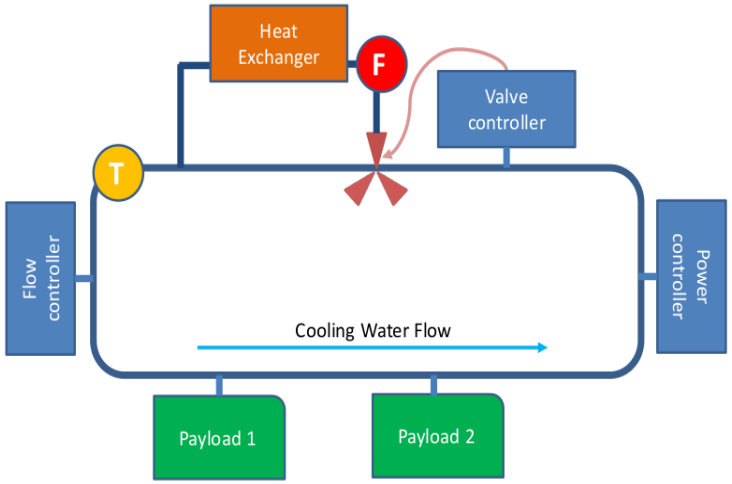
Understanding in terms of functions that depend on each other.



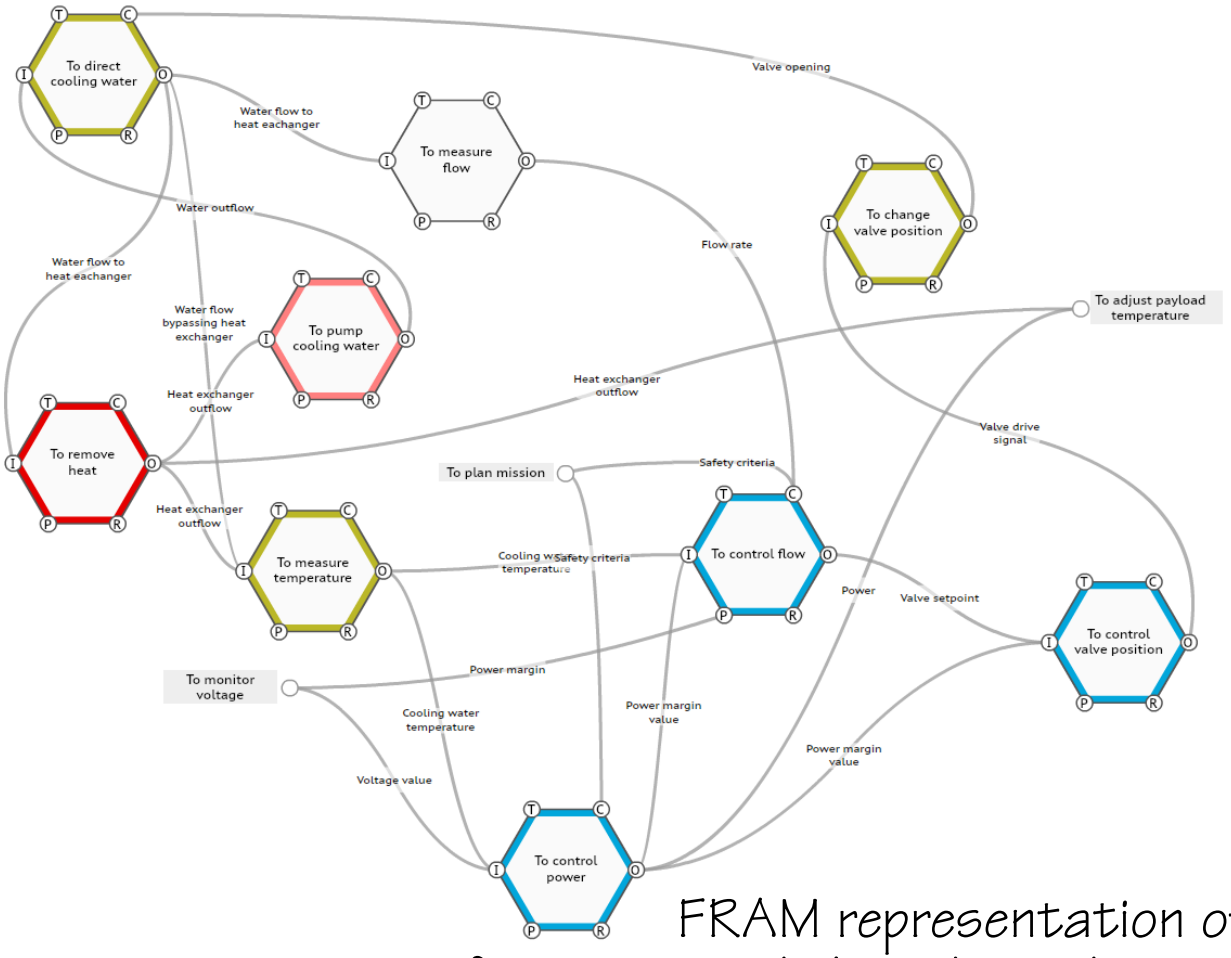
Few parts and well-defined (synchronous) connections

Many "parts" and ill-defined (asynchronous) connections.

System as parts or as functions



Classic description of parts / components and their relations.



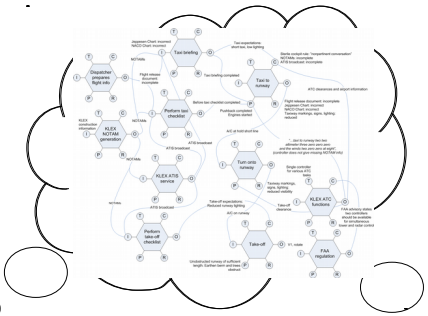
FRAM representation of functions and their dependencies

Functional non-linear model

Safety Synthesis

Non-linear models

If accidents can be understood as emerging from everyday performance adjustments ...

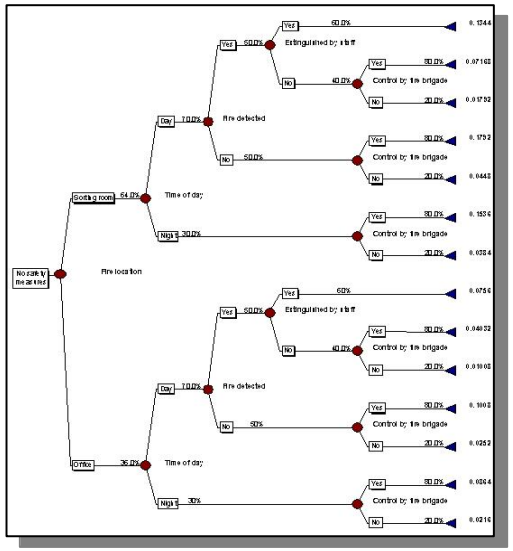


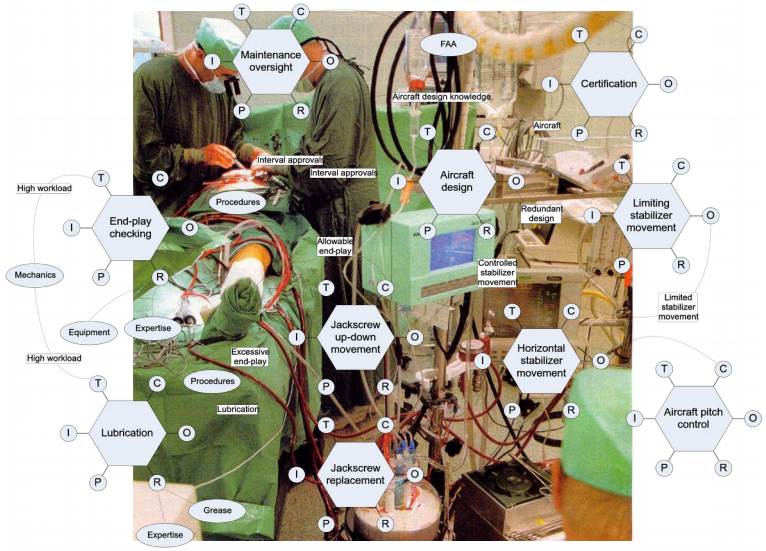
... then risks can be understood as emerging from everyday performance adjustments

Systems at risk are intractable rather than tractable.

The future can be understood by considering the characteristic variability of the present.

Common assumptions -then and now





System can be decomposed into meaningful elements (parts, events)

The function of each element is bimodal (true/false, work/fail)

The failure probability of elements can be analysed/described individually

The order or sequence of events is predetermined and fixed

Systems cannot be understood by decomposing them.

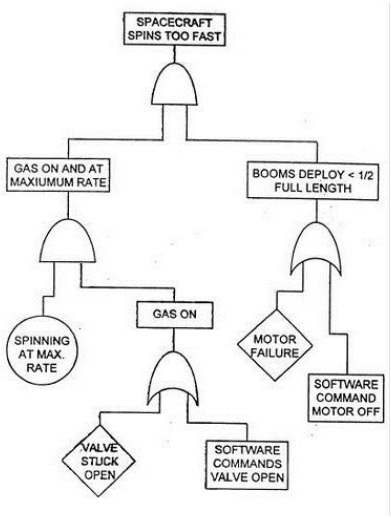
The functioning is not bimodal, performance is always variable.

Performance variability, is a source of success as well as of failure.

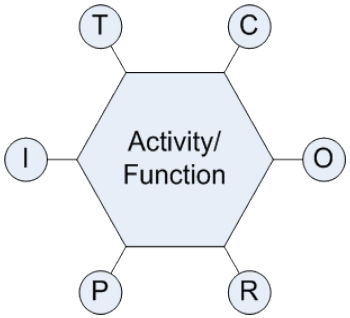
The order of activities must be flexible to fit the conditions.

Models and methods

An analysis of something inevitably involves some assumptions about how that something happens.
These assumptions correspond to a model: a simplified explanation of how something can happen and of how the 'world' is organised. The organisation usually implies some kind of hierarchical ordering of layers, parts, or components (structural models).



The model defines what the method can be used for, and therefore also sets the limits of the method.



The FRAM is a method to develop a representation or model of how something happens. This model can then be the basis for various kinds of analyses (reactive, proactive). A FRAM model represents the functions that sufficient and necessary for an activity to take place – not when it goes wrong but when it goes right.

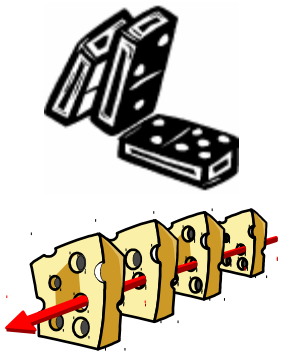
Three kinds of analysis

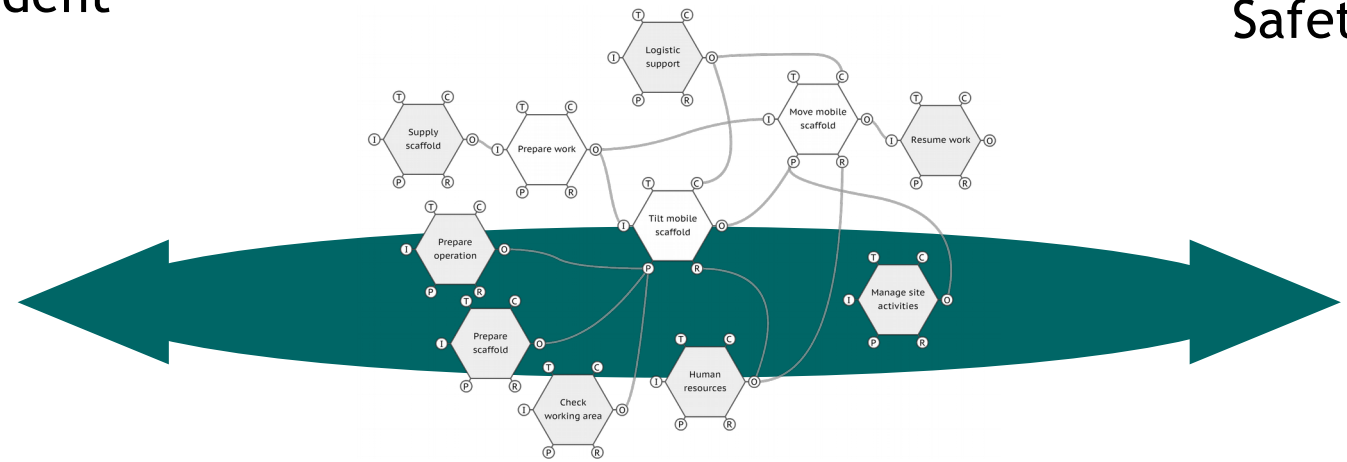
Analysis of the past (retrospective)

Analysis of the present (work-as-done)

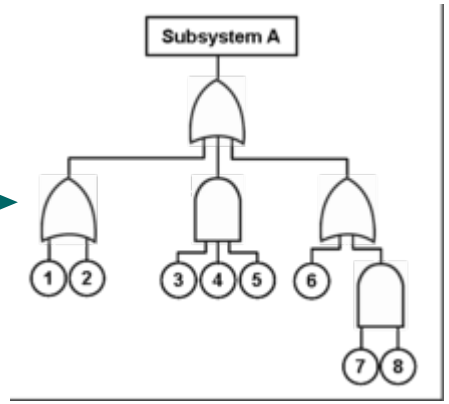
Analysis of the future (predictive)

Safety-I: Accident analysis





Safety-I: Risk analysis



Functional model of everyday activities.

Safety-II: Patterns of work, heuristics, habits

A FRAM model can be used for both retrospective and predictive analyses, and for Safety-I as well as Safety-II.

Safety-II: Feasibility, fitness for purpose, bottlenecks