GREAT BUT FLAWED EXPECTATIONS: ON THE IMPORTANCE OF PRESUMPTIONS AND ASTONISHMENT IN MODEL AND SIMULATION BASED RISK MANAGEMENT

Marko Hofmann[(a)], **Thomas Krieger**[(b)]

Institut für Technik Intelligenter Systeme (ITIS e.V.)
Universität der Bundeswehr München
Werner-Heisenberg-Weg 39
Germany - 85577 Neubiberg

marko.hofmann@unibw.de[(a)], thomas.krieger@unibw.de[(b)]

ABSTRACT

Unexpected detrimental events probably pose the most dangerous threat to every planning activity. They are the consequence of both explicit and unconscious presumptions made during the planning process. These presumptions are the manifestation of the modeler's own expectations, which can be seriously flawed. Model and simulation based risk management tries to identify potentially dangerous presumptions (for the real world planning) by looking for astonishing results in models in general and simulations in particular. The astonishment is triggered by (simulation) events that are violations of model assumptions (the model specific instantiations of the presumptions) or events which are simply counter-intuitive. The main idea of this approach is illustrated using examples taken from reliability theory. This choice has been made for didactical purposes: the analytical perspicuity of these examples is much better than the one of complex simulation models. Subsequently, the benefits of the approach are demonstrated for military conflict simulation models.

Keywords: uncertainty, simulation supported risk management, modeling assumptions, analysis and exploration, reliability theory

1. INTRODUCTION: MANAGING RISK

Risk management has become a paramount task of modeling and simulation in a great variety of applications. Hitherto finance markets and military endeavors have been the most prominent domains of risk management, but it seems to be indispensable for industrial applications, too. From a generalized point of view, risk (in the broader sense) has two dimensions: risk in the narrower sense and real uncertainty (Knight 1921). Risk (in the narrower sense) is associated with known dangerous events and the possibilities of their occurrence. These possibilities are regarded to be assessable, using "hard facts" (e.g. frequencies) (Risk type 1). Real uncertainty is an attribute of known dangerous events for which the possibility of occurrence is indeterminable on objective grounds (Type 2) and of completely unexpected events, which reveal their detriment only after they have happened (Type 3).

The more modern classification of parametric and structural uncertainty (which can not be attributed to a single origin) has a slightly different meaning, but addresses approximately the same distinction as type 2 and 3 uncertainty. Parametric uncertainty means that we know the relevant factors for a given phenomenon, but miss the exact (initial) values of these factors. In other words, we have good empirical evidence that the causal reasoning of the model we use is an adequate representation of the relations in the real world. What is sometimes hard to find are the "right" parameters for the model. Parametric uncertainty is roughly equivalent to uncertainty type 2 in the modified Knight's classification, but it also includes type 1, if we see probabilities as special expressions of parameter uncertainty. Structural uncertainty means that we are not sure if we know all the relevant factors and that we most probably do not know their interdependencies. Or, in other words, there are serious reasons to believe that the model we use to represent the phenomenon is at least incomplete. Hence, structural uncertainty ideally reflects the concept of type 3 uncertainty in the former classification, only substituting unknown factors for unknown events.

Ex ante, all three types of risk are equally important. In practice, on the contrary, the amount of work dealing with the first type of risk dominates the other two. This disequilibrium is due to the human inclination to operate on mathematically treatable information. Type 2 events are therefore often transformed into type 1 events using subject matter experts and their estimations as substitutes for objectively generated probabilties. Slightly simplifying matters, this paper treats risk type 2 as parametric uncertainty and risk type 3 as structural uncertainty.

The remainder of this paper is organized as follows: Section 2 outlines the difference we make between planning presumptions and modeling assumptions. Section 3 introduces the concept of astonishment. Section 4 provides the reader with an example of an astonishing result in the field of parametric uncertainty. Section 5 gives a second example which can be attributed to structural uncertainty. Section 6 discusses tactical wargaming, in general. Section 7 presents a simulation based wargaming method for parametric uncertainty, section 8 repeats it for structural uncertainty. Section 9 concludes the paper by reiterating its contributions and suggests some conclusions and future research directions.

2. PLANNING PRESUMPTIONS – MODELING ASSUMPTIONS

Making plans for the future of complex social systems is always affected by personal presumptions. They are unavoidable for many reasons. First, the perception and conception of human beings are limited by their experience and cognitive constraints (bounded rationality, see, e.g., Gigerenzer and Selten 2002). Second, planning in social systems is impossible without making predictions on the behavior of other humans. These predictions are necessarily unreliable.

Examples for such presumptions are:

- the equivalence of political “solutions” in different cultures,
- the superiority of western thinking and governing to all other forms,
- the impossibility of anomalies in hitherto well-understood systems,
- taking statistical correlations as causal dependencies.

In general, only a small part of such presumptions is mentioned during the planning process, a much greater part is unconscious, but detectable via questioning. Some presumptions are even difficult to detect, because they are deeply hidden in attitudes and beliefs. Such presumptions are classical examples of risk type 3.

If we deliberately reflect on our plans and make planning support models, the conscious part of these presumptions become modeling assumptions.

Modeling is seen here as a process that, if successful, helps a subject S to solve a problem P situated in an original (system) O at a given time T. The model M of the original O is always regarded as an abstraction (Stachowiak 1973). The notion of assumptions is concentrated in the following definition: An assumption A is a hitherto not (or insufficiently) empirically corroborated statement (in the sense of an assertion)

- of a subject S,
- about an original O,
- with the intention (purpose) I relative to
    - a problem P situated in O and
    - a timeframe T.

Some typical assumptions used in many (simulation) models are:

- Stability of processes,
- Uniformity of interactions over time and space,
- Linearity of interrelations,
- Processes have reached equilibrium,
- Empirical data fit approximately uniform, exponential, normal etc. distribution,
- Independence of statistical parameters.

In the narrow sense used here (for a broader view see Hofmann 2003), model assumptions are always explicit and therefore conscious, whereas some planning presumptions can still remain unconscious even after the modeling.

Model assumptions are examples for uncertainty of type 2.

In the following sections and examples, this distinction between explicit modeling assumptions as instantiations of risk type 2 and hidden personal presumptions as examples of risk type 3 will be further clarified.

3. ASTONISHMENT

With regard to the two types of risk scrutinized here, astonishment can be triggered by the violation of an explicit assumption or a hidden presumption. In the first case, during the realization phase an event occurs that seemed highly improbable during the planning phase. Consequently, the contrary had been assumed. In the second case, an event occurs which has not been expected during the planning. Some of such events can be advantageous for the own goals, but, in general, they are detrimental, because when things turn out different as supposed to be, plans cannot be accomplished optimally. Planners and decision makers must therefore try to reduce the amount of astonishing events during the realization phase.

The methodology to cope with events that have been considered improbable (type 2) differs from that applicable to unexpected events (type 3). In the first case (parametric uncertainty) it is always possible to reflect on the assumptions made. In tactical wargaming, a method developed for the German army, for example, all explicit modeling assumptions are deliberately taken as violated. The criticality and plausibility of this violation is estimated and contingency plans can be developed (see section 7).

Since we simply do not know unexpected events we cannot use this approach for type 3 risk (structural uncertainty). The default strategy of military decision makers for type 3 risk, for example, is the creation of reserves. Nevertheless, it would be advantageous to find as many potentially detrimental events in advance, because the planning of countermeasures could be much more specific (see section 8). The basic idea is to generate some astonishing results/events in a simulation, thereby reducing the amount of astonishing results/events in reality.

Before we discuss the application of these concepts within the framework of wargaming based on complex military simulation models, we will highlight the core idea with two analytical examples presented in the next two sections. These examples are taken for their simplicity, clearness and mathematical accuracy. Although the hard probabilistic reasoning might be somewhat misleading, the examples focus on the essential. The first example will demonstrate an unexpected effect in parameter variation and the second example is intended to illustrate a structural surprise.

The main difference between these examples and risk management in real world applications, like wargaming, is, that the astonishing results in the examples can be proven to be hard facts, whereas the simulation results in wargaming are generally only conjectures.

4. ON THE SAFETY OF AIRCRAFT WITH TWO OR FOUR ENGINES

In this section we consider an airline which wants to buy a new aircraft and has to decide between two types. One aircraft type (A_2) has two engines and the other type (A_4) four engines. It is assumed, that the decision of the airline's manager which aircraft will be bought, is made on the sole basis of the reliability of the aircraft. That also implies, for instance, that the different numbers of passengers which can be carried, are not taken into account.

The airline defines with respect to own experiences that aircraft A_2 (A_4) is working (i.e., is still flying and can touch down safe), if and only if at least one (at least two) of the engines are still working. Furthermore, it is assumed that

1. the engines are working independently of each other (an assumption which may not be fulfilled in all practical situations), and
2. the probability p, that an engine is working is the same for both aircraft and all their engines (one might think of one type of engine which is used in both aircraft).

In the past the airline has already bought a lot of aircraft. For making their decision they use a simulation which – for a given value of p – computes the reliability of the aircraft. It is important, that in the past the decision were made on basis of simulations for values of p close to 1, because the engines itself are very reliable under good flight conditions. Also, a changing of p during the flight was not taken into account. Now, in recent years it turned out that this assumptions have been too optimistic. In extremely bad weather the reliability of the engines can significantly decrease. Due to fierce competition on the air transportation market airlines have to fly even under such extreme conditions.

Let $P_2(p)$ resp. $P_4(p)$ be the reliability of aircraft A_2 resp. A_4 (which only depends on p), i.e., the probability that the aircraft is still flying and can touch down safe. For values of p close to 1 the M&S-Team of the research group had observed in the past the relation $P_2(p) < P_4(p)$ holds for all p close to 1. They now make – most probably unconscious - the implicit assumption – coming from their experience - that $P_2(p) < P_4(p)$ for all $p \in [0,1]$. This would lead to an absolute preference of aircraft A_4.

Skeptical about this reasoning the airline's manager demands critical rethinking of all assessments made on reliability. Using their simulation system the research group starts with p = 1 and obtain $P_2(1) = P_4(1)=1$. This is an obvious result, since all engines are working with probability 1 and therefore – per definition – the aircraft is working. Taking p = 0.9, an already low reliability rate, they get $P_2(0.9) = 0.99$ and $P_4(0.9) = 0.9963$, which satisfies their expectation that $P_2(p) < P_4(p)$ for all $p \in [0,1]$. Assured in their opinion about the superior reliability of four-engine aircraft, they try p = 0.6 and obtain $P_2(0.6) = 0.84$ and $P_4(0.6) = 0.8208$, i.e., $P_2(0.6) > P_4(0.6)$, and are absolutely astonished. This should be impossible! It seems unbelievable to them, that an aircraft with four engines can be less safe than a two-engine aircraft.

In fact, this result is not surprising at all. Using elementary probability calculus we get (see, e.g., Rohatgi 1976)

$$P_2(p) = 1 - (1 - p)^2$$

and

$$P_4(p) = 1 - (1 - p)^4 - 4p(1 - p)^3.$$

Both functions are presented graphically in Figure 1.

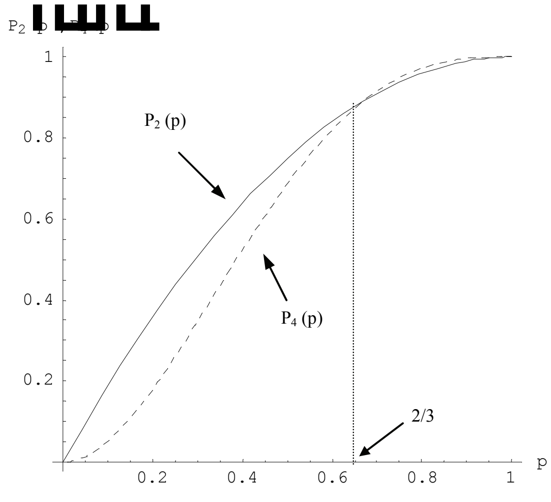

Figure 1: The Reliability Functions of Aircraft with two resp. four Engines.

We see that in case of $p \in (2/3,1]$ A_4 is more reliable than A_2, while in case of $p \in [0,2/3)$ A_2 is indeed more reliable than A_4. Consequently, the expectation
$P_2(p) < P_4(p)$ for all $p \in [0,1]$ is obviously false. The maximum difference between the system reliabilities P_2 and P_4 are 0.0128917 for $p \in (2/3,1]$ and 0.179558 for $p \in [0,2/3)$. That implies that a two-engine aircraft is, at worst, only 0.98 times less safe than a four-engine aircraft under good conditions but a four-engine aircraft can be as far as 0.52 times less safe than a two-engine aircraft under extremely critical conditions.

Of course, this reasoning cannot be a decisive argument to buy (or construct or use) only two-engine aircraft, since it completely neglects the actual distribution of p in reality (which may even be hard to find). However, the research group should have realized, that some assumptions they have made in judging the reliability of aircraft, may not hold in all cases. Such a self-critical attitude is the cornerstone of risk management in real applications.

By the way, in a group of 13 computer scientist and mathematicians we asked to estimate the two reliability functions, only one person has made the right guess.

5. NON-MONOTONIC RELIABILITY FUNCTIONS

A company wants to set up a network with three nodes and three edges (see Figure 2).

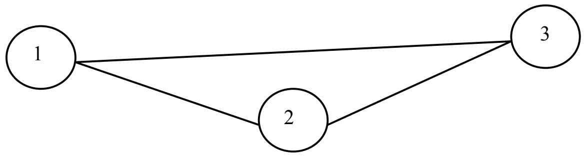

Figure 2: The Network.

It is known that

1. The nodes are working independent of each other with probability p.
2. The edges are faultless independent of each other with probability q.
3. Nodes and edges are working independent of each other.

The company defines that the network is working (i.e., reliable) if and only if there is a connection between all faultless nodes. Furthermore it is defined that if all nodes are failed the network is supposed to be not working. Let $P(q,p)$ be the probability that the network is working. The company´s M&S-Experts wants to get an idea of the reliability of the whole system depending on q and p. Due to their experience they expect that for a given q the reliability of the system is a monotone-increasing function of p. Moreover, they do not know any system that exhibits non-monotonic reliability behavior with respect to its components reliability. Furthermore they expect $P(q,1)=1$, since their first guess is, that if all nodes are failing then the whole system is failing and if all nodes are working with probability 1 the whole system should work with probability 1. Please, stop reading for a few seconds: What is your opinion? Which behavior of the reliability function $P(q,p)$ are you expecting?

The M&S-Experts of the company uses a simulation in order to get insights into the behavior of the system. In case of $q = 0.5$ they obtain $P(0.5,0.45) = 0.621$ and
$P(0.5,1) = 0.5$, i.e., the reliability of this system is in case of $p = 0.5$ larger than in case of $p = 1$. This outcome seems to them quite dubious. It is against the common understanding and the usual properties of reliability functions (see, e.g., Barlow and Proschan 1975). As mentioned above the experts are expecting a monotone-increasing function with $P(0.5,1) = 1$, which obviously is not fulfilled. These facts contradict their expert knowledge and this makes the system even more interesting to them.

The reliability function of the system can be computed as follows:

$$\begin{aligned} P(q,p) = {} & (1-q)^3\, 3\,(1-p)^2\, p \\ & + 3\,(1-q)^2\, q\,\left(3\,(1-p)^2\, p + (1-p)\, p^2\right) \\ & + 3\,(1-q)\, q^2\left(3\,(1-p)^2\, p + 2\,(1-p)\, p^2 + p^3\right) \\ & + q^3\left(3\,(1-p)^2\, p + 3\,(1-p)\, p^2 + p^3\right). \end{aligned}$$

The function $P(q,p)$ is drawn in Figure 3 for the cases $q = 0, 0.5, 1$.

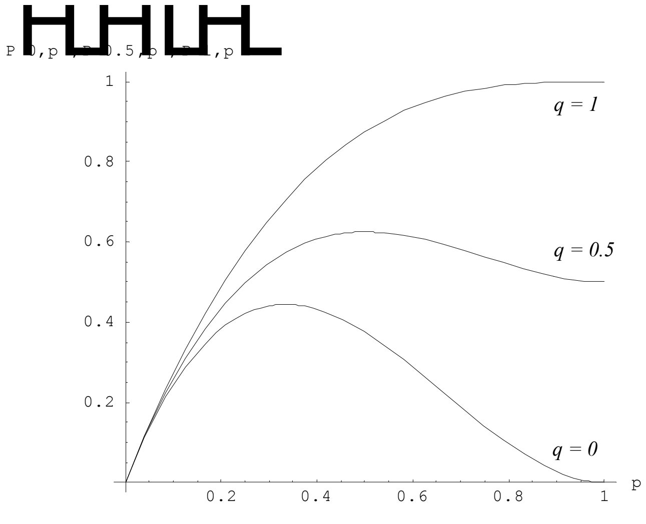

Figure 3: The Reliability Functions of the Network for the Cases q = 0, 0.5, 1.

As we can see we have P(0.5,0.5) > P(0.5,1). The reliability function P(q,p) – not only for fixed values of q - of the system as a function of q and p is depicted in Figure 4.

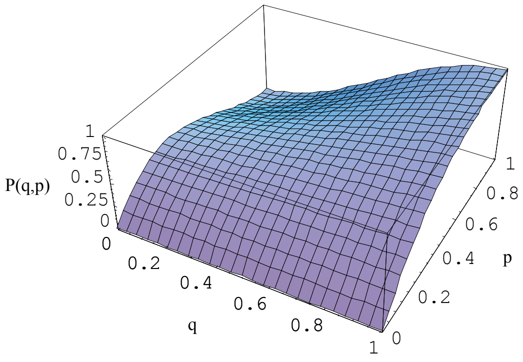

Figure 4: The Reliability Function P(q,p) of the Entire System.

The reason why this system does not behave like the experts expected it to do, is that it is not a monotonic system, i.e., if a failed element is repaired, then the system may change from a working system to a failed system. The reliability analysis of such systems is, in our opinion, a very interesting challenge for the future.

6. WARGAMING – A MILITARY TOOL OF RSIK MANAGEMENT

“Wargaming is a flexible instrument designed to develop, compare, and improve courses of action (COA)”. This is a definition given by the German General Staff College. It can also be seen as one of the most important methods of risk management within the military domain. The origin of institutionalized Wargaming dates back to emerging Prussia and its General Staff in 19th century and can be described as means to interactively play the *uncertain* development of a military (and later on non-military as well) operations. Both, Clausewitz and Moltke the elder saw the potential of Wargaming in a staff being better prepared for the incalculable course of an operation (v. Clausewitz 1832). They also advised it for the “free play” of its members’ creativity and the intensified examination of planned COA (Hofmann and Lehmann 2007).

Military leadership is a domain where personal presumptions and, consequently, model assumptions have always played and are continuing to play a decisive role. The “model” used from commanders and their staffs is usually an abstract representation of own and enemy units on a terrain representation. Most of the explicitly mentioned presumptions of commander and staff are depicted within this environment. They are therefore visible or visualizable model assumptions (like the currently perceived situation or the supposed enemy commander’s intent), which can be directly addressed. Some presumptions (like, just for example, the general superiority of mission-type tactics over order-type tactics and the higher value of special forces even in standard combat), in contrast, are not mentioned and therefore also not challenged.

Wargaming incorporates, even traditionally, two rather different aspects of investigation: analysis and exploration. In general, wargaming for COA *analysis* is based on a sequential process of some largely independent cycles (repetitions) of “moves” from at least two different teams within a gaming environment (sand table, map or computer based system). Within this paradigm it follows an Action – Reaction – Counteraction pattern using an impartial umpire to judge the outcome. The outcome can also be calculated using fixed rules (like in chess) or simulations (see section 7).

The central rationale of this analytical aspect of wargaming is the scrutiny of a succession of events that constitute a COA. Since all COA are given before the wargaming starts, the range of possible events is limited to what the team members regard as possible within that COA. From a generalized point of view, this kind of wargaming tries to reduce the uncertainty of known events and is therefore dealing with a special kind of parametric uncertainty (the decision relevant parameters of the COA). This technique can be applied on explicit model assumptions, too. One simply negates them and scrutinizes the effects.

The explorative aspect of wargaming is often hidden to the external observer, since it follows no general methodological rules and is basically a mental activity. Exploration means to think outside the standard analytical evaluation scheme. It is a questioning of own beliefs and assumptions and a speculative search for chances and exceptional risks. In such, explorative wargaming is an attempt to deal with structural uncertainty.

Both aspects of wargaming can be supported by combat or other military simulation systems. However, there is a significant distinction in the methodologies used to deal with parametric (already in use) and structural (currently tested within the German army) uncertainty (Hofmann and Junge 2008).

7. SIMULATION BASED WARGAMING FOR RISK TYPE 2 (PARAMETRIC UNCERTAINTY)

It is possible to use a stochastic simulation system instead of an umpire to evaluate the outcome of different COA. The wargaming factions implement their plan (COA) by fine-tuning the deployment of their units and giving initial orders. After the simulation has started all elementary combat processes, including off-road combat mode movement, reconnaissance and mutual attrition, are automatically simulated using random generators to mimic the effect of parametric uncertainty. In an interactive simulation it is possible and necessary to command the units by interrupting the simulation run, whereas in a closed simulation no human intervention is possible after the start.

The major advantage of interactive simulations is more flexibility in the command of the COA realization. However, this flexibility has two drawbacks for wargaming: First the evaluation of COA is dependent on human knowledge about the intricacies of the special simulation model and dependent on the human skills in the operation of the simulation. Second, interactive simulations are much slower than closed simulations (the difference can amount to several orders of magnitude). As a consequence of this limitation and of learning effects, it is seldom possible to make more than a few repetitions with the same (own and enemy) COA combination. Thus, there is little variation included into the evaluation. The parametric uncertainty of real combat is therefore often underestimated.

With closed simulation it is possible to make hundreds of simulation runs using different random numbers for all kind of elementary processes (movement, attrition, reconnaissance, communication etc.). Although this variation is somewhat compensated by its military "blindness" with respect to an interactive human commander, it can nevertheless capture a huge range of parametric uncertainty. The standard routine to take this uncertainty into consideration is the computation of measures of central tendency and dispersion (mean and variance, for example) (Ross 2002). For simplicity, let us assume that every simulation run ends with an exactly measurable result somewhere between a clear success (100) and a complete failure (0). A frequently used setting in standard military wargaming consists of three own COA which are compared with two enemy COA (usually the most likely and the most dangerous COA). A possible simplified result of a closed-simulation based statistical evaluation within this setting could look like Table 1.

Table 1: Simulation based Statistical Evaluation of COA (Example)

Most likely Enemy COA	Number of runs	mean	variance
COA 1	100	60	30
COA 2	100	55	20
COA 3	100	50	10

Most dangerous Enemy COA	Number of runs	mean	variance
COA 1	200	50	30
COA 2	200	40	10
COA 3	200	50	10

In order to get a definite result from these numbers it is necessary to weight between the two cases and the statistical measures. If we, for example, equally weight most likely and most dangerous enemy COA and ignore variance, COA 1 must be favoured because it has the highest average mean (55). A risk aversive decider would discount means by a certain proportion of the variance and would put a higher weight on the second case (the most dangerous enemy COA). This could lead to the preference of COA 3.

It is obvious that uncertainty in this approach is attributed as a statistical parameter to the given COA. The random effects in such a simulation system are produced by well understood random generators on the micro level of elementary process. However, the complicated interrelation and interaction of these processes can lead to a macro phenomenon (the overall combat, the outcome of a COA pair) which is astonishing. Since no combat simulation system can claim to be a valid representation of the future, the uncertainty measures created by this method are only measures of risk type 2 and should not be misinterpreted as measures of type1.

This standard reasoning can be easily transformed to a new method of dealing with modelling assumptions. Instead of analyzing the COA, explicit model assumptions are negated, implemented and the consequences of this negation analysed via stochastic simulation. The main methodological difference between the standard procedure and the special method of assumption-centred stochastic simulation is to take variance as decision criteria in the latter much more serious than in the former, since our focus of interest is the uncertainty attributed to the assumptions and not a disputable mean. The idea is to classify the analysed assumptions according to their criticality and plausibility (see Dewar 2002 and Hofmann 2007 for further information).

It should be mentioned that military simulation systems have reached an unprecedented level of complexity (Hofmann 2005). The example and reasoning presented are simplifications. Therefore, it now should have become clear, why we have chosen the aircraft example in section 4. It would have been extremely difficult to describe a wargaming example with the same completeness.

8. SIMULATION BASED WARGAMING FOR RISK TYPE 3 (STRUCTURAL UNCERTAINTY)

The basic idea of this approach is to use closed simulations *to detect* (not analyze) critical assumptions (and thereby actively dealing with structural uncertainty). The major advantage of closed simulations in comparison to interactive ones is, as already mentioned, the much greater speed of the former. It is therefore possible to run hundreds or even thousand of simulation runs within the time available for the decision making. The crucial question is, how *structural variability* can be introduced into the systems? By the use of random generators for the elementary process it is only possible to generate parametric variance. What is needed are random effects on the level of events and within the command and control modules. Random events can be easily generated if the demand for valid representations is completely given up. Then, a random event can be, for example, a regular event (detection, shot, etc.) without cause or a randomly chosen event from a historical data base. Introducing randomness into the command and control modules is somewhat more difficult, but nevertheless possible, if the notion of optimal behavior is neglected. By deliberately generating suboptimal behavior via random functions, the behavior of a command and control module becomes incalculable. Which is exactly what is intended. The key concept of the evaluation of such simulation runs is the exclusive debriefing of their extremes and the abdication of statistical reasoning. Disastrous simulation runs (from the perspective of the own planning) are taken as possible threats (implying critical assumptions), extremely advantageous runs are taken as possible chances. Means and measures of variance are not investigated, because the model is invalid anyway. All extreme simulation runs have to be checked by human experts, which can quickly discard them as completely implausible or further scrutinize them, because they appraise the chain of events as possible regardless its inconsequent creation within the simulation system. Their main function of this approach is to broaden the view of the planner with respect to unexpected future trajectories.

Invalid representations, deliberately modeling suboptimal behavior of automated forces and the renouncement of statistical evaluation for the benefit of mere extremes may seem absurd at first glance, but taken together and seen from the perspective of structural uncertainty they make perfect sense. Such simulations can be seen as explorations into the hidden realm of personal presumptions, that might be challenged by some extreme runs.

However, it is necessarily clear, that even this approach can not capture the whole range of possibilities spanned by real systems. Their major contribution in military (and maybe other kinds of) risk management might be, to open the decision maker's view to the completely unexpected by simply confronting him astonishing courses of action.

Reflecting this section it is now easily possible to explain the connection between the example in section 5 and simulation based tactical wargaming for structural uncertainty. Reliability functions that are non-monotonic in their components' reliability are extremely rare in practice and teaching. Thus, such an example will be astonishing for most engineers and students as well, recognizing their own presumptions about the subject (Hofmann and Lehmann 2007).

9. SUMMARY, CONCLUSION AND OUTLOOK

In most real applications risk is a multifaceted problem comprising objective and subjective probabilistic dangers as well as completely unknown threats. Some of the most critical uncertainties can be attributed to hidden personal presumptions and hitherto unquestioned modeling assumptions. Thus, an exhaustive methodology of risk management has to incorporate some strategies to cope with this challenge. In order to operationalize presumptions and assumptions we propose to see personal presumptions as part of what is called structural uncertainty, and the systematic questioning of explicit modeling assumptions as part of dealing with what is called parametric uncertainty.

With two simple examples we tried to demonstrate the importance of astonishing model/simulation results for the detection of (parametric model) assumptions and (structural) presumptions in the reliability of technical systems.

As a serious domain of application we subsequently introduced tactical wargaming. Two different methods of simulation and experimental design have been presented for parametric and structural uncertainty. The first approach follows traditional statistical reasoning for COA comparison, while stressing the importance of variance as opposed to means. In complete contrast to classical stochastic simulation and its evaluation by means of statistical reasoning, the second method is focused on the unexpected, uncommon, exceptional, ignoring validity and optimality. It is a method of thought triggering and definitely not of hard deduction.

Recent experiments during three major exercises have demonstrated the value of the unusual approaches, broadening the view on risk in general, and unexpected events in particular.

We are convinced that the methodologies do not only apply to military problems, but also to all kinds of other domains with parametric and structural uncertainty, including economic and industrial applications.

However, much work remains to do. First of all, more experiments have to be done in order to fine-tune the methods. Second, the value of the methods, especially the second one, heavily depend on unpredictable but sensible variability of the simulation system. For that purpose, much expertise has to be integrated into the respective modules. Third, we would really like to demonstrate the benefits of our approach in other domains.

REFERENCES

Barlow, R., Proschan, F., 1975. *Statistical Theory and Life Testing (Probability Models)*, New York: Holt, Rinehart and Winston.

Clausewitz, C. v., 1832. *Vom Kriege*. Reinbek: Rowohlt Tb. 1978.(also available on www.clausewitz.com)

Dewar, J. A., 2002 *Assumption-Based Planning: A Tool for Reducing Avoidable surprises*. Cambridge UK: Cambridge Press.

Gigerenzer, G., Selten, R., 2002. *Bounded Rationality: The Adaptive Toolbox*. Massachusetts : MIT Press.

Hofmann, M. 2003. Modeling Assumptions: How they affect Validation and Interoperability. *Proceedings of the 5. European Simulation Interoperability Workshop*, Stockholm.

Hofmann, M., 2005. On the Complexity of Parameter Calibration in Simulation Models. In: *Journal of Defense Modeling and Simulation 2, (4).*

Hofmann, M., Lehmann, T. 2007. On simulation-based Wargaming: Comparison of two different methodological approaches. *Information and Security.* Vol. 22, 64-72.

Hofmann, M., Lehmann, T., 2007. *Tactical Wargaming, the Coup D' Oeil and IT-Support-Systems*. Third International Conference On Military Technology Miltech 3, 129-146, Stockholm, Sweden.

Hofmann, M., Junge, B., 2008. *Dealing with Structural Uncertainty in Tactical Wargaming*. In: *Proceedings of the 9. European Simulation Interoperability Workshop*, Edinburgh, Great Britain.

Knight, F. H., 1921. *Risk, uncertainty and profit.* Hart, Schaffner & Marx, MA: Houghton Mifflin Company Boston.

Rohatgi, V. K., 1976. *An Introduction to Probability Theory and Mathematical Statistics.* New York: John Wiley & Sons.

Ross., S. M., 2002. *Simulation.* 3rd ed., San Diego: Academic Press.

Stachowiak, H., 1973. *Allgemeine Modelltheorie.*, Wien, New York: Springer Verlag.

AUTHORS BIOGRAPHIES

MARKO HOFMANN is Project Manager at the Institut für Technik Intelligenter Systeme (ITIS e.V.) an der Universität der Bundeswehr München, Germany. After his studies of computer science at the Universität der Bundeswehr München, he served two years in an army battalion staff. From 1995 to 2000 he was research assistant at the Institute for Applied System Analysis and Operations Research (IASFOR), where he got his Ph. D. in computer science. Since April 2000 he is responsible for basic research in applied computer science at the Institut für Technik Intelligenter Systeme (ITIS e.V.). He gives lectures in operations research, computer science and mathematics at the Universität der Bundeswehr München and at the University of Applied Science in Kufstein.

THOMAS KRIEGER is a Research Assistant at the Institut für Technik Intelligenter Systeme (ITIS e.V.) an der Universität der Bundeswehr München, Germany. He received his Diploma in Mathematics in 2000 from the Technische Universität Dresden, Germany and his Doctorate in Mathematics in 2003 from the Universität der Bundeswehr München. His research interests include Modelling and Simulation as well as Game Theory.