Modeling Uncertainty in Requirements Engineering Decision Support

Pedrito Maynard-Zhang, James D. Kiper
Department of Computer Science and Systems Analysis
Miami University
Oxford, Ohio 45056
USA
{ maynarp, kiperjd }@muohio.edu

Martin S. Feather
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California
USA
martin.s.feather@jpl.nasa.gov

Abstract

One inherent characteristic of requirements engineering is a lack of certainty during this early phase of a project. Nevertheless, decisions about requirements must be made in spite of this uncertainty. Traditionally, we have handled this uncertainty by continually revisiting requirements at later stages of a project when some of the uncertainty has been eliminated. However, we can supply a better level of decision support if we can model and tolerate uncertainty in various requirements.

Here we describe the context in which we are exploring this, and some initial work to support elicitation of uncertain requirements, and to deal with the combination of such information from multiple stakeholders.

1. Introduction

One of the most important techniques for requirements engineering and early design phases in the engineering of complex systems is the capture of human expertise. The expertise of its engineers is a critical resource for the Jet Propulsion Laboratory (JPL) and NASA whose use must be exploited and leveraged effectively. In early lifecycle phases, this expertise takes the form of expert judgments. These early phase judgments are vital, but are, by nature, characterized by some level of uncertainty. However, in many cases, expert engineers not only can make judgments, but can also make meta-judgments. That is, they can make judgments about the degree of certainty that they have in their judgments.

Uncertainty can be precisely modeled through an appropriate probability distribution. Most engineers understand this idea. However, many are not sufficiently expert in this area of statistical modeling to permit them to specify all of the parameters of a probability distribution to model their own uncertainty. They clearly would understand the meaning of the mean of a probability distribution, and have a rough feeling for the standard deviation of a distribution (although they may not be able to specify a number for the appropriate deviation or variance.) Typically, their understanding of other distribution parameters is even less complete than that of the standard deviation.

In this work, we propose ways of capturing this uncertainty through some visualizations of probability distributions. By allowing engineers to manipulate these visualizations, they can capture, for example, their degree of uncertainty in a judgment of the amount of risk in a particular action without their having to understand probability distribution parameters at a deep level.

This work is couched in the context of requirements engineering and early-phase design activities for space mission and supporting technology at JPL. This context specifically includes the Defect Detection and Prevention (DDP) tool and method that has been developed and is used at JPL. The DDP tool and methods are discussed in the next section.

The primary focus of this work is the importance of uncertainty modeling in requirements engineering and early-phase design of complex systems. In section 3 we describe uncertainty modeling through probability distributions, our visual tool for capture of this information, ways of combining information from multiple experts, and potential uses of this meta-information. In section 4, we describe more specifically how this meta-information can be used in the context of the DDP process and tool; section 5 covers related work.

2. Context – DDP and requirements

"Defect Detection and Prevention" (DDP) is a risk analysis process and custom software developed and used at JPL. DDP's primary use is risk assessment in the early stages of projects. To do this, risk information is elicited from experts, using the DDP software to capture their information on-the-fly, to present the growing aggregation of their information via various visualizations, and assist them in decision-making based on that information. For more details, see [4].

The risk information gathered in the course of DDP sessions is an instance of the human expertise with which this paper is concerned. In DDP, this includes experts' judgments on risk-related information: the risks themselves - what are they, which "objectives" (a.k.a. "goals" or "requirements") they threaten (and by how much – in common risk parlance, their "severity"), and their likelihoods; the aforementioned objectives, and their relative importance; and finally the so-called "mitigations" (options for reducing risks) – their costs, which risks they reduce, and by how much.

The sum total cost of all the identified options for reducing risks usually far exceeds the resources available; hence one of the primary purposes of using a method such as DDP is to determine a cost-effective selection of mitigations. It is apparent that this is an optimization problem – objectives, risks and mitigations comprise a model: mitigations are *options* (each one can be chosen, or not), and for a given selection of mitigations, the *cost* (sum of costs of the selected mitigations) can be computed, and the *benefit* (sum of expected attainment of objectives, taking into account the risks that threaten them, where those risks have been reduced by the selected mitigations) can also be computed.

In practical applications of DDP the experts have provided us with dozens to hundreds each of objectives, risks and mitigations, and thousands of connections among them. This highlights the need to follow an effective process for elicitation of such information. It also highlights the need for computer support for investigations of the accumulation of this information – (e.g., for 50 mitigations, there are 2^{50} ways of selecting from among them). For this purpose we have applied heuristic search techniques – the familiar ones of Simulated Annealing and Genetic Algorithms, and some novel machine-learning based techniques involving abduction and *treatment learning*. See [2, 10].

An important point to note is that these studies are typically done early in the development lifecycle – the time when there is maximum influence on the development to follow (because few commitments have been made), and yet foresight is hazy. Typically the novel aspects of whatever we are considering means that information from past projects is of only partial guidance, and that expert judgments must be utilized. As a result *uncertainty* pervades this process.

Our reaction to date has been to deal with this at the *decision* end of the process. The DDP process gathers expert judgments in the form of discrete point-estimates (e.g., this risk's likelihood is 0.7), or, on occasion, min/max range estimates (e.g., this risk's likelihood is at least 0.6 and at most 0.8). The DDP software computes with these point (or range) values. During the decision-making phase we compensate for the uncertainty that lies behind these numbers: for example, when using heuristic search to locate a (nearly) optimal solution for a given cost bound, we do not limit out attention to the one "best" solution returned by the search, rather, we explore interesting alternatives in the neighborhood of that "best" solution. These include sensitivity analysis, data clustering, and visualization of these clusters. [3]

There are good reasons to seek to deal with uncertainty from the information elicitation phase onwards. How this can be done efficiently and effectively is the goal of the remainder of this paper.

3. Uncertainty Modeling

One of the weaknesses of the DDP ontology is that it does not attempt to model the uncertainty in the expert opinion that is collected. However, experts frequently entertain a degree of uncertainty about the opinions they express. Furthermore, they are often quite capable of reflecting deeply on this uncertainty and describing it in a way that can be represented mathematically. (In a later section, we refer to some of the literature in related fields in which experts have been able to provide this type of insight.)

For example, although an expert may be uncertain about the exact cost of a mitigation, she may be willing to specify a simple distribution over the possible costs most closely matching her uncertainty, e.g., the mean and standard deviation of a Gaussian distribution. Similarly, when rating risk, experts can often quantify their certainty over the risk likelihood via a distribution. Eliciting and explicitly representing this uncertainty provides us with a wealth of information that can potentially reduce the space of viable options further by preferring those based on greater certainty. Having a measure of the certainty of experts' judgments, or lack thereof, is surely a valuable asset to a project manager. With minimal training about the meaning of probability distributions, we believe that experts at NASA and elsewhere will be able to provide such information about their certainty.

Successfully solving this problem requires addressing four issues: representation of uncertainty, elicitation of uncertainty, resolution of disagreements between experts, and effective utilization of this information in the decision-making process.

3.1. Uncertainty representation

By far the most common formalism for representing uncertainty is probability theory. This is due in no small part to its clear, well-understood semantics. In contrast, many alternative proposals for codifying expert uncertainty (e.g., certainty factors) can create unintended inconsistencies in a knowledge base. [13] Much of the recent work in artificial intelligence on representing uncertainty has focused on compact representations of multivariate distributions (e.g., Bayesian networks). This work is largely irrelevant from the perspective of our work, however, because the distributions describing certainty are always over single variables. Thus, it is sufficient to select distributions from standard parametric families such as Gaussians (for continuous variables) or binomials (for discrete variables).

3.2 Uncertainty elicitation and capture

The traditional approaches to probability elicitation developed in the decision analysis community include visual methods such as probability scales or wheels and indirect methods based on lotteries. [11] Again, much of the recent work on probability elicitation has focused on the challenging problem of effective elicitation techniques for complex, multivariate distributions [1, 14]. We do not encounter many of these challenges due to the fact that we only elicit distributions over single variables. Providing experts with standard parameterized families to choose from further simplifies the elicitation process.

We have developed a modest tool to help capture estimates of uncertainty as represented by probability distributions. As a starting point, we present a picture of a standard probability distribution whose mean is the value previously reported by the engineer. (See Figure 1.) We then allow the users to manipulate this visualization in ways that characterize their understanding of the uncertainty in this judgment. In this version of the tool, the distributions are from the Beta distribution family.

For example, the users may conclude that, if their judgment is wrong, then they have most likely underestimated the value. That is, the probability distribution is skewed to the right. Or they may assert that they have a high degree of certainty, so that the standard deviation of the probability distribution should be smaller. The interface makes this manipulation easy. Figures 2 and 3 below illustrate this interface with these choices.

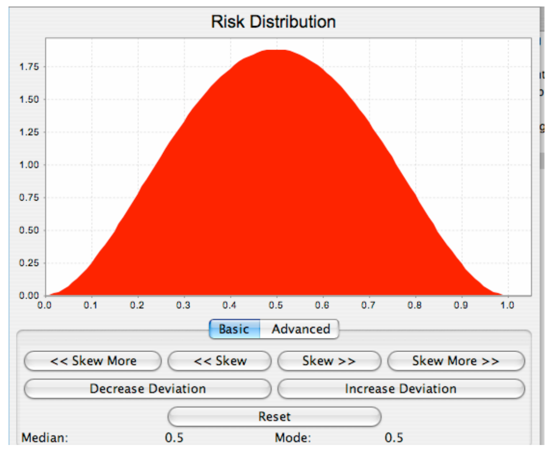

Figure 1: Initial probability distribution given a mean

The pertinent parameters for the selected probability distribution are captured from the interface, and saved with other project information. Note that experts are allowed to manipulate the parameters in a visual way until they are satisfied. They are not required to understand the associated parameters. (However, these parameters are available for their review, and for direct modification by those with a deeper level of knowledge and experience with such parameters.)

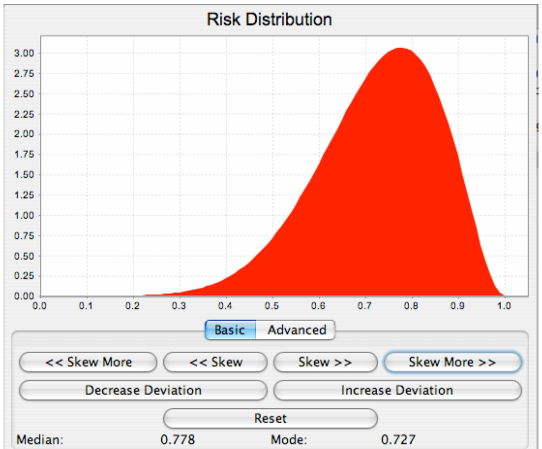

Figure 2: Probability distribution skewed to right

In using a process like DDP or similar collaborative methods and requirements elicitation or early-phase design, hundreds of opinions are elicited from experts. We cannot hope to ask these experts to evaluate the uncertainty probability for every one of these. Instead, we can ask for this insight on an as-needed basis, that is, on the critical issues as determined by the abduction, treatment learning, and decision clustering phases. Of course, we also allow for experts to specify uncertainty even when not asked. Opinions for which no distribution has been explicitly specified we assume to have the default distribution of certainty, i.e., where the given value has a probability of 1 and all other values have a probability of 0.

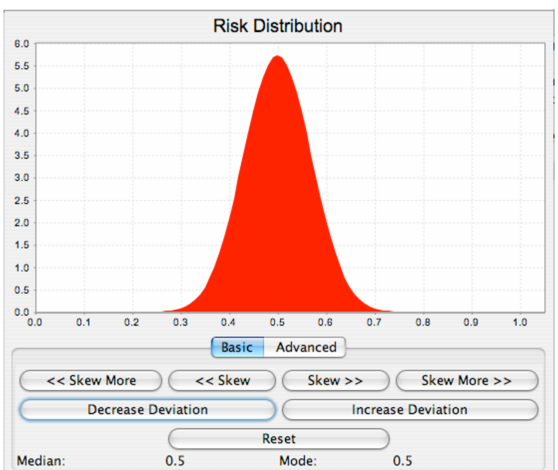

Figure 3: Probability with variance reduced

3.3 Resolving differences among experts

Given a measure of uncertainty over each variable gives the system additional power to distinguish between preferable worlds, as we will describe momentarily. However, an immediately obvious problem arises: multiple experts may specify different distributions over the same variable, but the system relies on each variable having a unique certainty distribution being associated with it. What is needed is a measure of the aggregate certainty with regards to each variable. The problem of aggregating probability distributions has long been studied in statistics and, more recently, in artificial intelligence. One simple technique that has received a great deal of attention is the Linear Opinion Pool or LinOP which computes a distribution that is a weighted sum of the experts' distributions.

Definition *Given probability distributions* $P_1, \ldots, P_L$ *and non-negative parameters* $\beta_1, \ldots, \beta_L$ *such that* $\Sigma_i \beta_i = 1$, *the LinOP operator is defined such that, for any marginal instantiation* $\boldsymbol{X}=\boldsymbol{x}$,

$$\text{LinOP}(\{\beta_1,P_1\}, \ldots, \{\beta_L,P_L\})(\boldsymbol{x}) = \Sigma_i \beta_i P_i(\boldsymbol{x}).$$ ♦

For example, assume that we are trying to estimate the failure rate of a device called devX. We have 2 experts A and B, whose opinions we seek to achieve our goal.

Expert A:

He thinks that devX will break with a probability of at most 0.2. (This is probably too high for a failure probability, but for the sake of simplicity we will use it in the example). When asked to clarify what he means by 'at most P=0.2', he explains:

- With a 30% probability, *devX* will fail to operate with a probability of 0.10
- With a 30% probability, *devX* will fail to operate with a probability of 0.15
- With a 40% probability, *devX* will fail to operate with a probability of 0.20

Expert B:

She thinks that devX will break down with a probability of at most 0.1. But she adds that there is an equal possibility that it may break with a probability of 0.05.

- With a 50% probability, devX will fail to operate with a P=0.05
- With a 50% probability, devX will fail to operate with a P=0.10

In order to combine the experts' opinions into one distribution, we take the weighted average their distributions for each value of the failure rate variable.

Suppose we trust both experts equally so that we weight their opinions equally, that is, $w_A = w_B = 0.5$. Then, for example, the LinOP aggregate probability for a 10% failure rate would be

$$\begin{aligned} &\text{LinOP}(A, B, \text{devX fail at } 10\%) \\ &= \sum w_i P_i(\text{devX fail at } 10\%) \\ &= w_A P_A(\text{devX fail at } 10\%) + w_B P_B(\text{devX fail at } 10\%) \\ &= 0.5(0.3) + 0.5(0.5) = 0.4 \end{aligned}$$

Doing the same calculation for all values of the failure rate variable, we get the distribution describing the aggregate uncertainty.

In spite of its simplicity, intuitiveness, and popularity, LinOP has often been dismissed as a normative aggregation operator, primarily because it fails to satisfy some properties deemed reasonable of all such operators (such as commutativity of aggregation and conditioning – that is, the operations of aggregating a distribution with another distribution and conditioning it with evidence should commute), but also because the weights are often chosen in an ad hoc manner. However, Maynard-Reid II and Chajewska [9] show that when it is accurate to describe the experts as having arrived at their uncertainty by learning from their experience in a way that approximates maximum likelihood estimation, then LinOP is exactly the correct operator to use where the weights correspond to the amount of experience each expert has had. We believe that the experts in Team X approximate these criteria sufficiently well to legitimize the use of LinOP.

3.4 Uncertainty Utilization

Now, given that every variable has an uncertainty distribution associated with it (whether explicitly or implicitly), if the system encounters two solutions that are in all ways equally preferable but where one exhibits a lesser degree of certainty on some of its variables, the system can safely remove that solution from consideration. More precisely, we can use the additional uncertainty information to eliminate solutions estimated to have lower *expected utility*. The expected utility of a world is its probability multiplied by its utility. Decision theory has shown that the most rational choice in any decision problem is to select a world that maximizes expected utility. In our context, utility is a function that seeks to maximize requirement benefits while minimizing mitigation costs. However, we do not have direct access to the probability of each solution. Because of the difficulties inherent in eliciting such a global joint distribution, we elicit local distributions over variables instead as described above. Presumably, these would be the corresponding marginal distributions of the joint had it been specified. In general, we cannot compute the joint distribution – and, thus, expected utility – given only marginal probabilities. However, if we assume that all variables are mutually independent (both conditionally and unconditionally), then the expected utility of a solution reduces to the product of the probabilities of the individual variable values that define the solution and the utility of the solution. This independence assumption is strong, but we believe the approximation will prove acceptable in this domain and that the benefits in elicitation and expected utility computation savings will far outweigh the slight loss of accuracy. Finally, we replace the utility measure used by the search to rank solutions with the estimated expected utility measure so that solutions are chosen that maximize expected utility.

Taken as a whole, our proposal incorporating uncertainty can be seen as an approximation of the decision-theoretic methodology. The local elicitation of uncertainty approximates the elicitation of a distribution over the solutions and the iterative combination of search, clustering, and localized expected utility maximization approximates the global expected utility maximization computation. We expect experimental results to show this strategy to outperform the original system that ignores uncertainty, modulo the validity of our assumptions.

4. Application to DDP

In subsection 3.4, we described expected utility maximization in general terms. Here, we apply this to the specific case of incorporating uncertainty in DDP. The DDP process requires engineers and domain experts to make scores, or even hundreds, of judgments about requirements, their weights, risks, their *a priori* likelihoods, mitigations, their costs, and strengths of relationship among requirements and risks and among risks and mitigations. We can clearly not ask these engineers and experts to take more of their valuable time to give probability estimates for each of these values. However, we have tools that can help to pare down the number of decisions to the vital ones. Menzies' treatment learning [2, 10] can determine which decisions are significant, and which are nearly irrelevant. Decision clustering and visualization [3] can help managers reduce the search space of viable solutions to a set that, in addition to meeting cost and effectiveness objectives, meet their less measurable criteria.

Thus, we propose that the modeling of uncertainty be applied after the other tools (treatment learning, decision clustering, etc.) have isolated the most important decision variables. We would then ask engineers and domain experts to give an estimate of the uncertainty in their previous single number estimates using the visual tool described in section 3.2. That is, we model their uncertainty with a probability distribution. For example, if an expert is reasonable sure about an estimated probability of 0.2, he would use the tool to model this with a normal distribution with a fairly small standard deviation. Thus, the expert would have to give three pieces of information about the probability distribution, e.g., normal, the mean, and the standard deviation. (Of course, the visualization of the probability distribution allows the engineer to specify these parameters with implicitly given numerical values.) If he was less sure, then he might give a larger estimate of the standard deviation.

In another situation, an expert might report that the estimate of 0.2 that she previously reported is the most likely value, but that if it is wrong, it is more likely to be an over-estimate. That is, the probability is skewed to the left. As demonstrated in section 3.2, our uncertainty visualization tool allows this type of manipulation.

5. Related Work

In Section 3, we described some of work related to the concepts discussed there. Here we mention some additional work related to the issue of combining probability distributions. Resolving differences among distributions so as to combine them has long been a topic of research in statistics [5] and has recently captured the interest of artificial intelligence researchers as well [9, 12]. Although the LinOP operator we use is perhaps the most used operator due to its simplicity, many other operators have been proposed including LogOp (similar to LinOp but based on a product of logarithms instead), Bayesian updating, and iterative elicitation methods. Each method has tradeoffs – indeed, the literature is full of impossibility theorems. However, recent advances have been made by taking advantage of information available about the experts (e.g., their expertise or how they acquired their beliefs) or of structure inherent in their beliefs.

Researchers have tackled this same problem in other domains including the fields of medicine (e.g., the highly successful Pathfinder system [6]), nuclear power [7], and fire management [8]. The approach of some of these efforts is to ask human experts to rate the uncertainty in their opinions. This training of experts in “the practice of expressing knowledge and beliefs as probability distributions” mentioned above is one of the steps in the NUREG-1150 methodology described in Hora and Iman [7].

6. Future directions

For this to be an effective addition to the DDP process, users need a short introduction to probability distributions to give them an understanding of the typical properties of each. Some other previous efforts have found such training to be possible and effective. The current prototype is a separate Java application. To fully integrate it with the DDP tool, we need to add components to the DDP interface to support selection of probability.

One question that remains to be addressed is this: Is there evidence that the added information about uncertainty makes a difference in the process. We can make an intuitive argument that it does. Suppose that experts agree that a large group of important decisions each have an uncertainty probability that is skewed uniformly in one direction. Intuition would infer that the combinations of these values might conclude that the most likely solution is considerably shifted because of this skewed uncertainty. Although this intuitive argument may be more or less compelling, a more precise study of this would have to be done to determine the real value of this extra information.

7. Conclusions

Requirements engineering and early-phase design of complex systems are inherently uncertain. Engineers and domain experts use their training and experience to produce the best judgments possible in light of this uncertainty. Typically, we are forced to endure this uncertainty early with hope of making more accurate judgments in later phase of the project. Here we treat this uncertainty as an opportunity by collecting meta-knowledge about uncertainty from engineers and experts and mining this information. This gives project managers insight into potential strengths and possible areas of weakness in the requirements or design. Since this meta-information about uncertainty is captured in early phases of a project, it can be leveraged to significantly improve the developing system.

8. Acknowledgements

The authors would acknowledge the contribution of Steve Conford as the originator for the DDP process and for his on-going participation as a source of requirements for this tool. We also acknowledge the senior project team from Miami University's Department of Computer Science and Systems Analysis for building the probability visualization and manipulation tool. This team consisted of Gianni Carrero, Zach Hendershot, Lucas Lee, Ron McClain, Chris Sabatini, and Kyle Wells.

The research described in this paper was carried out at Miami University and at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

9. References

[1] Druzdzel, M. J. and L. C. van der Gaag, “Elicitation of Probabilities for Belief Networks: Combining Qualitative and Quantitative Information”, *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence (UAI95)*, Montreal, Quebec, Canada, 1995.

[2] Feather, M.S. and T. Menzies. Converging on the optimal attainment of requirements. In *IEEE Joint Conference On Requirements Engineering ICRE’02 and RE’02*, 9-13th September, University of Essen, Germany, 2002.

[3] Feather, M.S., Kiper, J.D. and Kalafat, S. “Combining Heuristic Search, Visualization and Data Mining for Exploration of System Design Spaces”, *INCOSE 2004 - 14th Annual International Symposium Proceedings*; Toulouse, France, Jun 20-24 2004.

[4] Feather, M.S., S.L. Cornford, K.A. Hicks & K.R. Johnson, “Applications of tool support for risk-informed requirements reasoning,” Computer Systems Science and Engineering (CRL Publishing Ltd) 5-17, January 2005.

[5] Genest, C. and J. V. Zidek, “Combining Probability Distributions: A Critique and an Annotated Bibliography”, Statistical Science, 1(1): 114-148, 1986.

[6] Heckerman, D., E. Horvitz, and B. Nathwani, “Toward Normative Expert Systems: Part I. The Pathfinder Project”, *Methods of Information in Medicine*, volume 31, pages 90-105, 1992.

[7] Hora, S. C. and R. L. Inman, “Expert Opinion in Risk Analysis: the NUREG-1150 Methodology”, *Nuclear Science and Engineering*, 102, 323-331, 1989.

[8] Johansson, Henrik, “Decision Making in Fire Risk Management”, Report 1022, Department of Fire Safety Engineering, Lund University, Sweden.

[9] Maynard-Reid II, P. and U. Chajewska, “Aggregating Learned Probabilistic Beliefs”, *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence (UAI01)*, pages 354-361, 2001.

[10] Menzies, T. and Y. Hu Data Mining for Very Busy People;; IEEE Computer; 2003; November;

[11] Morgan, M. G. and M. Henrion, *Uncertainty, A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis,* Cambridge University Press, Cambridge, 1990.

[12] Pennock, D. M. and M. P. Wellman, “Graphical Representations of Consensus Belief” *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence (UAI99)*, pages 531-540, 1999.

[13] Russell, S. J. and Peter Norvig. *Artificial Intelligence: A Modern Approach*, second Edition, Prentice Hall, New Jersey, 2003.

[14] van der Gaag, L. C., S. Renooij, C. L. M. Witteman, B. M. P. Aleman, and B. G. Taal, “Probabilities for a Probabilistic Network: A Case Study in Oesophageal Cancer”, *Artificial Intelligence in Medicine*, 25(2), 123-148, 2002.