Chapter 13
Formal Reliability Analysis of Engineering Systems

Naeem Abbasi
Concordia University, Canada

Osman Hasan
Concordia University, Canada

Sofiène Tahar
Concordia University, Canada

ABSTRACT

Reliability analysis of engineering systems has traditionally been done using computationally expensive computer simulations that cannot attain 100% accuracy due to their inherent limitations. The authors conduct a formal reliability analysis using higher-order-logic theorem proving, which is known to be sound, accurate, and exhaustive. For this purpose, they present the higher-order-logic formalization of independent multiple continuous random variables, their verified probabilistic properties, and generalized relations for commonly encountered reliability structures in engineering systems. To illustrate the usefulness of the approach, the authors present the formal reliability analysis of a single stage transmission of an automobile.

1. INTRODUCTION

The reliability of an engineering system is very important as an unreliable system usually translates to loss of both money and time and a considerable amount of inconvenience. Such reliability analysis is conducted using probabilistic techniques while considering the individual reliabilities of sub-components of the given engineering systems. This analysis usually involves building a model of the given engineering system using random variables and various continuous physical parameters. Computer simulations have traditionally been used for the reliability analysis of engineering systems. Computer simulations are automatic and thus user friendly and can be used to analyze analytically complex systems including the ones that cannot be modeled in a closed mathematical form. However, they cannot guarantee 100% accurate results, because 1)

DOI: 10.4018/978-1-4666-4789-3.ch013

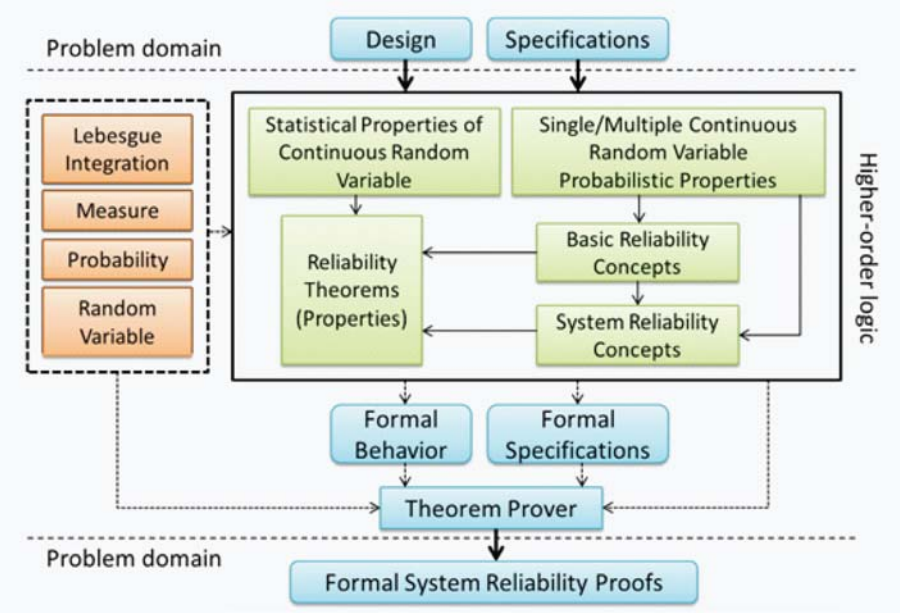
infinite precision real numbers, corresponding to the physical parameters of the system, cannot be precisely modelled in computer memory, 2) due to the enormous size of the present-age engineering systems, e.g., a modern power plant is composed of over a million components, exhaustive testing of all possible input scenarios is not possible due to limited computational resources, and 3) random variables are usually approximated using pseudo random number generators that are not truly random. The accuracy of reliability analysis of engineering systems has become a dire need these days due to their extensive usage in safety-critical applications where an incorrect reliability estimate may lead to disastrous situations including the loss of innocent lives.

Formal methods based techniques such as probabilistic model checking (Baier, Haverkort, Hermanns, & Katoen, 2003) and probabilistic theorem proving (Hasan, 2008) can alleviate some of the inaccuracy limitations of computer simulations. Model checking based techniques can handle finite sized systems or finite models of infinite systems. Some basic probabilistic and statistical reliability properties of an engineering system can be verified using this technique. However, such results cannot be considered truly formal as the decision procedures used in this process depend on numerical computations.

Theorem proving based techniques, on the other hand, do not suffer from these limitations, however, they lack the foundational formalization required for the reliability analysis of engineering systems. This includes formal reasoning support for multiple independent continuous random variables and relations that describe the reliability of engineering systems in terms of their sub-components.

This paper is targeted towards developing these foundations to facilitate formal reliability analysis using theorem proving. The proposed reliability analysis framework is shown in Figure 1. The reliability modeling and analysis process begins with the construction of a formal model of the system and its environment. The functional and reliability requirements of the system are then formally stated. The proposed reliability analysis framework then facilitates verification, computa-

Figure 1. Formal reliability analysis framework



tion, reasoning, and documentation of the reliability proofs in the sound environment of the HOL theorem prover. Finally, the formal functional and reliability analysis results are unformalized and interpreted and stated in an appropriate language in the problem domain.

The two main contributions of this paper are that: 1) it presents the formalization of multiple independent continuous random variables, and the verification of the standard cumulative distribution function properties of multiple continuous random variables in the sound core of the HOL theorem prover; 2) it presents the formalization of various commonly used reliability structures such as series, parallel, series-parallel and parallel-series in higher-order logic. These contributions play a vital role in conducting formal reliability analysis as the multiple continuous random variables can be used to model the randomness associated with each sub component of an engineering system while the reliability structure related formalization can be utilized to construct formal models of the given system using its sub-components as modules and reasoning about its associated properties.

To illustrate the usefulness of our work, we present the formal reliability analysis of a single stage transmission of an automobile. In our analysis, we utilize formalized multiple independent random variables with different distributions and verified reliability relations. The analysis is done mechanically and interactively in the sound environment of the HOL theorem prover. Such analysis, until now, was only possible using inaccurate simulation based techniques.

One of the earliest examples of detailed reliability studies in engineering systems dates back to 1938 (Dean, 1938). In this study, factors for the improvement of service reliability for electrical power systems were considered. In the field of electronics, the concepts related to reliability were initially introduced after the second world war to improve the performance of communication and navigational systems (Myers, & Ball, 1964).

In order to predict reliability, one must model a system and its constituent components in a way that captures failure mechanisms. For example, in the case of electronic systems, a method called the part failure method has been shown to be very accurate (US Department of Defense, 1991). This method has been extensively used by military engineers to predict useful lifetimes of systems and to develop highly reliable systems and equipments. This method is based on calculating failure rates of individual components of the system and then using appropriate formulas to determine the reliability of the whole system. Standards such as MIL-HDBK-2173 (US Department of Defense, 1998), FIDES (FIDES, 2012), and IEEE-1332 (Institute of Electrical & Engineers, 1998) are some of the examples which specify adequate performance requirements and environmental conditions for reliability modeling, analysis, and risk assessment.

Many formal methods based techniques have been extended to analyze reliability of systems during the last two decades. Many expressive formalisms such as stochastic petri nets (Labadi, Saggadi, & Amodeo, 2009) and process algebras (Ciocchetta & Hillston, 2009) along with various probabilistic (Kwiatkowska, Norman, Segala, & Sproston, 2002) and stochastic temporal logics (Baier et al., 2003), and compositional and guarded command notations (Hurd & Morgan, 2005) have been used in modeling, specification and analysis of complex engineering (Hasan & Tahar, 2008) and applied science problems (Barnat, Brim, & Safranek, 2010). They were either not designed to deal with reliability analysis problems or lack the capability to handle reliability problems due to lack of infrastructure. Probabilistic model checking can be used to analyze reliability; however, it does not support the verification of statistical properties (moments and variance) of the commonly used lifetime distributions (Baier et al., 2003; Rutten, Kwaiatkowska, Normal, & Parker, 2004). The proposed approach on the other hand is

capable of handling these and other probabilistic and statistical properties.

The accuracy of reliability analysis depends on both the field data gathering and the methods and tools used for analysis. In this paper, we do not address the problem of field data gathering. Our focus is on the higher-order-logic formalization of fundamental concepts of the reliability theory. Until recently it was only possible to reason about reliability problems that involved discrete random variables in a theorem proving environment (Hasan, 2008). Hurd (Hurd, 2002) formalized a probability theory along with discrete random variables in the HOL theorem prover (Gordon & Melham, 1993).

Building upon Hurd's work (Hurd, 2002), Hasan (Hasan, 2008) formalized statistical properties of single and multiple discrete random variables. Hasan (Hasan, 2008) also formalized a class of continuous random variables for which the inverse CDF functions can be expressed in a closed form. Hasan et al. (Hasan, Tahar, & Abbasi, 2010) presented higher-order-logic formalizations of some core reliability theory concepts and successfully formalized and verified the conditions for consistent repairability for reconfigurable memory arrays in the presence of stuck-at and coupling faults. However, all these existing works do not support reasoning about multiple continuous random variables and reliability structures, which is the main scope of the present paper.

The rest of the chapter is organized as follows. The formalization of multiple continuous random variables is described in Section 2. Section 3 presents the formalization of the relations for various reliability structures. Section 4 presents the reliability analysis of an automotive transmission as an illustrative example and Section 5 concludes the paper.

2. PROBABILITY DISTRIBUTION PROPERTIES OF MULTIPLE RANDOM VARIABLES

Hurd (Hurd, 2002) formalized discrete random variables as independent probabilistic algorithms in HOL. Hasan (Hasan, 2008) defined a standard uniform continuous random variable as a probabilistic algorithm with a standard uniform probability mass function utilizing a very large number of random bits (Hasan, 2008). Using this approach, it is possible to model multiple discrete random variables and a maximum of a single continuous random variable as this method exhausts the complete sequence of random bits in the standard continuous random variable. We build on these foundations; we, first, split the random Boolean sequence into a number of disjoint random Boolean sequences, then using Hasan's formalization of continuous random variables, formalize multiple continuous random variables. In our formalization, each random variable receives a disjoint segment of the random Boolean sequence, which ensures that the resulting random variables will be independent.

In the rest of this section, we describe the formalization of multiple random variables as lists of random variables. We verify their CDF properties. Moreover, we formalize the notion of independence of multiple random variables.

2.1 Formal Specification of CDF of Lists of Random Variables

In order to formally specify the CDF of a list of random variables in higher-order logic, we first define two list functions. They are rv_val and rv_lf. The higher-order logic recursive definitions of the two functions rv_val and rv_lf are as follows:

Definition 1: Random Variable Value Function

$\vdash \forall s.\ rv\_val\ [] \ s = [] \wedge$

$\forall\ h\ X\ s.\ rv\_val\ (h::\ X)\ s = h\ s::\ rv\_val\ X\ s$

Definition 2: Random Variable Logical Formula Function

$\vdash (rv\_lf\ [] \ [] = T) \wedge$

$(rv\_lf\ (h1::\ t1)\ (h2::\ t2) = h1 \leq h2 \wedge rv\_lf\ t1\ t2)$

The function rv_val takes a list of random variables, X, and the random Boolean sequence, s, and returns a list of real values. The function rv_lf takes two real lists as input and returns a Boolean expression consisting of conjunction of several terms formed from the corresponding elements of the two input lists.

Each inequality in this Boolean expression is of the form ((EL X i) s ≤ (EL x i)). The function EL takes a list and a natural number as input arguments (for example, EL X i) and returns the corresponding element of the list as output (in this case it would return the ith element of the list X).

Now using Definitions 1 and 2, we formally specify the joint CDF of a list of random variables in Definition 3.

Definition 3: Joint CDF of a List of Random Variables

$\vdash \forall X\ x.\ mcrv\_cdf\ X\ x = prob\ bern\ \{s \mid rv\_lf\ (rv\_val\ X\ s)\ x\}$

where X is a list of random variables of type (((num→bool)→real) list), and x is a list of real numbers of type (real list).

2.2 Formal Verification of CDF Properties of Lists of CRVs

Using the formal specification of the CDF function for a list of random variables, we formally verify the classical properties of the CDF of a list of random variables, given in Table 1.

We verify these properties under the assumption that the set {s | X s x}, where X represents a list of random variables under consideration, is measurable for all values of the list. The formal proof of these properties was mainly based on reasoning from probability and set theories in HOL and real analysis. The details can be found in (Abbasi, 2012). The formal proofs of these properties not only confirm our formal specifications of the CDF but also can be used to reason about probability distribution properties of multiple random variables.

2.3 Independent Random Variables

The notion of independence for a list of random variables $X = [X_0; X_1; X_2; \ldots ; X_{N-1}]$ is defined as:

$$P\left(\bigcap_{i=N-1} X_0 \leq x_0 \wedge X_1 \leq x_1 \wedge \ldots \wedge X_{N-1} \leq x_{N-1}\right) = \prod_{i=0} P\left(X_i \leq x_i\right)$$

where $x = [x_0; x_1; x_2; \ldots ; x_{N-1}]$ is a list of real numbers. The subscript in the above equation represents the index of the random variable in the list. N represents the length of the list of random variables X. In order to formalize a list of independent continuous random variables, we first define the notion of a list of disjoint random Boolean sequences using higher-order logic functions s_arb and s_split in Definitions 4 and 5, respectively.

Table 1. Formally verified joint CDF properties in HOL

Description	Joint CDF Property
CDF Bounded	$0 \le F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) \le 1,\ \forall x_1,\ldots,x_n \in R$
CDF Monotonic-Non decreasing	$F_{X_1,\ldots,X_n}(x_1,\ldots,a,\ldots,x_n) \le F_{X_1,\ldots,X_n}(x_1,\ldots,b,\ldots,x_n), a \le b;$ $\forall x_1,\ldots,x_n,a,b \in R$
Marginal CDF	$\lim_{x_i\to\infty} F_{X_1,\ldots,X_{i-1},X_i,X_{i+1},\ldots,X_n}(x_1,\ldots,x_{i-1},x_i,x_{i+1},\ldots,x_n) \le$ $F_{X_1,\ldots,X_{i-1},X_{i+1},\ldots,X_n}(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)$ $\forall x_1,\ldots,x_n \in R$
CDF at positive Infinity	$\lim_{x_1\to\infty},\ldots,\lim_{x_n\to\infty} F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) \le F_{X_1,\ldots,X_n}(\infty,\ldots,\infty) = 1,$ $\forall x_1,\ldots,x_n \in R$
CDF at Negative Infinity	$\lim_{x_i\to-\infty} F_{X_1,\ldots,X_i,X_n}(x_1,\ldots,x_i,\ldots,x_n) \le F_{X_1,\ldots,X_n}(x_1,\ldots,-\infty,\ldots,x_n) = 0,$ $\forall x_1,\ldots,x_n \in R$

Definition 4: Boolean Sequence Split Function

⊢ (∀ s M i. s_arb M i 0 = s i) ∧

∀ s n M i. s_arb s M i (SUC n) = s (M*SUC n + i)

The function s_arb takes three arguments. The first argument is a Boolean sequence s. The second and third arguments are natural numbers M and i. The function s_arb can split the input Boolean sequence s into M disjoint Boolean sequences. The third argument i is used to pick every ith element from the input infinite Boolean sequence and the function s_arb returns that Boolean sequence as output. This way we can provide each random variable in the list of random variables with a different infinite random Boolean sequence. This fact also guarantees independence of random variables in the list (Williams, 1991).

Definition 5: List of Disjoint Boolean Sequences

⊢ ∀ M s. s_split 0 M s = [(λx. s_arb s x M) 0] ∧

∀ N M s. s_split (SUC N) M s = (λx. s_arb s x M) (SUC N):: s_split N M s

The function s_split takes a Boolean sequence as input and returns a list consisting of M+1 disjoint Boolean sequences. For example, s_split 2 2 s would return a list of three disjoint Boolean sequences given by [s_arb s 2 2; s_arb s 1 2; s_arb s 0 2].

In order to define the notion of independence of a list of random variables, we first define a list function that we call rv_val indep. This function merges two lists element by element and generates a list. The first list argument of this function is a list of random variables of type ((num->bool)->real) list) and the second list argument is a list

consisting of random boolean sequences of type ((num->bool) list). The function merges the two lists element by element and returns a list of real independent random variables.

Definition 6: List Function rv_val_indep

⊢ (rv_val_indep [] [] = []) ∧

(rv_val_indep (h1::t1) (h2::t2) = h1 h2::rv_val_indep t1 t2)

Finally, the HOL formalization of the notion of independence is given in Definition 7.

Definition 7: Independent Random Variable List

⊢ ∀ X x. indep_rv_list X x =

(prob bern {s | rv_lf (rv_val_indep X (s_split (PRE (LENGTH X)) (LENGTH X) s)) x} =

prod1 (0, LENGTH X)

(λi. prob bern {s | EL i (rv_val_indep X (s_split (PRE (LENGTH X)) (LENGTH X) s) ≤ EL i x}))
∧

{s | rv_lf (rv_val_indep X (s_split (PRE (LENGTH X)) (LENGTH X) s)) x} IN events bern ∧

∀ i. {s | EL i (rv_val_indep X (s_split (PRE (LENGTH X)) (LENGTH X) s)) <= EL i x} IN events bern

where X and x are of types (((num -> bool) -> real) list) and (real list), respectively. prod1 is a product of a sequence function and represents the big pi operator ($\prod$). The function s split splits the random Boolean sequence s and returns a list of disjoint random boolean sequences. PRE is a function of type (num->num) and is defined as: ∀m. PRE m = (if m = 0 then 0 else @n. m = SUC n), where @ is the hilbert's choice operator. The list function EL takes two arguments, a natural number i and a list and returns the ith element of the list.

The second and the third logical terms in Definition 7 state that the respective events are measurable in the probability space. The higher-order logic formalization presented in this section facilitates the verification of expressions for various reliability structures and the reliability analysis of automotive transmission described in Sections 3 and 4, respectively.

3. RELIABILITY ANALYSIS OF COMPLEX SYSTEMS

The reliability structure of an engineering systems is determined by its functional and non-functional requirements. We have verified relations for series, parallel, series-parallel, and parallel-series reliability structures in HOL.

In the following, we briefly describe these results. Formalization and detailed proof descriptions of these results can be found in (Abbasi, 2012).

3.1 Series Connected Systems

In a series connected system (Figure 2) with N components, the system functions as long as all its components are functioning. As soon as any of the system component fails, the system fails as well. The reliability of such a system is mathematically described as:

$$R_S\left(t\right) = \prod_{i=1}^{N} R_i\left(t\right) \tag{1}$$

In Definition 8, we define a series system structure that consists of N components. These components are modeled using a list of random variables of type ((num->bool)->real) list).

Figure 2. Reliability structure of series connected systems

Definition 8: N Series System Structure Function

⊢ ∀ X x s t N. N_series_system X x s t N =

list_conj_gt (rv_val_indep X (s_split (PRE N) N s)) (FILL_LIST_R x t)

The function rv_val_indep takes two lists as arguments and constructs a single list. The first argument of this function is the list of random variables X. The second argument is another list generated by the function s split. This generated list consists of disjoint segments of the random boolean sequence s.

The function list_conj_gt constructs a conjunction of logical terms, each of which is a greater than inequality and consists of corresponding terms from its two list arguments. Both list arguments of list_conj_gt are real lists. The second argument of list_conj_gt is constructed by the list function LIST_FILL_R, which fills the list x with a real value t.

We define the survival function (the probability of failure at a certain time) of a series connected system with N components in Definition 9.

Definition 9: N Series System Survival Function

⊢ ∀ X x N. N_series_survival_function X x N =

(λt. prob bern {s| list_conj_gt (rv_val_indep X (s_split (PRE N) N s)) (FILL_LIST_ R x t)})

In Theorem 1, we verify the N series system reliability property (Equation 1).

Theorem 1: N Series System Reliability

⊢ ∀ X x t N. indep_rv_list X (FILL_LIST_R x t) ==>

(N_series_survival_function X x N t =

(λt. prod1 (0,N) (λi. prob bern {s| t < EL i (rv_val_indep X (s_split (PRE N) N s))})) t)

The proof of this theorem follows from the definitions of the series survival function and the independence of a list of random variables and involves reasoning from real, measure, probability, and set theories in the HOL theorem prover.

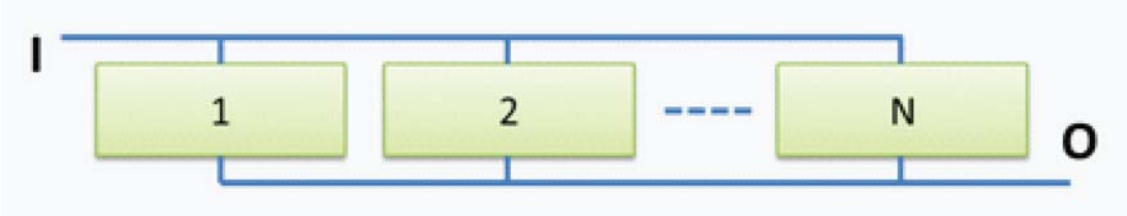
3.2 Parallel Connected Systems

If N components of a system are connected in parallel (as shown in Figure 3), the system will function properly as long as at least one of the components is functioning. Such a system stops functioning when all the system components fail. The reliability of such a system is mathematically described as:

$$R_P\left(t\right) = 1 - \prod_{i=1}^{N}\left(1 - R_i\left(t\right)\right) \tag{2}$$

In Definition 10 the parallel system structure function is formalized.

Figure 3. Reliability structure of parallel connected systems



Definition 10: N Parallel System Structure Function

⊢ ∀ X x s t N. N_parallel_system X x s t N =

list_disj_gt (rv_val_indep X (s_split (PRE N) N s)) (FILL_LIST_R x t)

In this definition, rv_val_indep constructs a list of independent random variables as described in the case of series connected systems. The function list_disj_gt constructs a disjunction of logical terms, each of which is a greater than inequality and consists of corresponding terms from its two list arguments.

In Definition 11, we define a parallel connected system with N elements.

Definition 11: N Parallel System Survival Function

⊢ ∀ X x N. N_parallel_survival_function X x N =

(λt. prob bern { s| list_disj_gt (rv_val_indep X (s_split (PRE N) N s)) (FILL_LIST_R x t)})

The first argument is a list of random variables of type ((num->bool)->real) list). The function list_disj_gt takes two lists as arguments and creates a logical expression that consists of disjunction of greater than inequalities involving the corresponding terms of the two input lists. The first list (rv_val_indep X (s split (PRE (LENGTH X)) (LENGTH X) s)) argument of list_disj_gt is a list of real random variables constructed in a similar manner as explained in Definition 8. The function FILL_LIST_R returns the list x after filling it with the variable t. The third argument N represents the number of components in the parallel reliability structure.

The reliability expression for a N parallel connected system (Equation 2) is verified in Theorem 2.

Theorem 2: N Parallel System Reliability

⊢ ∀ t X x. indep_rv_list X (FILL_LIST_R x t) ==>

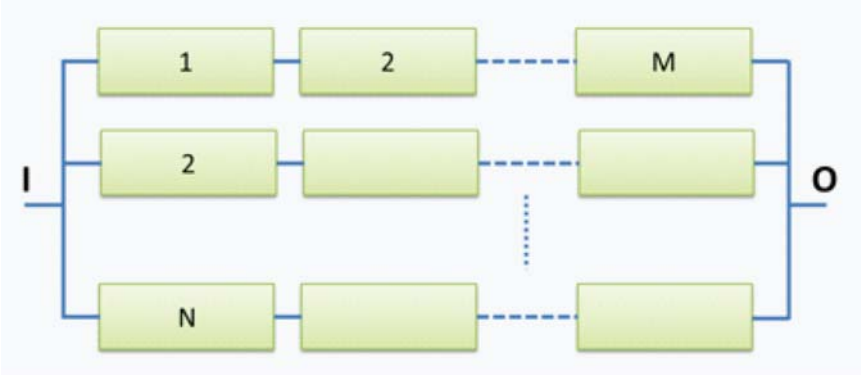
(N_parallel_survival_function X x N t =

1- prod1 (0,N) (λi. 1 – prob bern {s|t < EL i (rv_val_indep X (s_split (PRE N) N s))}))

3.3 Series Parallel Connected Systems

If a system consists of N components in parallel, where each of such parallel connected component has M components connected in series then such a system is called a series-parallel system. One such example is shown in Figure 4 and the reliability of such a system is given by:

Figure 4. Reliability structure of parallel-series connected systems



$$R_{SP}(t) = 1 - \prod_{i=1}^{N}\left(1 - \prod_{j=1}^{M} R_{ij}(t)\right) \quad (3)$$

where R_{ij} is the reliability of the j-th component in the i-th branch of the system. Such a system configuration is typically used to enhance the reliability at the system level.

An N x M series parallel structure has N components connected in parallel such that each of these components has M sub components connected in series. Definition 12 shows how such a system structure function can be formally specified.

Definition 12: NxM Series Parallel System Structure Function

⊢ ∀ N M X s. NxM_series_parallel_system N M X s =

LIST_SPLIT (FILL_LIST_NM M N) (rv_val_indep (FLAT X) (s_split (PRE (N*M)) (N*M) s))

Definition 13 formally describes the series parallel survival function of a N x M system.

Definition 13: NxM Series Parallel System Survival Function

⊢ ∀ X N M. NxM_series_parallel_survival_function X N M =

(λt. prob bern {s| series_parallel_system (LIST_SPLIT (FILL_LIST_NM M N) (rv_val_indep (FLAT X) (s_split (PRE (LENGTH (FLAT X))) (LENGTH (FLAT X)) s))) t s })

The reliability expression for a N x M parallel series system is verified in Theorem 3.

Theorem 3: Series Parallel System Reliability

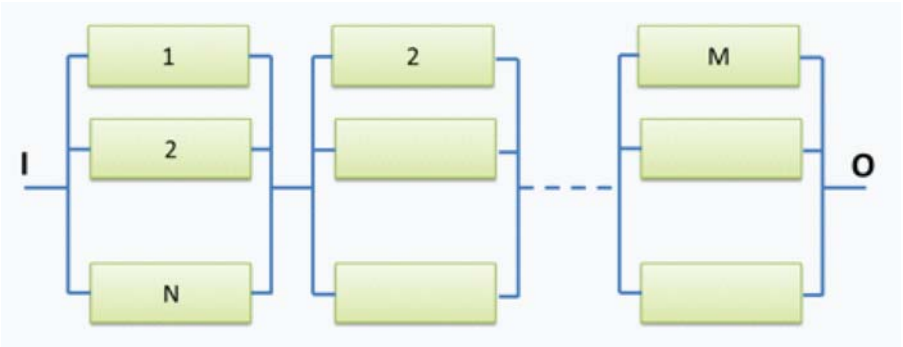
⊢ ∀ t x. (∀ X x t. indep_rv_list (FLAT X) (FILL_LIST_R x t)) ==>

NxM_series_parallel_survival_function L N M t = 1 – prod1 (0,N) (λi. 1 – prod1 (0,M) (λj. prob bern {s|t < ELEL i j (LIST_SPLIT (FILL_LIST_NM M N) (rv_val_indep (FLAT X) (s_split (PRE (N*M)) (N*M) s)))}))

3.4 Parallel Series Connected Systems

If a system consists of M components connected in series such that each of the series component consists of N sub components connected in parallel. Such a system is called a parallel-series system and is shown in Figure 5. The reliability of such a system is given by:

Figure 5. Reliability structure of series-parallel connected systems



$$R_{PS}(t) = \prod_{j=1}^{M}\left(1 - \prod_{i=1}^{N}\left(1 - R_{ij}(t)\right)\right) \quad (4)$$

where Rij is the reliability of the ij th component of the system.

Parallel-series connections can be considered as introducing component level redundancy. It can be shown mathematically that such a redundancy improves the reliability of the system more than the system level redundancy [9].

An N x M parallel series structure has been formally described in Definition 14.

Definition 14: NxM Parallel Series System Structure Function

⊢ ∀ N M X s. NxM_parallel_series_system N M X s = LIST_SPLIT (FILL_LIST_NM M N) (rv_val_indep (FLAT X) (s_split (PRE (N*M)) (N*M) s))

Definition 15 formally describes the parallel series survival function of a N x M system.

Definition 15: NxM Parallel Series System Survival Function

⊢ ∀ X N M. NxM_parallel_series_survival_function X N M = (λt. prob bern {s | parallel_series_system (LIST_SPLIT (FILL_LIST_NM M N) (rv_val_indep (FLAT X) (s_split (PRE (LENGTH (FLAT X))) (LENGTH (FLAT X)) s))) t s})

The reliability expression for a N x M parallel series system is verified in Theorem 4.

Theorem 4: NxM Parallel Series System Reliability

⊢ ∀ t x a. (∀ X x t. indep_rv_list (FLAT X) (FILL_LIST_R x t)) ==>

NxM_parallel_series_survival_function X N M t =

prod1 (0,M) (λj. 1-prod1 (0,N) (λi. 1 –prob bern {s | t < ELEL i j (LIST_SPLIT (FILL_LIST_NM M N) (rv_val_indep (FLAT X) (s_split (PRE (N*M)) (N*M) s)))}))

The proofs of theorems 1 through 4 are primarily based on reasoning from the probability, set, measure, Boolean, and real theories in HOL. Such analysis has traditionally been done using simulation based techniques and suffers from accuracy problems and their inability to model true independent random behavior. The higher-order logic formalization presented in this section enables analysis of reliability behavior of many simple and complex engineering systems as will be demonstrated in the next section.

4. AUTOMOBILE TRANSMISSION

The automobile transmission transfers mechanical power from the input shaft to the output shaft using a pair of gears. The power is transmitted from a larger gear on the input shaft to a smaller gear on the output shaft. The reliability of the automobile transmission is very important and is usually determined using three main steps. The first step identifies the reliability relevant components and determines their reliability. The second step determines the reliability structure of the system. Finally, based on the reliability structure of the system, we calculate the overall reliability of the system.

ABC (Naunheimer, Bertsche, Ryborz, Novak, & Kuchle, 2010) and FMEA (Langford, 2006) analysis are qualitative analysis methods commonly used to separate reliability relevant components from reliability neutral components. In this application, the ABC analysis suggests that out of the 27 parts in the automotive transmission only twelve are reliability relevant components of the system (Bertsche & Ingenieure, 2008). These components include the shafts, the bearings, the gears, the fitting keys and the seals. Moreover, the mechanical transmission has a pure serial reliability structure as shown in Figure 6. Therefore, the system reliability R_{TRAN} is given by the product of the reliability of the individual components.

$$R_{TRAN} = R_{IS}.R_{OS}.R_{G1B}.R_{G2B}.R_{RB1}.R_{RB2}.R_{FK}.R_{G12P}.R_{RB3}.R_{RB4}.R_{SS1}.R_{SS2} \quad (5)$$

4.1 Formal Reliability Analysis of the Automotive Transmission

In this analysis, we assume that Weibull random variable (Bertsche & Ingenieure, 2008) is used to model the reliability behavior of various components of the automotive transmission.

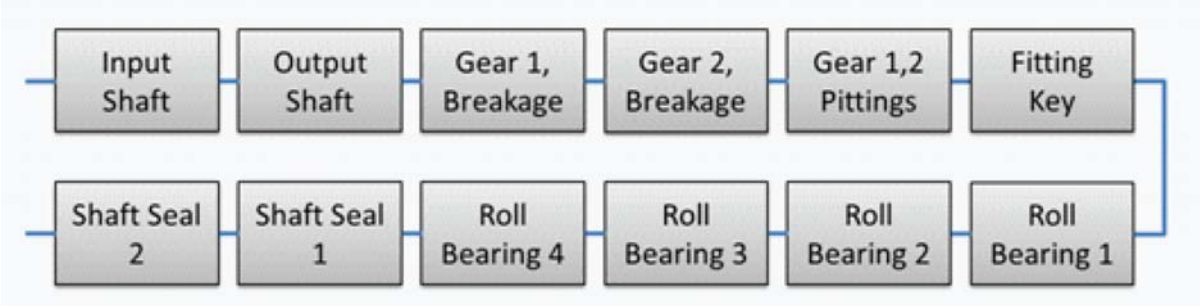
The Weibull distribution is commonly used in such analysis. We first construct a list of N independent Weibull random variables to model the automotive transmission as given in Definition 16.

Definition 16: Automotive Transmission Reliability Model

⊢ ∀ a b N s. auto_rv_list a b N s = rv_val_indep (WB_RV_LIST a b) (s_split (PRE N) N s)

In Definition 16, a and b are lists that contain shape and scale parameters of the Weibull random variables in the WB_RV_LIST. x is a real list, N represents the number of components in the series reliability structure and t is a positive real value. Each element of this list represents the lifetime of a component of the transmission.

Figure 6. Reliability structure of the automobile transmission



Definitions 17 shows our formalization of list of random variables with weibull distribution.

Definition 17: List of Weibull Random Variables

⊢ (WB_RV_LIST [] [] = []) ∧

(WB_RV_LIST (ah::at) (bh::bt) = [(λa b s. weibull_rv a b s) ah bh] ++ WB_RV_LIST at bt)

The function WB_RV_LIST takes two real lists as arguments and returns a list of independent Weibull random variables. The two real lists contain the corresponding scale and shape parameters of the Weibull distributions of the random variables in the returned list.

Definition 18 formally describes the reliability model of the automotive transmission. The series reliability structure is modeled using the series reliability structure definition (N series survival function).

Definition 18: Automotive Transmission Reliability Model

⊢ ∀ a b x N t. auto_trans_rel_model_N a b x N t = N_series_survival_function (WB_RV_LIST a b) x N t

4.2 Lifetime Reliability Analysis in HOL

The survival function $S_T(t)$ is defined as:

$$S_T\left(t\right) = 1 - F_T\left(t\right) \tag{6}$$

where $F_T(t)$ is the cumulative distribution function of the random variable T .

The survival function represents the probability that a component is functioning at one particular time t and is formalized in HOL as follows:

Definition 19: Survival Function

⊢ ∀ rv. survival_function rv = (λt. 1 – CDF rv t)

where CDF is the cumulative distribution function of random variable rv. Both survival function and CDF in HOL are of type (((num → bool) → real) → real → real).

Theorem 5: Automotive Transmission System Reliability

⊢ ∀ a b t. (∀ a b t. indep_rv_list (WB_RV_LIST a b) (FILL_LIST_R x t)) ∧

(∀ i. 0 < (EL i a)) ∧ (∀ i. 0 < (EL i b)) ∧ (0 ≤ t) ∧ (LENGTH (WB_RV_LIST a b) = 12) ==>

(auto_trans_rel_model_N a b x 12 t = prod1 (0,12) (λi. survival_function (EL i (WB_RV_LIST a b)) t))

Theorem 5 formally states that for an automotive transmission, consisting of 12 critical reliability relevant components, the overall system reliability is given by the product of reliability of its individual components (Equation 5), provided the components of the transmission fail independent of each other.

The proof of Theorem 5 required rewriting with Definition 17 and reasoning from Theorem 1 for the series connected system. Theorem 5 provides a formal proof of correctness of the reliability specification of an automotive transmission.

The HOL code describing our formalization of multiple continuous random variables and their probabilistic properties consists of approximately 2000 lines of code and took over 240 hours to complete. The formalization of reliability structures consists of approximately 2500 lines of code and took over 160 manhours to complete. The formalization and the verification of the automotive transmission reliability took only 300 lines

of HOL code and only 30 hours to complete. As can be seen that, the automotive transmission reliability analysis results took an order of magnitude less effort to prove than the infrastructure development work. This shows the strength of our work and that it will be very useful for engineers building on this work to attempt larger and more involved hardware and embedded system software reliability analysis problems.

The reliability expressions we presented are guaranteed to be accurate, unlike the simulation based analysis, and are generic due to the universally quantified variables. Such analysis was not possible in the HOL theorem prover earlier.

5. CONCLUSION

In this paper, we presented an approach for formal reliability analysis of engineering systems using higher-order-logic theorem proving. In this context, we also presented the formalization of multiple continuous random variables and verified their classical Cumulative Distribution Function properties. Then building on these foundations, we described the formalization and analysis of commonly used reliability structures. The proposed formalization is general and can facilitate performance and reliability analysis of problems in many domains of engineering and applied sciences. It does not have any theoretical limitations as far as the number of system components and the modeling of complexity of structure is concerned. The results presented are guaranteed to be accurate, unlike simulation based analysis, and are generic due to the universally quantified variables in the proven reliability properties. For illustration purposes, we presented the analysis of an automobile transmission system. The formal reasoning about this system was straightforward and required very little user intervention, which demonstrates the usefulness of our work.

We are currently working on several interesting and large reliability analysis applications. In one such application we are analyzing the reliability of a multiprocessor system consisting of various types of processing units, such as field programmable gate arrays, general purpose processors and memories. In many real-world applications correlated random variables are required. In order to be able to tackle the formal analysis of these applications, we plan to develop formal reasoning support for multiple correlated random variables (Snedecor & Cochran, 1989). Other related future directions of research include the formalization of availability and maintainability theories (Ebeling, 1997) in HOL.

REFERENCES

Abbasi, N. (2012). *Formal reliability analysis using theorem proving*. (PhD Thesis). Concordia University, Montreal, Canada.

Baier, C., Haverkort, B., Hermanns, H., & Katoen, J. (2003). Model checking algorithms for continuous time Markov chains. *IEEE Transactions on Software Engineering*, *29*(4), 524–541. doi:10.1109/TSE.2003.1205180

Barnat, J., Brim, L., & Safranek, D. (2010). 05). High-performance analysis of biological systems dynamics with the DiVinE model checker. *Briefings in Bioinformatics*, *11*(3), 301–312. doi:10.1093/bib/bbp074 PMID:20478855

Bertsche, B., & Ingenieure, V. D. (2008). *Reliability in automotive and mechanical engineering: Determination of component and system reliability*. Berlin: Springer.

Ciocchetta, F., & Hillston, J. (2009). Bio-PEPA: A framework for the modelling and analysis of biological systems. *Theoretical Computer Science*, *410*(33), 3065–3084. doi:10.1016/j.tcs.2009.02.037

Dean, S. (1938). Considerations involved in making system investments for improved service reliability. *EEI Bulletin*, (6), 491-496.

Ebeling, C. (1997). *An introduction to reliability and maintainability engineering*. Hoboken, NJ: McGraw Hill.

FIDES. (2012, April 30). *The FIDES methodology*. Retrieved from http://tinyurl.com/d5a2bn6

Gordon, M., & Melham, T. (1993). *Introduction to HOL: A theorem proving environment for higher-order logic*. Cambridge, UK: Cambridge University Press.

Govil, A. (1983). *Reliability engineering*. Hoboken, NJ: TATA McGraw-Hill Publishing Company.

Hasan, O. (2008). *Formal probabilistic analysis using theorem proving*. (PhD Thesis). Concordia University, Montreal, Canada.

Hasan, O., & Tahar, S. (2008). Performance analysis of ARQ protocols using a theorem prover. In *Proceedings of the International Symposium on Performance Analysis of Systems and Software* (pp. 85-94). IEEE Computer Society.

Hasan, O., Tahar, S., & Abbasi, N. (2010). Formal reliability analysis using theorem proving. *IEEE Transactions on Computers*, *59*(5), 579–592. doi:10.1109/TC.2009.165

Hurd, J. (2002). *Formal verification of probabilistic algorithms*. (PhD Thesis). University of Cambridge, Cambridge, UK.

Hurd, J., McIver, A., & Morgan, C. (2005). Probabilistic guarded commands mechanized in HOL. *Theoretical Computer Science*, *346*(1), 96–112. doi:10.1016/j.tcs.2005.08.005

Institute of Electrical and Electronics Engineers. (1998). *IEEE standard reliability program for the development and production of electronic systems and equipment*. IEEE.

Kwiatkowska, M., Norman, G., Segala, R., & Sproston, J. (2002). Automatic verification of real-time systems with discrete probability distributions. *Theoretical Computer Science*, *282*(1), 101–150. doi:10.1016/S0304-3975(01)00046-9

Labadi, K., Saggadi, S., & Amodeo, L. (2009). PSA-SPN - A parameter sensitivity analysis method using stochastic petri nets: Application to a production line system. *AIP Conference Proceedings*, *1107*(1), 263–268. doi:10.1063/1.3106483

Langford, J. (2006). *Logistics: Principles and applications*. Hoboken, NJ: SOLE Press/McGraw-Hill.

Myers, R. H., & Ball, L. W. (1964). *Reliability engineering for electronic systems*. Hoboken, NJ: J. Wiley.

Naunheimer, H., Bertsche, B., Ryborz, J., Novak, W., & Kuchle, A. (2010). *Automotive transmissions: Fundamentals, selection, design and application*. Berlin: Springer.

Rutten, J., Kwaiatkowska, M., Normal, G., & Parker, D. (2004). *Mathematical techniques for analyzing concurrent and probabilisitc systems* (Vol. 23). American Mathematical Society.

Snedecor, G. W., & Cochran, W. (1989). Statistical methods (No. v. 276). Iowa State University Press.

US Department of Defence. (1991). *Reliability prediction of electronic equipment, military handbook (MIL-HDBK-217F)*. Washington, DC: US Department of Defence.

US Department of Defense. (1998). *Reliability-centered maintenance (RCM) requirements for naval aircraft, weapon systems, and support equipment (MIL-HDBK-2173)*. Washington, DC: US Department of Defense.

Williams, D. (1991). *Probability with martingales*. Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511813658