Formalization of RBD-based Cause Consequence Analysis in HOL

Mohamed Abdelghany
Electrical and Computer Engineering
Concordia University
Montréal, QC, Canada
m_eldes@ece.concordia.ca

Sofiène Tahar
Electrical and Computer Engineering
Concordia University
Montréal, QC, Canada
tahar@ece.concordia.ca

Abstract—**Cause consequence analysis is a safety assessment technique that is traditionally used to model the causes of subsystem failures in a critical system and their potential consequences using Fault Tree and Event Tree (ET) dependability modeling techniques, combined in a graphical Cause-Consequence Diagram (CCD). In this paper, we propose a novel idea of formal CCD analysis based on Reliability Block Diagrams (RBD). Unlike Fault Trees, RBDs allow to model the success relationships of subsystem components to keep the entire subsystem reliable. To this end, we formalize in higher-order logic novel mathematical formulations of CCD functions for the RBD modeling of generic n-subsystems using HOL4. This formalization enables universal n-level CCD analysis, based on RBDs and ETs, that can determine the probabilities of multi-state safety classes. For illustration purposes, we apply our formalization on a Smart Grid system, where we determine in HOL4 all possible safety classes of accident events, and compare our results with those obtained from manual approaches and MATLAB Monte-Carlo simulation.**

Index Terms—**Cause-Consequence Diagram, Reliability Block Diagram, Event Tree, Higher-Order Logic, Theorem Proving.**

I. Introduction

Since the late 90's, various types of dependability modeling techniques have been developed to determine the safety assessment of safety-critical systems, such as smart grids [1] and automotive industry [2]. These include predominantly graph theory based approaches such as Fault Trees (FT) [3], Reliability Block Diagrams (RBD) [4] and Event Trees (ET) [5]. FTs mainly provide a graphical model for analyzing the factors causing a complete system failure upon their occurrences. On the other hand, RBDs provide a schematic structure for analyzing the success relationships of system components that keep the entire system reliable. In contrast to FTs and RBDs, ETs provide a tree model for all possible complete/partial failure and success scenarios at the system-level so that one of these possible scenarios can occur [5]. More recently, an approach has been proposed to conduct ET analysis in conjunction with FTs to identify all subsystem failure events in a critical system and their cascading dependencies on the entire system [6]. This analysis method is known as cause-consequence analysis, using a combined hierarchical structure of Cause-Consequence Diagrams (CCD) [6].

Traditionally, CCD analysis based on FTs and ETs is carried out by using paper-and-pencil approaches (e.g., [7], [8]) or computer simulation tools (e.g., [9], [10]). The major limitations of the manual analysis approach are its human-error proneness and scalability to handle large complex systems [8]. On the other hand, while simulation-based analysis approaches, such as MATLAB Monte-Carlo Simulation (MCS), can be used for CCD analysis for faster computation. They, however, lack the rigor of detailed proof steps and absolute accuracy (i.e., results approximation) due to an explosion of the test cases [9]. A more practical way to remedy the shortcomings of informal reasoning approaches of cause-consequence analysis is to use formal generic mathematical formulations that can analyze large-scale CCD graphs. Only a few works have previously considered using formal methods for cause-consequence analysis. For instance, Ortmeier et al. in [11] developed a formal framework for Deductive Cause-Consequence Analysis (DCCA) using the SMV model checker [12] to formally verify probabilistic properties for CCD analysis. However, according to the authors of [13], there is a scalability problem of showing the completeness of DCCA due to the exponential growth of the number of proof obligations with large complex CCD graphs. For that reason, to overcome the above-mentioned limitations, we endeavor to solve the scalability problem of CCDs by using theorem proving, in particular the HOL4 proof assistant [14], which provides the ability of verifying generic probabilistic expressions constructed in higher-order logic (HOL).

Prior to this work, there were three notable projects for building formal infrastructures in HOL to formally model and analyze FTs, RBDs and ETs. For instance, Ahmad [15] used the HOL4 theorem prover to formalize ordinary (static) FT and RBD structures. While, Elderhalli [16] had formalized dynamic versions of FTs and RBDs in HOL4. These formalizations have been used for the reliability analysis of several engineering systems. However, they formally analyze either a critical system static/dynamic failure or static/dynamic success only. Therefore, Abdelghany et al. in [17] developed a HOL4 theory to reason about ETs considering all failure and success events of system-level components simultaneously. They proposed a new datatype `EVENT_TREE` consisting of ET basic constructors that can build large scale ET diagrams and provides us the ability to obtain a verified system-level failure/operating consequence expression. Moreover, Abdelghany et al. in [18] proposed a formal approach for state-of-the-art CCD analysis using the above static FT and ET formalizations, which enables safety analysts to perform formal failure analysis for n-level subsystems of a complex system and obtain all possible complete/partial failure and success consequences events that can occur in HOL4.

In this paper, we provide a formalization of a novel graph theory of CCDs based on RBD and ET theories in HOL4. Unlike FT-based CCD analysis, RBDs allow to model all success relationships of n-subsystems to keep them reliable and obtain multi-state consequence safety classes, i.e., complete/partial failure and complete/partial success, that can occur in the entire critical system at the subsystem level. To the best of our knowledge, the idea of using RBD modeling in conjunction with the graph theory of CCDs has not been proposed before. We propose new mathematical formulations that can analyze scalable CCDs associated with different RBD configurations to n-subsystems. In order to check the correctness of the newly-proposed equations, we verified them within the sound environment of the HOL4 theorem prover. To this end, we formalize in HOL4 cause-consequence functions for the formal modeling of the graph theory of RBDs corresponding to generic n-subsystems. Also, our proposed formalization enables the formal probabilistic assessment of large scale n-level CCD structures based on any probabilistic distribution, which makes our work the first of its kind. To demonstrate the practical effectiveness of the proposed RBD/ET-based CCD formalization in HOL4, we conduct the formal CCD analysis of a real-world Smart Grid (SG) system, where we formally determine all possible SG safety classes of complete/partial reliability and failure consequence events that can occur at the subsystem level. To assess the accuracy of our results, we compare them with paper-and-pencil and MATLAB Monte-Carlo Simulation (MCS) approaches.

The rest of the paper is organized as follows: In Section II, we describe some preliminaries to facilitate the understanding of the rest of the paper. Section III presents the proposed formalization of CCDs based on RBDs and ETs, including the newly introduced probabilistic formulations. In Section IV, we present the formal RBD/ET-based CCD analysis of a Smart Grid system. Lastly, Section V concludes the paper.

II. Preliminaries

In this section, we briefly summarize the fundamentals of existing RBD and ET formalizations in HOL4 to facilitate the reader's understanding of the rest of the paper.

A. RBD Formalization

Reliability Block Diagram [4] (RBD) analysis is one of the commonly used safety assessment techniques for critical systems. It mainly provides a schematic diagram for analyzing the success relationships of subsystem components that keep the entire subsystem reliable. An RBD structure consists of blocks that represent the subsystem components and connectors that indicate the connections between these components. An RBD has two main types of configuration patterns *series* and *parallel*. The reliability of a subsystem when its components are connected in series configuration is considered to be reliable at time *t* only if all of the components are functioning reliably at time *t*, then the overall reliability $\mathcal{R}$ of the subsystem can be mathematically expressed as [4]:

$$\mathcal{R}_{series}(t) = Pr\left(\bigcap_{i=1}^{N} X_i(t)\right) = \prod_{i=1}^{N} \mathcal{R}_i(t) \qquad (1)$$

Similarly, the reliability of a subsystem where its components connected in parallel will continue functioning at a specific time *t* as long as at least one of its components remains functional, which can be mathematically expressed as [4]:

$$\mathcal{R}_{parallel}(t) = Pr\left(\bigcup_{i=1}^{N} X_i(t)\right) = 1 - \prod_{i=1}^{N} (1 - \mathcal{R}_i(t)) \qquad (2)$$

Ahmad et al. in [19] presented the RBD formalization by defining a new datatype `rbd`, in HOL4 as:

```
Hol_datatype rbd = series of (rbd list)   |
                   parallel of (rbd list) |
                   atomic of (event)
```

The RBD constructors `series` and `parallel` are recursive functions on `rbd`-typed lists, while the RBD constructor `atomic` operates on an `rbd`-type variable. A semantic function is then defined over the `rbd` datatype that can yield mathematically the corresponding RBD diagram as:

Definition 1:

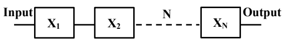
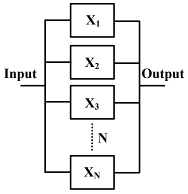
```
⊢ rbd_struct p (atomic X = X ∧
  rbd_struct p (series  (X::XN) =
  rbd_struct p X ∩ rbd_struct p (series XN) ∧
  rbd_struct p (parallel  (X::XN)) =
  rbd_struct p X ∪ rbd_struct p (parallel XN)
```

The function `rbd_struct` takes a single event `X`, identified by a basic type constructor `atomic`, and returns the given event `X`. If the function `rbd_struct` takes an arbitrary list of type `rbd`, identified by a type constructor `series`, then it performs the intersection of all elements after applying the function `rbd_struct` on each element of the given list. Similarly, if the function `rbd_struct` takes an arbitrary list of type `rbd`, identified by a type constructor `parallel`, then it returns the union of all elements after applying the function `rbd_struct` on each element of the list $\mathtt{X}_N$.

The formal verification in HOL4 for the reliability series and parallel probabilistic expressions Eq. 1 and Eq. 2, respectively, is presented in Table I [19]. These mathematical expressions (Theorems 1-2) are verified under the constraints that (a) all associated events in the given list X_N are drawn from the events space *p* (X_N ∈ `events p`); (b) *p* is a valid probability space (`prob_space p`); and lastly (c) the events in the given list X_N are independent (`MUTUAL_INDEP p` X_N). The function `PROB_LIST` takes an arbitrary list $[Z_1, Z_2, Z_3, \dots, Z_N]$ and returns a list of probabilities associated with the elements of the list [`prob p` Z_1,`prob p` Z_2,`prob p` $Z_3, \dots,$`prob p` Z_N], while the function `COMPL_LIST` takes a list

$[X_1, X_2, X_3, \ldots, X_N]$ and returns the complement of all elements in the list $[(1-X_1), (1-X_2), (1-X_3), \ldots, (1-X_N)]$. The function $\prod$ takes a list $[Y_1, Y_2, Y_3, \ldots, Y_N]$ and returns the product of the list elements $Y_1 \times Y_2 \times Y_3 \times \cdots \times Y_N$.

TABLE I: RBD Probabilistic Theorems

RBD Connection	Probabilistic Theorem
	**Theorem 1:** prob p (rbd_struct p (series X_N)) = $\prod$ (PROB_LIST p X_N)
	**Theorem 2:** prob p (rbd_struct p (parallel X_N)) = 1 - $\prod$ (PROB_LIST p (COMPL_LIST p X_N))

B. ET Formalization

Event Tree [5] (ET) is a widely used dependability modeling technique that can model all possible system-level components failure and success states and their cascading dependencies on the entire system in the form of a tree structure. The graph theory of an ET diagram starts by an initiating event called *Node* from which all possible consequence scenarios of an event that can occur in the system are drawn as *Branches* so that *only one* of these scenarios can occur (mutually exclusive). These ET constructors were formally modeled using a new recursive datatype EVENT_TREE, in HOL4 as [17]:

```
Hol_datatype EVENT_TREE = ATOMIC of (event) |
                  NODE of (EVENT_TREE list) |
     BRANCH of (event) (EVENT_TREE)
```

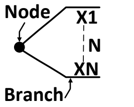
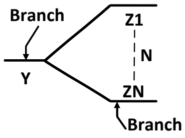
The type constructors NODE and BRANCH are recursive functions on EVENT_TREE-typed. A semantic function is then defined over the EVENT_TREE datatype that can yield a corresponding ET diagram as:

Definition 2:

```
⊢ ETREE (ATOMIC Y) = Y ∧
  ETREE (NODE (X::XN)) =
  ETREE X ∪ (ETREE (NODE XN)) ∧
  ETREE (BRANCH Y (Z::ZN)) = Y ∩ ETREE Z
```

The function ETREE takes a success/fail event Y, identified by an ET type constructor ATOMIC and returns the event Y. If the function ETREE takes a list XN of type EVENT_TREE, identified by a type constructor NODE, then it returns the union of all elements after applying the function ETREE on each element of the given list. Similarly, if the function ETREE takes a success/fail event X and a proceeding ET Z, identified by a type constructor of EVENT_TREE type, then it performs the intersection of the event X with the ET Z after applying the function ETREE. For the formal probabilistic assessment of each path occurrence in the ET diagram, HOL4 probabilistic properties for NODE and BRANCH ET constructors are presented in Table II [17]. These expressions are formally verified under the same RBD constrains, i.e., X_N ∈ events p, prob_space p, MUTUAL_INDEP p X_N, as well as ALL_DISTINCT X_N and disjoint X_N to ensure that each pair of elements in a given list X_N is distinct and mutually exclusive, respectively. The function $\sum$ takes a list $[X_1, X_2, X_3, \ldots, X_N]$ and returns the sum of the list elements $X_1 + X_2 + X_3 + \cdots + X_N$.

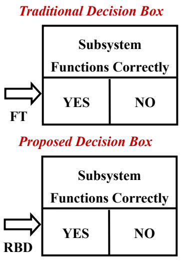
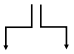
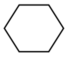
TABLE II: ET Probabilistic Theorems

ET Constructor	Probabilistic Theorem
	**Theorem 3:** prob p (ETREE (NODE X_N)) = $\sum$ (PROB_LIST p X_N)
	**Theorem 4:** prob p (ETREE (BRANCH Y(NODE $\mathcal{Z}_N$))) = (prob p Y) × $\sum$ (PROB_LIST p $\mathcal{Z}_N$)

III. Cause-Consequence Diagram Formalization

The graph theory of CCDs [20] uses three basic constructors *Decision box*, *Consequence path* and *Consequence box* [21]. The detailed description of the CCD constructors is illustrated in Table III. To present a clear understanding of these concepts, the traditional FT/ET-based CCD analysis for n-subsystems is described in Fig. 1. As shown in Fig. 1, FT *logic*-gates, such as AND and OR, are associated with all decision boxes to model the failure of generic n-subsystems, where the description of the used FT gates are presented in Table IV. It can be noticed from Fig. 1 that the output of each NO BOX for all decision boxes is equal to the subsystem FT

TABLE III: CCD Symbols and Functions

CCD Symbol	Function
 Traditional Decision Box Subsystem Functions Correctly FT YES NO Proposed Decision Box Subsystem Functions Correctly RBD YES NO	Decision Box: represents the status of functionality for a component or subsystem. (1) NO Box: describes the subsystem failure operation. An FT or RBD of the subsystem is connected to this box that can be used to determine the failure probability, i.e., $\mathcal{P}_{\text{NO}}$ = $\mathcal{P}_{\text{FT}}$ = 1 - $\mathcal{P}_{\text{RBD}}$ (2) YES Box: represents the correct functioning of the subsystem or reliability, which can be calculated by simply taking the complement of the failure operation, i.e., $\mathcal{P}_{\text{YES}}$ = 1 - $\mathcal{P}_{\text{FT}}$ = $\mathcal{P}_{\text{RBD}}$
	Consequence Path: models the next possible consequence scenarios due to the occurrence of subsystem failure or reliability
	Consequence Box: models the final outcome event due to a particular sequence of events for all connected subsystems

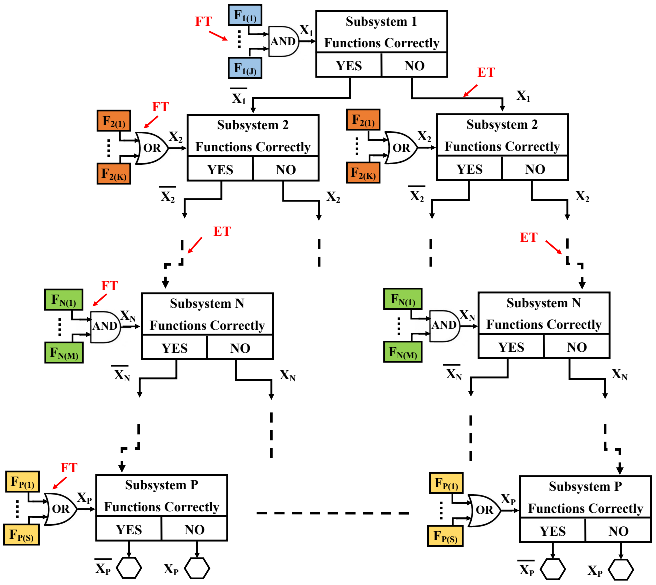


Fig. 1: FT/ET-based CCD Analysis

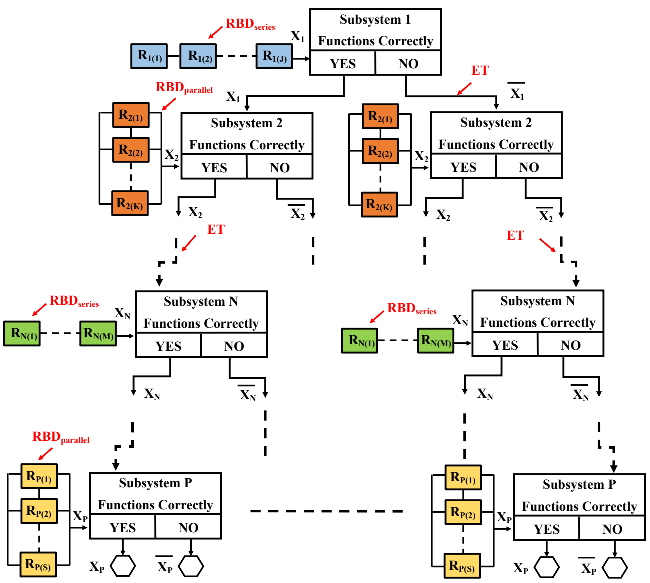


Fig. 2: Proposed RBD/ET-based CCD Analysis

TABLE IV: FT Symbols and Functions

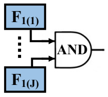
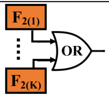
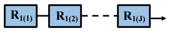
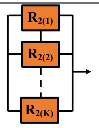
FT Symbol	Function
	AND Gate: models the complete failure of the subsystem if all of the input failure events $F_{1(1)}, \ldots, F_{1(J)}$ occur at the same time
	OR Gate: models the complete failure of the subsystem if any of the input failure events $F_{2(1)}, \ldots, F_{2(K)}$ occurs alone

TABLE V: RBD Symbols and Functions

RBD Symbol	Function
	Series: models the complete success of the subsystem if all of the input success events $R_{1(1)}, \ldots, R_{1(J)}$ occur at the same time
	Parallel: models the complete success of the subsystem if any of the input success events $R_{2(1)}, \ldots, R_{2(K)}$ occurs alone

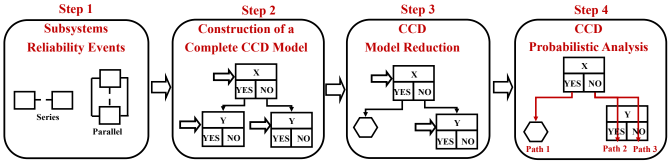

Step 1
Subsystems Reliability Events
Series
Parallel
Step 2
Construction of a Complete CCD Model
Step 3
CCD Model Reduction
Step 4
CCD Probabilistic Analysis
X
Y
YES
NO
Path 1
Path 2
Path 3

Fig. 3: Overview of RBD-based CCD Analysis

model (FT_X), while the YES BOX is the complement of the FT model ($\overline{FT_X}$). Analogously to Fig. 1, Fig. 2 illustrates the proposed RBD/ET-based CCD analysis, where different RBD configurations, such as *series* and *parallel* (see Table V), are associated with all CCD decision boxes to model the reliability of generic n-subsystems. As shown in Fig. 2, the output of each YES BOX for all decision boxes is equal to the RBD outcome (RBD_X), while the NO BOX is the complement of the RBD model ($\overline{RBD_X}$).

Fig. 3 depicts the overview of the developed *four* steps of cause-consequence safety analysis for complex systems [7]: (1) *Subsystems reliability events*: identify the success events for all subsystems using RBD models that keep the subsystems reliable in a complex system; (2) *Construction of a complete CCD*: build a full CCD diagram using its basic constructors (see Table III) considering that the order of components should follow the temporal action of the system; (3) *CCD model reduction*: remove the unnecessary decision boxes in the system to obtain its minimal CCD model representing the actual functional behavior of the complex system and reduce the number of test cases; and (4) *CCD probabilistic analysis*: determine the probabilities of all CCD consequence paths, which represent the likelihood of specific sequence scenarios that are possible to occur in a system so that *only one* scenario can occur [8]. This implies that all consequences in a CCD are mutually exclusive [21]. Assuming that all failure/reliability events associated with the decision box RBDs in a CCD model are mutually independent, then the probabilities of CCD consequence paths can be quantified as follows:

1) Evaluate the probabilities of each outgoing branch stemming from a *decision box*, i.e., quantifying the associated RBD models for all subsystems of a complex system
2) Compute the probability of each *consequence path* by multiplying the individual probabilities of all YES/NO events resulted from the decision boxes associated with it
3) Determine the probability of a particular *consequence box* corresponding to a complete/partial reliability or failure

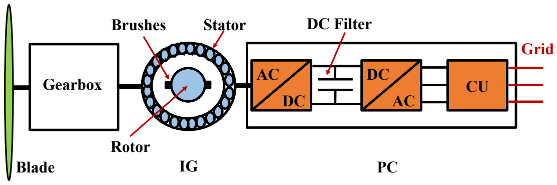


Fig. 4: Wind Turbine System [23]

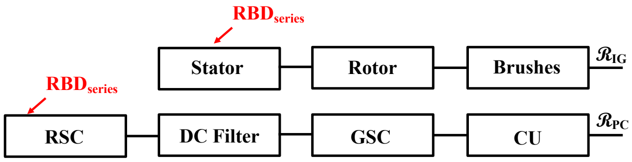


Fig. 5: RBD Models of Wind Turbine Subsystems

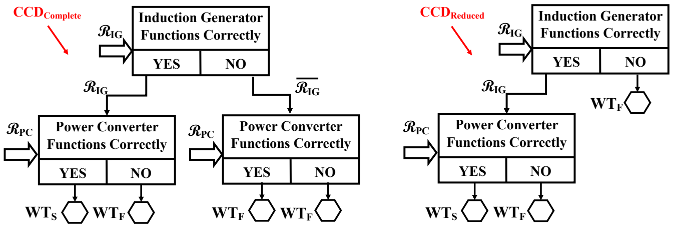

CCD_Complete
ℛ_IG
Induction Generator Functions Correctly
YES
NO
ℛ_IG
ℛ̄_IG
ℛ_PC
Power Converter Functions Correctly
YES
NO
WT_S
WT_F
ℛ_PC
Power Converter Functions Correctly
YES
NO
WT_F
WT_F
CCD_Reduced
ℛ_IG
Induction Generator Functions Correctly
YES
NO
ℛ_IG
WT_F
ℛ_PC
Power Converter Functions Correctly
YES
NO
WT_S
WT_F

Fig. 6: Wind Turbine Complete and Reduced CCD Models

event in the system by summing the probabilities of all consequence paths ending with that consequence event

As an example, consider a Wind Turbine system [22] consisting of two main subsystems Induction Generator (IG) and Power Converter (PC), as shown in Fig. 4 [23]. An IG consists of three components *Stator*, *Rotor* and *Brushes*, while a PC consists of four components *Rotor Side AC/DC Converter* (RSC), *DC Filter*, *Grid Side DC/AC Converter* (GSC) and *Control Unit* (CU). The *four* main steps of the above-mentioned RBD/ET-based cause-consequence analysis for the wind turbine system can be done as follows:

1) *Components reliability events*: Assign an RBD series configuration to each subsystem in the wind turbine, i.e., $\mathcal{R}_{\text{IG}}$, $\mathcal{R}_{\text{PC}}$, as shown in Fig. 5 [23], which can be expressed mathematically as:

$$\mathcal{R}_{\text{IG}} = \mathcal{R}_{\text{stator}} \times \mathcal{R}_{\text{rotor}} \times \mathcal{R}_{\text{brushes}} \tag{3}$$

$$\mathcal{R}_{\text{PC}} = \mathcal{R}_{\text{RSC}} \times \mathcal{R}_{\text{filter}} \times \mathcal{R}_{\text{GSC}} \times \mathcal{R}_{\text{CU}} \tag{4}$$

2) *Construction of a complete CCD*: Draw a complete CCD model of the wind turbine system, as shown in Fig. 6. For instance, if the condition of the IG decision box is either YES or NO, then the next subsystem PC is taken into consideration. Each consequence path in the CCD analysis ends with either a wind turbine success (WT_S) or a wind turbine failure (WT_F).
3) *CCD model reduction*: Apply the reduction operation on the constructed complete CCD model. For instance, if the condition of the IG decision box (IG functions correctly) is not satisfied, i.e., NO box, then the wind turbine fails regardless of the status of PC. Fig. 6 represents the minimal RBD/ET-based cause consequence analysis of the wind turbine operation.
4) *CCD probabilistic analysis*: The probabilistic assessment of the two consequence boxes WT_S and WT_F in Fig. 6 can be expressed mathematically as:

$$\begin{aligned}\mathcal{P}(Consequence\_Box_{WT_S}) &= \\ &\mathcal{P}(\text{IG}_{\text{YES}}) \times \mathcal{P}(\text{PC}_{\text{YES}})\end{aligned} \tag{5}$$

$$\begin{aligned}\mathcal{P}(Consequence\_Box_{WT_F}) &= \\ \mathcal{P}(\text{IG}_{\text{YES}}) \times &\mathcal{P}(\text{PC}_{\text{NO}}) + \mathcal{P}(\text{IG}_{\text{NO}})\end{aligned} \tag{6}$$

where $\mathcal{P}(\text{X}_{\text{YES}})$ is the reliability function outgoing from a subsystem decision box, i.e., $\mathcal{R}_{\text{X}}$ model, and $\mathcal{P}(\text{X}_{\text{NO}})$ is the unreliability function or the probability of failure, i.e., the complement of the $\mathcal{R}_{\text{X}}$ model ($\overline{\mathcal{R}_{\text{X}}}$).

A. Formal CCD Modeling

The CCD basic constructors *Decision box*, *Consequence path* and *Consequence box*, as described in Table III, were formally developed, in HOL4, respectively, as follows [18]:

Definition 3:

```
⊢ DECISION_BOX p X Y =
                if X = 1 then FST Y
           else if X = 0 then SND Y
           else p_space p
```

where `Y` is an ordered pair (`FST Y`, `SND Y`) representing the reliability and unreliability functions in a decision box, respectively. The condition `X = 1` represents the `YES Box` while `X = 0` represents the `NO Box`. If `X` is neither `1` nor `0`, for instance, `X = 2`, then this represents the irrelevance of the decision box, which returns the probability space *p* to be used in the reduction process of cause-consequence analysis.

Secondly, the CCD *Consequence path* is defined by recursively applying the `BRANCH` ET constructor (see Section II-B) on a given n-list of decision boxes ($\texttt{DECISION\_BOX}_{\mathcal{N}}$) using the HOL4 recursive function `FOLDL` from the *list* theory as:

Definition 4:

```
⊢ CONSEQ_PATH p (DECISION_BOX_1::DECISION_BOX_𝒩)
  = FOLDL (λa b. ETREE (BRANCH a b))
          DECISION_BOX_1 DECISION_BOX_𝒩
```

Finally, the CCD *Consequence box* is defined by mapping the function `CONSEQ_PATH` on a given *two-dimensional* list of consequence paths $\text{L}_{\mathcal{M}}$ using the HOL4 mapping function `MAP`, then apply the `NODE` ET constructor:

Definition 5:

```
⊢ CONSEQ_BOX p L_ℳ =
   ETREE (NODE (MAP (λa. CONSEQ_PATH p a) L_ℳ))
```

Using the above-mentioned CCD *generic* definitions, we can formally construct a complete CCD model (*Step 2* in Fig. 3) for the wind turbine shown in Fig. 6, in HOL4 as:

```
⊢ Wind_Turbine_COMPLETE_CCD $\mathcal{R}_{\text{IG}}$ $\mathcal{R}_{\text{PC}}$ =
  CONSEQ_BOX p [[DECISION_BOX p 1 ($\mathcal{R}_{\text{IG}}$,$\overline{\mathcal{R}_{\text{IG}}}$);
                  DECISION_BOX p 1 ($\mathcal{R}_{\text{PC}}$,$\overline{\mathcal{R}_{\text{PC}}}$)];
                 [DECISION_BOX p 1 ($\mathcal{R}_{\text{IG}}$,$\overline{\mathcal{R}_{\text{IG}}}$);
                  DECISION_BOX p 0 ($\mathcal{R}_{\text{PC}}$,$\overline{\mathcal{R}_{\text{PC}}}$)];
                 [DECISION_BOX p 0 ($\mathcal{R}_{\text{IG}}$,$\overline{\mathcal{R}_{\text{IG}}}$);
                  DECISION_BOX p 1 ($\mathcal{R}_{\text{PC}}$,$\overline{\mathcal{R}_{\text{PC}}}$)];
                 [DECISION_BOX p 0 ($\mathcal{R}_{\text{IG}}$,$\overline{\mathcal{R}_{\text{IG}}}$);
                  DECISION_BOX p 0 ($\mathcal{R}_{\text{PC}}$,$\overline{\mathcal{R}_{\text{PC}}}$)]]
```

In cause-consequence safety analysis [8], *Step 3* in Fig. 3 is to minimize the complete CCD model in the sense that the unnecessary decision boxes should be eliminated to decrease the number of test cases and model the accurate functional behavior of systems. Upon this, the reduced CCD model that actually represents the wind turbine system, as shown in Fig. 6, can be done formally by assigning `X` with neither `1` nor `0` options, for instance, `X = 2`, which represents the irrelevance of the decision box, in HOL4 as:

```
⊢ Wind_Turbine_REDUCED_CCD $\mathcal{R}_{\text{IG}}$ $\mathcal{R}_{\text{PC}}$ =
  CONSEQ_BOX p [[DECISION_BOX p 1 ($\mathcal{R}_{\text{IG}}$,$\overline{\mathcal{R}_{\text{IG}}}$);
                  DECISION_BOX p 1 ($\mathcal{R}_{\text{PC}}$,$\overline{\mathcal{R}_{\text{PC}}}$)];
                 [DECISION_BOX p 1 ($\mathcal{R}_{\text{IG}}$,$\overline{\mathcal{R}_{\text{IG}}}$);
                  DECISION_BOX p 0 ($\mathcal{R}_{\text{PC}}$,$\overline{\mathcal{R}_{\text{PC}}}$)];
                 [DECISION_BOX p 0 ($\mathcal{R}_{\text{IG}}$,$\overline{\mathcal{R}_{\text{IG}}}$);
                  DECISION_BOX p 2 ($\mathcal{R}_{\text{PC}}$,$\mathcal{R}_{\text{PC}}$)]]
```

Also, we can formally *verify* the above minimal CCD model of the wind turbine system after reduction, in HOL4 as:

```
⊢ Wind_Turbine_REDUCED_CCD $\mathcal{R}_{\text{IG}}$ $\mathcal{R}_{\text{PC}}$ =
  CONSEQ_BOX p [[DECISION_BOX p 1 ($\mathcal{R}_{\text{IG}}$,$\overline{\mathcal{R}_{\text{IG}}}$);
                  DECISION_BOX p 1 ($\mathcal{R}_{\text{PC}}$,$\overline{\mathcal{R}_{\text{PC}}}$)];
                 [DECISION_BOX p 1 ($\mathcal{R}_{\text{IG}}$,$\overline{\mathcal{R}_{\text{IG}}}$);
                  DECISION_BOX p 0 ($\mathcal{R}_{\text{PC}}$,$\overline{\mathcal{R}_{\text{PC}}}$)];
                 [DECISION_BOX p 0 ($\mathcal{R}_{\text{IG}}$,$\mathcal{R}_{\text{IG}}$)]]
```

B. Formal CCD Analysis

The last step in the cause-consequence analysis is to evaluate the probability of each path occurrence in the CCD model [21]. For that purpose, we propose the following novel CCD probabilistic mathematical formulations, based on RBD and ET modeling techniques, which have the capability to determine the probability of *n-level* CCD paths corresponding to n-subsystems in a critical system, where each subsystem consists of an arbitrary list of RBD events. Then, we provide the formalization of the proposed new formulas in HOL4.

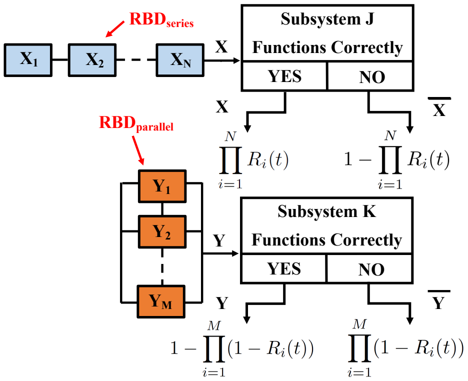


Fig. 7: CCD Decision Boxes with RBD Connections

a) One Decision Box: Fig. 7 depicts a single CCD decision box associated with either a series or a parallel RBD pattern. It can be observed that the `YES BOX` of the former CCD diagram with a series RBD model is the outcome of Eq. 1 and its `NO BOX` is the complement of Eq. 1. Similarly, the `YES BOX` of the later CCD diagram with a parallel RBD model is the outcome of Eq. 2 and its `NO BOX` is the complement of Eq. 2. The probability of a consequence path for each CCD decision box assigned with a *generic* RBD model consisting of n-events, i.e., series or parallel, as shown in Fig. 7, is verified under the constraints described in Table I (Section II-A), respectively, in HOL4 as:

Theorem 5:

```
⊢ let RBD$_{\text{series}}$ = rbd_struct p (series $X_N$)
   in prob_space p ∧ $X_N$ ∈ events p ∧
      MUTUAL_INDEP p $X_N$ ⇒
  prob p
    (CONSEQ_PATH p
        [DECISION_BOX p J
             (RBD$_{\text{series}}$, COMPL p (RBD$_{\text{series}}$))])
      = if J = 1 then ∏ (PROB_LIST p $X_N$)
  else if J = 0 then 1 - ∏ (PROB_LIST p $X_N$)
  else 1
```

Theorem 6:

```
⊢ let RBD$_{\text{parallel}}$ = rbd_struct p (parallel $Y_M$)
   in prob_space p ∧ $Y_M$ ∈ events p ∧
      MUTUAL_INDEP p $Y_M$ ⇒
  prob p
    (CONSEQ_PATH p
        [DECISION_BOX p K
             (RBD$_{\text{parallel}}$, COMPL p (RBD$_{\text{parallel}}$))])
      = if K = 1 then
  1 - ∏ (PROB_LIST p (COMPL_LIST p $Y_M$))
  else if K = 0 then
  ∏ (PROB_LIST p (COMPL_LIST p $Y_M$))
  else 1
```

where the function `COMPL` is defined to take a set X, which is the output of the RBD function `rbd_struct`, and returns the complement of the set X in the probability space p.

For a complex graph of CCDs consisting of n-level decision boxes, where each decision box is associated with a series/parallel RBD model consisting of an arbitrary list of success events, we define *three* types *A*, *B* and *C* with all possible CCD consequence scenarios that can occur.

b) N Decision Boxes (Type A): The probability of n-level decision boxes assigned to a consequence path corresponding to n-subsystems of a complex system, where each decision box is associated with a *generic* RBD model consisting of an arbitrary list of k-events in a *series* connection, can be expressed mathematically for *three* cases as:

(A1) All outcomes of n decisions boxes are YES

$$\mathcal{R}_{A1}(t) = \prod_{i=1}^{n} \prod_{j=1}^{k} \mathcal{R}_{ij}(t) \qquad (7)$$

(A2) All outcomes of n decisions boxes are NO

$$\mathcal{R}_{A2}(t) = \prod_{i=1}^{n}(1 - \prod_{j=1}^{k}\mathcal{R}_{ij}(t)) \tag{8}$$

(A3) Some outcomes of m decisions boxes are YES and the rest outcomes of p decisions boxes are NO, as shown in Fig. 8, respectively, as follows:

$$\mathcal{R}_{A3}(t) = \left(\prod_{i=1}^{m}\prod_{j=1}^{k}\mathcal{R}_{ij}(t)\right) \times \left(\prod_{i=1}^{p}(1 - \prod_{j=1}^{k}\mathcal{R}_{ij}(t))\right) \tag{9}$$

To formalize the above-proposed new cause-consequence mathematical formulations in HOL4, we formally define two *generic* functions $\mathcal{SS}^{YES}_{series}$ and $\mathcal{SS}^{NO}_{series}$ that can recursively generate the outcomes YES and NO of the RBD function rbd_struct, identified by the RBD basic constructor series, for a given arbitrary list of subsystems (SS) events, respectively as:

Definition 6:

⊢ $\mathcal{SS}^{YES}_{series}$ p (SS1::SSN) =
    rbd_struct p
     (series (rbd_list SS1))::$\mathcal{SS}^{YES}_{series}$ p SSN

Definition 7:

⊢ $\mathcal{SS}^{NO}_{series}$ p (SS::SSN) =
   COMPL p
    (rbd_struct p
     (series (rbd_list SS1)))::$\mathcal{SS}^{NO}_{series}$ p SSN

Using the above defined functions, we can verify *two-dimensional* and *scalable* CCD probabilistic properties corresponding to the proposed formulas Eq. 7, Eq. 8 and Eq. 9, respectively, in HOL4 as:

Theorem 10:

⊢ prob_space p ∧ MUTUAL_INDEP p SSN ∧
 ∀y. y ∈ SSN ⇒ y ∈ events p ∧ ⇒
 prob p (CONSEQ_PATH p ($\mathcal{SS}^{YES}_{series}$ p SSN)) =
   ∏ (MAP (λ a. ∏ (PROB_LIST p a)) SSN)

Theorem 11:

⊢ prob_space p ∧ MUTUAL_INDEP p SSN ∧
 ∀y. y ∈ SSN ⇒ y ∈ events p ∧ ⇒
 prob p (CONSEQ_PATH p ($\mathcal{SS}^{NO}_{series}$ p SSN)) =
 ∏ (MAP (λ b. (1 - ∏ (PROB_LIST p b))) SSN)

Theorem 12:

⊢ prob_space p ∧ MUTUAL_INDEP p (SSM ++ SSP) ∧
 ∀y. y ∈ (SSM ++ SSP) ⇒ y ∈ events p ∧ ⇒
 prob p
   (CONSEQ_PATH p
     [CONSEQ_PATH p ($\mathcal{SS}^{YES}_{series}$ p SSM);

      CONSEQ_PATH p ($\mathcal{SS}^{NO}_{series}$ p SSP)]) =
$\left(\prod \text{(MAP (}\lambda\text{ a. }\prod\text{ (PROB\_LIST p a)) SSM)}\right)$ ×
$\left(\prod \text{(MAP (}\lambda\text{ b. (1 - }\prod\text{ (PROB\_LIST p b))) SSP)}\right)$

where the assumptions of Theorems 10-12 are similar to the ones used in Theorems 1-4 (see Section II).

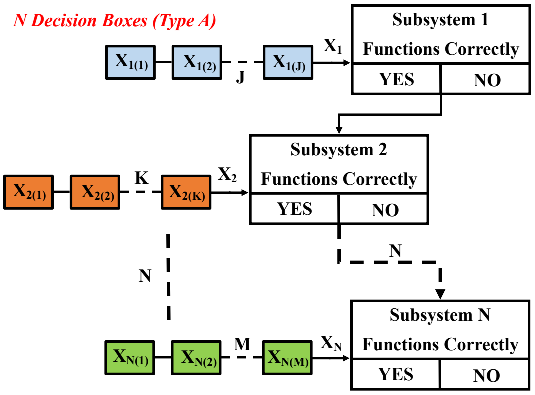


Fig. 8: N-level Decision Boxes for CCD Analysis of Type A

c) N Decision Boxes (Type B): Similarly, the probabilistic assessment of n-level decision boxes assigned to a CCD consequence path, where each decision box is associated with a *generic* RBD model consisting of k-events connected in *parallel*, can be expressed mathematically for *three* cases: (B1) All outcomes of n decisions boxes are YES; (B2) All outcomes of n decisions boxes are NO; and (B3) Some outcomes of m decisions boxes are YES and some outcomes of p decisions boxes are NO, as shown in Fig. 9, respectively, as follows:

$$\mathcal{R}_{B1}(t) = \prod_{i=1}^{n}(1 - \prod_{j=1}^{k}(1 - \mathcal{R}_{ij}(t))) \tag{10}$$

$$\mathcal{R}_{B2}(t) = \prod_{i=1}^{n}\prod_{j=1}^{k}(1 - \mathcal{R}_{ij}(t)) \tag{11}$$

$$\mathcal{R}_{B3}(t) = \left(\prod_{i=1}^{m}(1-\prod_{j=1}^{k}(1-\mathcal{R}_{ij}(t)))\right) \times \left(\prod_{i=1}^{p}\prod_{j=1}^{k}(1-\mathcal{R}_{ij}(t))\right) \tag{12}$$

To verify the correctness of the above-proposed new CCD mathematical formulas in HOL4, we define two *generic* functions $\mathcal{SS}^{YES}_{parallel}$ and $\mathcal{SS}^{NO}_{parallel}$ to recursively generate the outcomes YES and NO of the function rbd_struct, identified by the RBD basic constructor parallel, for a given list of subsystems events.

Definition 8:

⊢ $\mathcal{SS}^{YES}_{parallel}$ p (SS1::SSN) =
   rbd_struct p
   (parallel (rbd_list SS1))::$\mathcal{SS}^{YES}_{parallel}$ p SSN

Definition 9:

⊢ $\mathcal{SS}^{NO}_{parallel}$ p (SS::SSN) =
  COMPL p
   (rbd_struct p
    (parallel (rbd_list SS1)))::$\mathcal{SS}^{NO}_{parallel}$ p SSN

Using above defined functions, we can formally verify three *scalable* probabilistic properties corresponding to Eq. 10, Eq. 11, and Eq. 12, respectively, in HOL4 as:

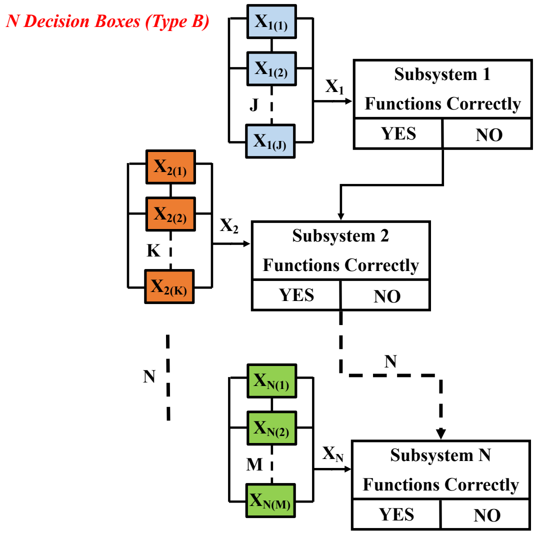


Fig. 9: N-level Decision Boxes for CCD Analysis of Type B

Theorem 13:

```
⊢ prob_space p ∧ MUTUAL_INDEP p SSN ∧
  ∀y. y ∈ SSN ⇒ y ∈ events p ∧ ⇒
  prob p
      (CONSEQ_PATH p ($\mathcal{SS}^{YES}_{parallel}$ p SSN)) =
   ∏
    (MAP
      (λ a.
        (1 - ∏
               (PROB_LIST p
                        (compl_list p a)))) SSN)
```

Theorem 14:

```
⊢ prob_space p ∧ MUTUAL_INDEP p SSN ∧
  ∀y. y ∈ SSN ⇒ y ∈ events p ∧ ⇒
  prob p
      (CONSEQ_PATH p ($\mathcal{SS}^{NO}_{parallel}$ p SSN)) =
   ∏
    (MAP
      (λ b.
        ∏ (PROB_LIST p (compl_list p b))) SSN)
```

Theorem 15:

```
⊢ prob_space p ∧ MUTUAL_INDEP p (SSM ++ SSP) ∧
  ∀y. y ∈ (SSM ++ SSP) ⇒ y ∈ events p ∧ ⇒
  prob p
      (CONSEQ_PATH p
          [CONSEQ_PATH p ($\mathcal{SS}^{YES}_{parallel}$ p SSM);

           CONSEQ_PATH p ($\mathcal{SS}^{NO}_{parallel}$ p SSP)]) =
   ∏
    (MAP
      (λ a.
        (1 - ∏
               (PROB_LIST p
                        (compl_list p a)))) SSM) ×
   ∏
    (MAP
      (λ b.
        ∏ (PROB_LIST p (compl_list p b))) SSP)
```

d) N Decision Boxes (Type C): The probabilistic assessment of n-level decision boxes assigned to a consequence path for a very complex system, where some m decision boxes are associated with *generic* RBD models consisting of k-events connected in *series*, while other p decision boxes are associated with *generic* RBD models consisting of z-events connected in *parallel*, as shown in Fig. 2, can be expressed mathematically for *nine* cases as:

(C1) All outcomes of m and p decisions boxes are YES.

$$\mathcal{R}_{C1}(t) = \left(\prod_{i=1}^{m}\prod_{j=1}^{k}\mathcal{R}_{ij}(t)\right) \times \left(\prod_{i=1}^{p}(1-\prod_{j=1}^{z}(1-\mathcal{R}_{ij}(t)))\right) \tag{13}$$

(C2) All outcomes of m and p decisions boxes are NO.

$$\mathcal{R}_{C2}(t) = \left(\prod_{i=1}^{m}(1-\prod_{j=1}^{k}\mathcal{R}_{ij}(t))\right) \times \left(\prod_{i=1}^{p}\prod_{j=1}^{z}(1-\mathcal{R}_{ij}(t))\right) \tag{14}$$

(C3) All outcomes of m decisions boxes are YES and all outcomes of p decisions boxes are NO.

$$\mathcal{R}_{C3}(t) = \left(\prod_{i=1}^{m}\prod_{j=1}^{k}\mathcal{R}_{ij}(t)\right) \times \left(\prod_{i=1}^{p}\prod_{j=1}^{z}(1-\mathcal{R}_{ij}(t))\right) \tag{15}$$

(C4) All outcomes of m decisions boxes are NO and all outcomes of p decisions boxes are YES.

$$\mathcal{R}_{C4}(t) = \left(\prod_{i=1}^{m}(1-\prod_{j=1}^{k}\mathcal{R}_{ij}(t))\right) \times \left(\prod_{i=1}^{p}(1-\prod_{j=1}^{z}(1-\mathcal{R}_{ij}(t)))\right) \tag{16}$$

(C5) Some outcomes of s out of m decisions boxes are YES, some outcomes of u out of m decisions boxes are NO and all outcomes of p decisions boxes are YES.

$$\mathcal{R}_{C5}(t) = \left(\prod_{i=1}^{s}\prod_{j=1}^{k}\mathcal{R}_{ij}(t)\right) \times \left(\prod_{i=1}^{u}(1-\prod_{j=1}^{k}\mathcal{R}_{ij}(t))\right) \times \left(\prod_{i=1}^{p}(1-\prod_{j=1}^{z}(1-\mathcal{R}_{ij}(t)))\right) \tag{17}$$

(C6) Some outcomes of s out of m decisions boxes are YES, some outcomes of u out of m decisions boxes are NO and all outcomes of p decisions boxes are NO.

$$\mathcal{R}_{C6}(t) = \left(\prod_{i=1}^{s}\prod_{j=1}^{k}\mathcal{R}_{ij}(t)\right) \times \left(\prod_{i=1}^{u}(1-\prod_{j=1}^{k}\mathcal{R}_{ij}(t))\right) \times \left(\prod_{i=1}^{p}\prod_{j=1}^{z}(1-\mathcal{R}_{ij}(t))\right) \tag{18}$$

(C7) Some outcomes of s out of p decisions boxes are YES, some outcomes of u out of p decisions boxes are NO and all outcomes of m decisions boxes are YES.

$$\mathcal{R}_{C7}(t) = \left(\prod_{i=1}^{m}\prod_{j=1}^{k}\mathcal{R}_{ij}(t)\right) \times \left(\prod_{i=1}^{u}\prod_{j=1}^{z}(1-\mathcal{R}_{ij}(t))\right) \times \left(\prod_{i=1}^{s}(1-\prod_{j=1}^{z}(1-\mathcal{R}_{ij}(t)))\right) \tag{19}$$

(C8) Some outcomes of s out of p decisions boxes are YES, some outcomes of u out of p decisions boxes are NO and all outcomes of m decisions boxes are NO.

$$\mathcal{R}_{C8}(t) = \left(\prod_{i=1}^{m}(1-\prod_{j=1}^{k}\mathcal{R}_{ij}(t))\right) \times \left(\prod_{i=1}^{u}\prod_{j=1}^{z}(1-\mathcal{R}_{ij}(t))\right) \times \left(\prod_{i=1}^{s}(1-\prod_{j=1}^{z}(1-\mathcal{R}_{ij}(t)))\right) \tag{20}$$

Using Theorems 5-15, we formally *verify* in HOL4 all the above-newly proposed formulas from Eq. 13 to Eq. 20 for RBD/ET-based cause consequence safety analysis (see Theorems 16-23, respectively, in [24]), which is evidence for the correctness of the proposed mathematical formulations.

(C9) Some outcomes of s out of m decisions boxes are YES, some outcomes of u out of m decisions boxes are NO, some outcomes of v out of p decisions boxes are YES and some outcomes of w out of p decisions boxes are NO.

$$\mathcal{R}_{C9}(t) = \left(\prod_{i=1}^{s}\prod_{j=1}^{k}\mathcal{R}_{ij}(t)\right) \times \left(\prod_{i=1}^{v}(1-\prod_{j=1}^{z}(1-\mathcal{R}_{ij}(t)))\right) \times \left(\prod_{i=1}^{u}(1-\prod_{j=1}^{k}\mathcal{R}_{ij}(t))\right) \times \left(\prod_{i=1}^{w}\prod_{j=1}^{z}(1-\mathcal{R}_{ij}(t))\right) \tag{21}$$

Theorem 24:

```
⊢ prob p
     (CONSEQ_PATH p
         [CONSEQ_PATH p ($\mathcal{SS}^{YES}_{series}$ p SSs);

          CONSEQ_PATH p ($\mathcal{SS}^{NO}_{series}$ p SSu);

          CONSEQ_PATH p ($\mathcal{SS}^{YES}_{parallel}$ p SSv);

          CONSEQ_PATH p ($\mathcal{SS}^{NO}_{parallel}$ p SSw)]) =
   ∏ (MAP (λ a. ∏ (PROB_LIST p a)) SSs)
× ∏ (MAP (λ b. 1 - ∏ (PROB_LIST p b)) SSu)
× ∏
    (MAP
      (λ c.
        (1 - ∏
              (PROB_LIST p
                         (compl_list p c)))) SSv)
× ∏
    (MAP
      (λ d.
        ∏ (PROB_LIST p (compl_list p d))) SSw)
```

e) A Consequence Box: Lastly, we verify a *generic* probabilistic formulation of a CCD CONSEQ_BOX for a certain event occurrence in the given complex system as the sum of all individual probabilities of all $\mathcal{M}$ CCD consequence paths ending with that event:

Theorem 25:

```
⊢ Let
   PATHS $L_{\mathcal{M}}$ = MAP (λa. CONSEQ_PATH p a) $L_{\mathcal{M}}$)
   in   prob_space p ∧ MUTUAL_INDEP p $L_{\mathcal{M}}$ ∧
disjoint (PATHS $L_{\mathcal{M}}$) ∧ ALL_DISTINCT (PATHS $L_{\mathcal{M}}$)
⇒ prob p (CONSEQ_BOX p $L_{\mathcal{M}}$) =
            ∑ (PROB_LIST p (PATHS $L_{\mathcal{M}}$))
```

where the assumptions of the above-theorem are quite similar to those used in Theorems 3 and 4 (see Section II-B). The verification of all the above-mentioned theorems was a bit challenging as we are dealing with all four types of different RBD configurations, i.e., series, the complement of series, parallel, and the complement of parallel, where each type is consisting of *generic* n-decision boxes and each decision box is associated with *generic* m-events, simultaneously in HOL4.

The proof-script of the formalization work presented in this section amounts to about 5,500 lines of HOL4 code and can be downloaded from [24]. In the next section, we present the formal RBD/ET-based CCD analysis of a smart grid system to illustrate the applicability of our proposed formal approach.

IV. Case Study: Smart Grid System

Smart Grid (SG) [25] is an interconnected network for delivering electricity from producers to customers. An SG system consists of three major sectors [26]: (i) generating stations; (ii) transmission grid; and (iii) distribution system. According to the policy of Renewable Energy Network for the 21st Century (REN21) [27], generating power from Renewable Energy Sources (RES), such as solar Photo-Voltaic (PV) and Wind Turbine (WT) farms, has become a mandatory requirement to be the best alternative for expanding fossil fuel generators [28]. The endeavor is to use 100% RES for generation stations by 2050 due to global warming [29]. A major challenge in SGs incorporating RES is to keep them safe from all disturbances and failures that could happen due to the intermittent nature of RES [30]. Therefore, it is a dire need to perform an adequate safety assessment of

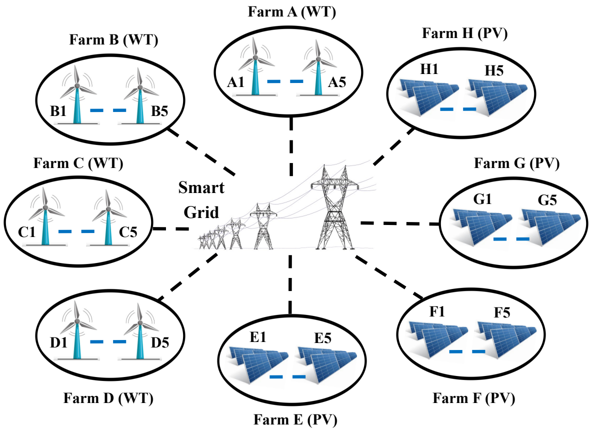


Fig. 10: Smart Grid System with RES Farms

the central SGs including RES subsystems at the subsystem level to determine the probabilities of disconnecting the least priority loads, which is well-known as *load-shedding* [31], and hence maintain the stability of SG and prevent it from an undesirable blackout. Fig. 10 depicts a smart microgrid system [32] supplied by 100% RES WT and PV farms, where each farm is consisting of *five* generating units connected in *parallel* and *series*, respectively. We can apply our proposed CCD formalization to verify the expressions all possible complete/partial safety classes for the SG system, i.e., SG *load-shedding* 0% (complete success), 12.5%, 25%, etc., 87.5% and 100% (complete failure), in HOL4 as:

Step 1 (Subsystem reliability events):
We formally define the RBD models of PV and WT farms, which are connected in parallel configurations, in HOL4 as:

Definition 10:

```
⊢ $\mathcal{R}_{\text{WTA}}$ p [WT_A1, WT_A2, ..., WT_A5] =
   rbd_struct p
        (parallel [WT_A1, WT_A2, ..., WT_A5])
```

Definition 11:

```
⊢ $\mathcal{R}_{\text{PVE}}$ p [PV_E1, PV_E2, ..., PV_E5] =
   rbd_struct p
        (series [PV_E1, PV_E2, ..., PV_E5])
```

where the other farms $\mathcal{R}_{\text{WTB}}$-$\mathcal{R}_{\text{WTD}}$ and $\mathcal{R}_{\text{PVF}}$-$\mathcal{R}_{\text{PVH}}$ are quite similar to Definitions 10 and 11, respectively.

Steps 2 and 3 (Construction of a CCD diagram):
We formally model the *8-level* CCD diagram of the SG system with 256 test cases, as shown in Fig. 11, in HOL4 as:

Definition 12:

```
⊢ CCD_SMART_GRID $\mathcal{R}_{\text{WTA}}$ ... $\mathcal{R}_{\text{WTD}}$ $\mathcal{R}_{\text{PVE}}$ ... $\mathcal{R}_{\text{PVH}}$
  = CONSEQ_BOX p
    [[DEC_BOX p 1 ($\mathcal{R}_{\text{WTA}}$,$\overline{\mathcal{R}_{\text{WTA}}}$); ...;
      DEC_BOX p 1 ($\mathcal{R}_{\text{WTD}}$,$\overline{\mathcal{R}_{\text{WTD}}}$);
      DEC_BOX p 1 ($\mathcal{R}_{\text{PVE}}$,$\overline{\mathcal{R}_{\text{PVE}}}$); ...;
      DEC_BOX p 1 ($\mathcal{R}_{\text{PVH}}$,$\overline{\mathcal{R}_{\text{PVH}}}$)];
                ⋮
     [DEC_BOX p 0 ($\mathcal{R}_{\text{WTA}}$,$\overline{\mathcal{R}_{\text{WTA}}}$); ...;
      DEC_BOX p 0 ($\mathcal{R}_{\text{WTD}}$,$\overline{\mathcal{R}_{\text{WTD}}}$);
      DEC_BOX p 0 ($\mathcal{R}_{\text{PVE}}$,$\overline{\mathcal{R}_{\text{PVE}}}$); ...;
      DEC_BOX p 0 ($\mathcal{R}_{\text{PVH}}$,$\overline{\mathcal{R}_{\text{PVH}}}$)]]
```

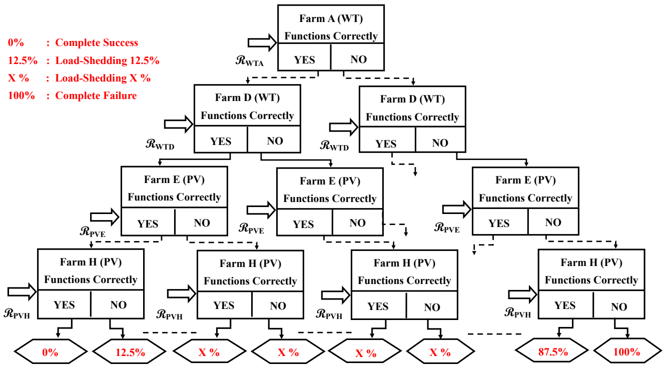

0% : Complete Success
12.5% : Load-Shedding 12.5%
X % : Load-Shedding X %
100% : Complete Failure
Farm A (WT) Functions Correctly
YES
NO
$\mathcal{R}_{\text{WTA}}$
Farm D (WT) Functions Correctly
YES
NO
$\mathcal{R}_{\text{WTD}}$
Farm D (WT) Functions Correctly
YES
NO
$\mathcal{R}_{\text{WTD}}$
Farm E (PV) Functions Correctly
YES
NO
$\mathcal{R}_{\text{PVE}}$
Farm E (PV) Functions Correctly
YES
NO
$\mathcal{R}_{\text{PVE}}$
Farm E (PV) Functions Correctly
YES
NO
$\mathcal{R}_{\text{PVE}}$
Farm H (PV) Functions Correctly
YES
NO
$\mathcal{R}_{\text{PVH}}$
Farm H (PV) Functions Correctly
YES
NO
$\mathcal{R}_{\text{PVH}}$
Farm H (PV) Functions Correctly
YES
NO
$\mathcal{R}_{\text{PVH}}$
Farm H (PV) Functions Correctly
YES
NO
$\mathcal{R}_{\text{PVH}}$
0%
12.5%
X %
X %
X %
X %
87.5%
100%

Fig. 11: CCD Analysis of the Smart Grid

Step 4 (Probabilistic analysis):
Using Theorems 5-25, we can *formally verify* the probabilistic expression at the subsystem-level for any of the safety classes that could occur in the SG. Assuming that all components of PVs and WTs are exponentially distributed (i.e., $\forall$ `t. 0` $\leq$ `t` $\Rightarrow$ `(Reliability p X t` $=$ $e^{(-\lambda_X t)}$`)`, where λ_X is the failure rate of the component X and t is a time index), we can, for example, verify the probabilistic expression of the outcome safety class *load-shedding 25%* of the SG system, in HOL4 as [24]:

Definition 13:

```
⊢ LOAD_SHED_25% $\mathcal{R}_{\text{WTA}}$ ... $\mathcal{R}_{\text{WTD}}$ $\mathcal{R}_{\text{PVE}}$ ... $\mathcal{R}_{\text{PVH}}$
  = CONSEQ_BOX p
    [[DEC_BOX p 1 ($\mathcal{R}_{\text{WTA}}$,$\overline{\mathcal{R}_{\text{WTA}}}$); ...;
      DEC_BOX p 0 ($\mathcal{R}_{\text{WTD}}$,$\overline{\mathcal{R}_{\text{WTD}}}$);
      DEC_BOX p 1 ($\mathcal{R}_{\text{PVE}}$,$\overline{\mathcal{R}_{\text{PVE}}}$); ...;
      DEC_BOX p 0 ($\mathcal{R}_{\text{PVH}}$,$\overline{\mathcal{R}_{\text{PVH}}}$)];
                ⋮
     [DEC_BOX p 1 ($\mathcal{R}_{\text{WTA}}$,$\overline{\mathcal{R}_{\text{WTA}}}$);
      DEC_BOX p 1 ($\mathcal{R}_{\text{WTB}}$,$\overline{\mathcal{R}_{\text{WTB}}}$); ...;
      DEC_BOX p 0 ($\mathcal{R}_{\text{PVG}}$,$\overline{\mathcal{R}_{\text{PVG}}}$);
      DEC_BOX p 0 ($\mathcal{R}_{\text{PVH}}$,$\overline{\mathcal{R}_{\text{PVH}}}$)]; ...]; ...]
```

Theorem 26:

```
⊢ prob p
   (LOAD_SHED_25% $\mathcal{R}_{\text{WTA}}$ ... $\mathcal{R}_{\text{WTD}}$ $\mathcal{R}_{\text{PVE}}$ ... $\mathcal{R}_{\text{PVH}}$)
```

$$
\begin{aligned}
= \Big(1 - & (1-e^{(-\lambda_{WT\_A1}t)}) \times \ldots \times (1-e^{(-\lambda_{WT\_A5}t)})\Big) \times \\
\ldots \times & \Big((1-e^{(-\lambda_{WT\_D1}t)}) \times \ldots \times (1-e^{(-\lambda_{WT\_D5}t)})\Big) \times \\
\ldots \times & \Big(e^{(-\lambda_{PV\_E1}t)} \times \ldots \times e^{(-\lambda_{PV\_E5}t)}\Big) \times \\
\ldots \times & \Big(1-e^{(-\lambda_{PV\_H1}t)} \times \ldots \times e^{(-\lambda_{PV\_H5}t)}\Big) + \ldots + \\
& \Big(1 - (1-e^{(-\lambda_{WT\_A1}t)}) \times \ldots \times (1-e^{(-\lambda_{WT\_A5}t)})\Big) \times \\
& \Big(1 - (1-e^{(-\lambda_{WT\_B1}t)}) \times \ldots \times (1-e^{(-\lambda_{WT\_B5}t)})\Big) \times \\
\ldots \times & \Big(1-e^{(-\lambda_{PV\_G1}t)} \times \ldots \times e^{(-\lambda_{PV\_G5}t)}\Big) \times \\
& \Big(1-e^{(-\lambda_{PV\_H1}t)} \times \ldots \times e^{(-\lambda_{PV\_H5}t)}\Big) + \ldots
\end{aligned}
$$

In the sequel, we compare our formally analysis results with those obtained from the paper-and-pencil approaches [33] as well as the MATLAB software based on Monte-Carlo Simulation (MCS) [9], which uses a random-based algorithm that predicts the real behavior patterns to estimate the average value of the various safety classes. We consider the failure rates of WTs and PVs λ_{WT_A-D} and λ_{PV_E-H} to be, respectively, 0.13 and 0.11 per year [34]. Assuming the study is undertaken after one year, i.e., t = 8760 hours, Table VI summarizes the manual, MATLAB and HOL4 results for some SG safety classes, where the HOL4 numerical values were obtained by a set of Standard Meta Language (SML) functions.

TABLE VI: Safety Classes Results of the Smart Grid

Safety Classes	Manual	MATLAB	HOL4
Load-Shedding 0%	11.08e-2	15.31e-2	11.0791226877e-2
Load-Shedding 12.5%	32.49e-2	26.45e-2	32.4963936847e-2
Load-Shedding 25%	35.74e-2	40.12e-2	35.7443052783e-2
Load-Shedding 75%	13.93e-11	20.93e-11	13.9295728515e-11
Load-Shedding 87.5%	24.99e-16	21.99e-16	24.9995145396e-16
Load-Shedding 100%	16.83e-21	25.83e-21	16.8254717074e-21
CPU Time (Seconds)	–	42.582	2.146

It can be noticed that the results of safety classes for the SG system obtained from our analysis are roughly equal to those calculated using paper-and-pencil, while MATLAB MCS uses a random-based algorithm, which estimates different results at every generation of a random number. This clearly elucidates that our approach provides the *first mechanical computation* of *n-level* cause-consequence probabilistic analysis ever augmented with the rigor of the HOL4 theorem prover. Moreover, the CPU time using the SML functions is much faster than MATLAB MCS (20X). The experiments were performed on core i5, 2.20 GHz, Linux VM with 1 GB of RAM.

V. Conclusion

In this paper, we proposed novel formulations of cause-consequence analysis, based on RBDs and ETs dependability modeling techniques, for the safety assessment of large systems. We provided a HOL4 formalization for the proposed equations that enables the formal probabilistic assessment of scalable CCD models associated with different RBD configurations and based on any probabilistic distribution and failure rates. Moreover, the proposed RBD/ET-based CCD formalization in HOL4 solves the scalability problem of n-level CCD analysis. We demonstrated the practical effectiveness of the proposed CCD formalization by performing the formal RBD/ET-based cause consequence analysis of a Smart Grid system, where we verified its probabilistic mathematical expressions for all possible multi-state safety classes of complete/partial reliability and failure consequence events. We also compared our approach with traditional CCD analysis techniques. As future work, we plan to develop an integrated framework with a GUI for CCD modeling and linking RBD/ET software simulation tools with the proposed CCD formalization in the HOL4 theorem prover.

References

[1] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart Grid—The New and Improved Power Grid: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 944–980, 2011.

[2] M. Rahman, "Power Electronics and Drive Applications for the Automotive Industry," in *Conference on Power Electronics Systems and Applications*. IEEE, 2004, pp. 156–164.

[3] M. Towhidnejad, D. R. Wallace, and A. M. Gallo, "Fault Tree Analysis for Software Design," in *NASA Goddard Software Engineering Workshop*, 2002, pp. 24–29.

[4] A. Brall, W. Hagen, and H. Tran, "Reliability Block Diagram Modeling-Comparisons of Three Software Packages," in *Rel. and Maintainability Symposium*, 2007, pp. 119–124.

[5] I. A. Papazoglou, "Mathematical Foundations of Event Trees," *Reliability Engineering & System Safety*, vol. 61, no. 3, pp. 169–183, 1998.

[6] M. Ridley, "Dependency Modelling Using Fault-Tree and Cause-Consequence Analysis," Ph.D. dissertation, Loughborough University, UK, 2000.

[7] J. Andrews and M. Ridley, "Reliability of Sequential Systems Using the Cause Consequence Diagram Method," *Part E: Journal of Process Mechanical Engineering*, vol. 215, no. 3, pp. 207–220, 2001.

[8] G. Vyzaite, S. Dunnett, and J. Andrews, "Cause-Consequence Analysis of Non-Repairable Phased Missions," *Reliability Engineering & System Safety*, vol. 91, no. 4, pp. 398–406, 2006.

[9] M. Wadi, M. Baysal, A. Shobole, and R. Tur, "Reliability Evaluation in Smart Grids via Modified Monte Carlo Simulation Method," in *International Conference on Renewable Energy Research and Applications*. IEEE, 2018, pp. 841–845.

[10] M. Bevilacqua, M. Braglia, and R. Gabbrielli, "Monte Carlo Simulation Approach for a Modified FMECA in a Power Plant," *Quality and Reliability Engineering International*, vol. 16, no. 4, pp. 313–324, 2000.

[11] F. Ortmeier, W. Reif, and G. Schellhorn, "Deductive Cause-Consequence Analysis," *IFAC Proceedings Volumes*, vol. 38, no. 1, pp. 62–67, 2005.

[12] SMV, 2020. [Online]. Available: http://www.cs.cmu.edu/~modelcheck/smv.html

[13] M. Güdemann, F. Ortmeier, and W. Reif, "Using Deductive Cause-Consequence Analysis (DCCA) with SCADE," in *International Conference on Computer Safety, Reliability, and Security*. Springer, 2007, pp. 465–478.

[14] HOL Theorem Prover, 2020. [Online]. Available: https://hol-theorem-prover.org

[15] W. Ahmad, "Formal dependability analysis using higher-order-logic theorem proving," Ph.D. dissertation, National University of Sciences & Technology, Pakistan, 2017.

[16] Y. Elderhalli, "Dynamic Dependability Analysis using HOL Theorem Proving with Application in Multiprocessor Systems," Ph.D. dissertation, Concordia University, Canada, 2019.

[17] M. Abdelghany, W. Ahmad, and S. Tahar, "A Formally Verified HOL4 Algebra for Event Trees," 2020. [Online]. Available: http://arxiv.org/abs/2004.14384

[18] M. Abdelghany and S. Tahar, "Formal FT-based Cause-Consequence Reliability Analysis using Theorem Proving," 2021. [Online]. Available: https://arxiv.org/abs/2101.07174

[19] W. Ahmed, O. Hasan, and S. Tahar, "Formalization of Reliability Block Diagrams in Higher-Order Logic," *Journal of Applied Logic*, vol. 18, pp. 19–41, 2016.

[20] B. Xin, L. Wan, J. Yu, and W. Dang, "Basic Event Probability Determination and Risk Assessment Based on Cause-Consequence Analysis Method," in *Journal of Physics*, vol. 1549, no. 5. IOP Publishing, 2020, p. 052094.

[21] J. Andrews and M. Ridley, "Application of the Cause-Consequence Diagram Method to Static Systems," *Reliability Engineering & System Safety*, vol. 75, no. 1, pp. 47–58, 2002.

[22] F. Porté-Agel, M. Bastankhah, and S. Shamsoddin, "Wind-Turbine and Wind-Farm Flows: a Review," *Boundary-Layer Meteorology*, vol. 174, no. 1, pp. 1–59, 2020.

[23] S. Jaiswal and G. Pahuja, "Effect of Reliability of Wind Power Converters in Productivity of Wind Turbine," in *International Conference on Power Electronics*. IEEE, 2014, pp. 1–6.

[24] "RBD/ET based Cause-Consequence Formalization in HOL4," 2020. [Online]. Available: https://github.com/hvg-concordia/CCD_RBD

[25] A. Keyhani and M. Albaijat, *Smart Power Grids*. Springer Sci. & Bus. Media, 2012.

[26] S. Xu, Y. Qian, and R. Q. Hu, "On Reliability of Smart Grid Neighborhood Area Networks," *IEEE Access*, vol. 3, pp. 2352–2365, 2015.

[27] R. Adib, H. E. Murdock, F. Appavou *et al.*, "Renewables 2015 Global Status Report," *REN21 Secretariat, Paris, France*, 2015.

[28] D. Connolly, H. Lund, and M. Mathiesen, B. V.and Leahy, "The First Step Towards a 100% Renewable Energy-System for Ireland," *Applied Energy*, vol. 88, no. 2, pp. 502–507, 2011.

[29] H. Lund and B. V. Mathiesen, "Energy System Analysis of 100% Renewable Energy Systems—The Case of Denmark in Years 2030 and 2050," *Energy*, vol. 34, no. 5, pp. 524–531, 2009.

[30] I. Vokony and A. Dán, "Examination of Smart Grids in Island Operation," in *Bucharest PowerTech*. IEEE, 2009, pp. 1–7.

[31] M. Marzband, M. M. Moghaddam, M. F. Akorede, and G. Khomeyrani, "Adaptive Load Shedding Scheme for Frequency Stability Enhancement in Microgrids," *Electric Power Syst. Research*, vol. 140, pp. 78–86, 2016.

[32] N. Hatziargyriou, H. Asano, R. Iravani, and C. Marnay, "Microgrids," *Power and Energy Magazine*, vol. 5, no. 4, pp. 78–94, 2007.

[33] R. Karki, R. Billinton, and A. K. Verma, *Reliability Modeling and Analysis of Smart Power Systems*. Springer Sci. & Business Media, 2014.

[34] R. Yokoyama, T. Niimura, and N. Saito, "Modeling and Evaluation of Supply Reliability of Microgrids including PV and Wind Power," in *Power and Energy Soc. General Meeting-Conversion and Delivery of Elect. Energy in the 21st Century*. IEEE, 2008, pp. 1–5.