Using Minimal Inductive Validity Cores to Generate Minimal Cut Sets

Danielle Stewart[1], Michael Whalen[1], Mats Heimdahl[1], Jing (Janet) Liu[2], and Darren Cofer[2]

[1] University of Minnesota, Minneapolis, MN, USA,
{dkstewar, mwwhalen, heimdahl}@umn.edu
[2] Collins Aerospace - Trusted Systems, Cedar Rapids, IO, USA,
{jing.liu, darren.cofer}@collins.com

Abstract. Risk and fault analysis are activities that help to ensure that critical systems operate in an expected way, even in the presence of component failures. As critical systems become more dependent on software components, analyses of error propagation through these software components becomes crucial. These analyses should be understandable to the analyst, scalable, and sound, in order to provide sufficient guarantees that the system is safe. A commonly used safety artifact is the set of all *minimal cut sets*, minimal sets of faults that may lead to a violation of a safety property. In this research, we define how minimal cut sets can be derived from certain results of model checking, the Minimal Inductive Validity Cores (MIVCs). Using a compositional model checking approach, we can incorporate both hardware and software failures and auto-generate safety artifacts. This research describes a technique for determining the minimal cut sets by the use of MIVCs and producing compositionally derived artifacts that encode pertinent system safety information. We describe our technique, prove that it is sound, and demonstrate it in an implementation in the OSATE tool suite for AADL.

1 Introduction

Risk and safety analyses are important activities used to ensure that critical systems operate in an expected way. From nuclear power plants and airplanes to heart monitors and automobiles, critical systems are vitally important in our society. These systems are required to not only operate safely under nominal (normal) conditions, but also under conditions when faults are present in the system. Guaranteeing that system safety properties hold in the presence of faults is an important aspect of critical systems development and falls under the discipline of safety analysis. Safety analysis produces various safety related artifacts that are often used during the development process of critical systems [**?**,**?**]. Many of these safety artifacts require the generation of *Minimal Cut Sets*, the minimal sets of faults that cause the violation of a system safety property. Since the introduction of minimal cut sets in the field of safety analysis [**?**], much research has been performed to address the generation of these sets [**?**, **?**, **?**, **?**, **?**]. One of the challenges with minimal cut set generation is scaling to industrial-sized systems. As the system gets larger, more minimal cut sets are possible with increasing cardinality. In recent years, symbolic model checking has been used to scale the analysis of systems with millions of minimal cut sets [**?**, **?**, **?**, **?**, **?**, **?**].

Recently, Ghassabani et al. developed an algorithm that traces a safety property to a minimal set of model elements necessary for proof; this is called the *all minimal inductive validity core* algorithm (`All_MIVCs`) [**?**,**?**,**?**]. Inductive validity cores produce the minimal set of model elements necessary to prove a property. Each set contains the *behavioral contracts* – the requirement specifications for components – of the model used in a proof. The `All_MIVCs` algorithm gives the minimal set of contracts required for proof of a safety property. If all of these sets are obtained, we have insight

into every proof path for the property. Thus, if we violate at least one contract from every MIVC set, we have in essence "broken" every proof path. This is the information that is used to perform fault analysis using MIVCs.

Safety analysts are often concerned with faults in the system, i.e., when components or subsystems deviate from nominal behavior, and the propagation of errors through the system. To this end, the model elements included in the reasoning process of the `All_MIVCs` algorithm are not only the contracts of the system, but faults as well. This will provide additional insight into how an active fault may violate contracts that directly support the proof of a safety property.

In complex critical systems, safety analysts are concerned with hardware faults, how these may propagate to software components reliant on the failed hardware, and other faults whose propagation requires insight into system dynamics. Scaling model checking of complex hardware and software is challenging; one way to address this problem is to take advantage of the architecture of the system model through a *compositional* approach [**?**, **?**, **?**]. Compositional model checking reduces the verification of a large system into multiple smaller verification problems that can be solved independently and which together guarantee correctness of the original problem. One way to structure compositional verification is hierarchically: layers of the system architecture are analyzed independently and their composition demonstrates a system property of interest.

This paper proposes a new method of minimal cut set generation using compositional model checking, allowing us to reason uniformly about faults in hardware and software and their impact (propagation) to system properties. The main contributions of this research are summarized as follows:

- We propose a novel method for minimal cut set generation using Minimal Inductive Validity Cores (MIVCs) generated during model checking.
- We provide proof of the soundness of this method.
- We discuss the implementation of the algorithm for compositional cut set generation.

The organization of the paper is as follows. Section 2 provides a running example, Section 3 provides the preliminaries for Section 4 which outlines the formalisation of this approach. The implementation of the algorithms is discussed in Section 5 and related work follows in Section 6. The paper ends with a conclusion and discussion of related work.

2 Running Example

We present a running example of a simplified sensor system in a Pressurized Water Reactor (PWR). In a typical PWR, the core inside of the reactor vessel produces heat. Pressurized water in the primary coolant loop carries the heat to the steam generator. Within the steam generator, heat from the primary coolant loop vaporizes the water in a secondary loop, producing steam. The steamline directs the steam to the main turbine, causing it to turn the turbine generator, which produces electricity. There are a few important factors that must be considered during safety assessment and system design. An unsafe climb in temperature can cause high pressure and hence pipe rupture, and high levels of radiation could indicate a leak of primary coolant.

The following sensor system can be thought of as a subsystem within a PWR that monitors these factors. A diagram of the model is shown in Figure **??** and represents a highly simplified version of a safety critical system. The temperature subsystem details are shown at the bottom of Figure **??**; each of the subsystems have a similar architecture.

The subsystems each contain three sensors that monitor pressure, temperature, and radiation. Environmental inputs are fed into each sensor in the model and the redundant sensors monitor tem-

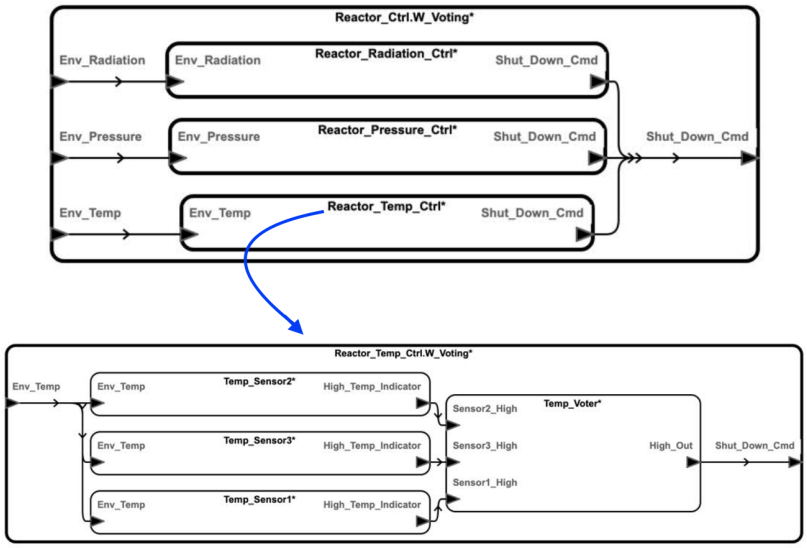

Fig. 1. PWR Sensor System

perature, pressure, or radiation respectively. If temperature, pressure, or radiation is too high, a shut down command is sent from the sensors to the parent components.

2.1 PWR Nominal Model

The temperature, pressure, and radiation sensor subsystems use a majority voting mechanism on the sensor values and will send a shut down command based on this output. The safety property of interest in this system is: *shut down when and only when we should*; the AGREE guarantee stating this property is shown in Figure **??**.

```
guarantee "Shut down when and only when we should":
    Shut_Down_Cmd =
        ((Env_Temp> HIGH_TEMPERATURE_THRESHOLD)  or
        (Env_Pressure > HIGH_PRESSURE_THRESHOLD) or
        (Env_Radiation > HIGH_RADIATION_THRESHOLD));
```

Fig. 2. Sensor System Safety Property

The safety of the system requires a shut down to take place if the temperature, pressure, or radiation levels climb beyond safe levels; thus, a threshold for each subsystem is introduced. If any sensor subsystem reports passing that threshold, a shutdown command is sent. Supporting guarantees are located in each sensor subsystem and correspond to temperature, pressure, and radiation sending a shut down command if sensed inputs are above a given threshold. Each sensor has a similar guarantee. For reference throughout this paper, we provide Figure **??** which shows the guarantees and faults of interest for this running example.

Note: the thresholds vary for pressure, temperature, and radiation. These are given as constants T_p, T_t, and T_r respectively. The overall (or "top level") shutdown command is defined notationally

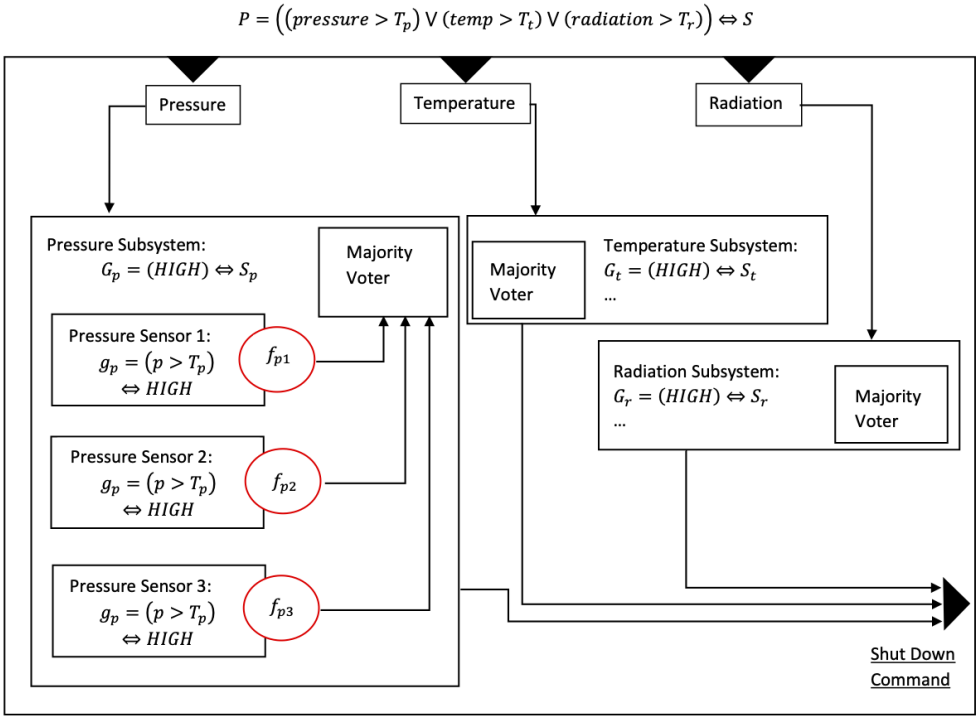

Fig. 3. Sensor System Nominal and Fault Model Details

as S; each sensor subsystem provides their own shutdown command, S_p for example. The faults are shown as "fail low" which correspond to the temp (or pressure or radiation) being high, but the sensor reports safe ranges. We also do not list all guarantees and assumptions that are in the model, but only the ones of interest for the illustration.

2.2 PWR Fault Model

The faults that are of interest in this example system are any one of the sensors failing high or low. If sensors report high and a shut down command is sent, we shut down when we should not. On the other hand, if sensors report low when it should be high, a shut down command is not sent and we do not shut down when we should. From the perspective of safety, a false report of low temperature is the main concern. For simplification in this paper, we focus on the failures when sensors report low when they should not.

A fault is defined for each sensor in the system using the safety annex. An example of a temperature sensor fault stuck at high is shown in Figure **??**.

The Safety Annex provides a way to weave the faults into the nominal model by use of the *inputs* and *outputs* keywords. This allows users to define a fault and attach it to the output of a component.

```
annex safety {**
    fault temp_sensor_stuck_at_high "temp sensor stuck at high": Common_Faults.stuck_true {
        inputs: val_in <- High_Temp_Indicator;
        outputs: High_Temp_Indicator <- val_out;
        probability: 1.0E-5 ;
        duration: permanent;
    }
**};
```

Fig. 4. Fault on Temperature Sensor Defined in the Safety Annex for AADL

The fault shown in Figure **??** is defined to be a *permanent* fault and has probability of occurrence set at 1.0×10^{-5}. If the fault is active, the error can possibly violate the guarantees of this component or the assumptions of downstream components [**?**]. The activation of a fault is not up to the user, but instead left up to the model checker, JKind, to determine if the activation of this fault will contribute to a violation of higher level guarantees. If so, it can be activated during the analysis.

3 Preliminaries

My suggestion is to place this at the front of the implementation section. It just seems like it is taking longer to get to the interesting part of the paper by having it here.

The algorithms in this paper are implemented in the Safety Annex for the Architecture Analysis and Design Language (AADL) and require the Assume-Guarantee Reasoning Environment (AGREE) [**?**] to annotate the AADL model in order to perform verification using the back-end model checker `JKind` [**?**].

Architecture Analysis and Design Language We are using the Architectural Analysis and Design Language (AADL) to construct system architecture models of performance-critical, embedded, real-time systems [**?**, **?**]. Language annexes to AADL provide a richer set of modeling elements for various system design and analysis needs, and the language definition is sufficiently rigorous to support formal analysis tools that allow for early phase error/fault detection.

Compositional Analysis One way to structure compositional verification is hierarchically: layers of the system architecture are analyzed independently and their composition demonstrates a system property of interest. Compositional verification partitions the formal analysis of a system architecture into verification tasks that correspond into the decomposition of the architecture [**?**]. A proof consists of demonstrating that the system property is provable given the contracts of its direct subcomponents and the system assumptions [**?**, **?**]. When compared to monolithic analysis (i.e., analysis of the flattened model composed of all components), the compositional approach allows the analysis to scale to much larger systems [**?**, **?**, **?**].

Assume Guarantee Reasoning Environment The Assume Guarantee Reasoning Environment (AGREE) is a tool for formal analysis of behaviors in AADL models and supports compositional verification [**?**]. It is implemented as an AADL annex and is used to annotate AADL components with formal behavioral contracts. Each component's contracts includes assumptions and guarantees about the component's inputs and outputs respectively. AGREE translates an AADL model and the behavioral contracts into Lustre [**?**] and then queries the `JKind` model checker to conduct the back-end analysis [**?**].

JKind JKind is an open-source industrial infinite-state inductive model checker for safety properties [**?**]. Models and properties in JKind are specified in Lustre [**?**], a synchronous dataflow language, using the theories of linear real and integer arithmetic. JKind uses SMT-solvers to prove and falsify multiple properties in parallel.

Safety Annex for AADL The Safety Annex for AADL provides the ability to reason about faults and faulty component behaviors in AADL models [**?**, **?**]. In the Safety Annex approach, AGREE is

used to define the nominal behavior of system components, faults are introduced into the nominal model, and JKind is used to analyze the behavior of the system in the presence of faults. Faults describe deviations from the nominal behavior and are attached to the outputs of components in the system.

3.1 Formal Background

Given a state space U, a transition system (I, T) consists of an initial state predicate $I : U \rightarrow bool$ and a transition step predicate $T : U \times U \rightarrow bool$. We define the notion of reachability for (I, T) as the smallest predicate $R : U \rightarrow bool$ which satisfies the following formulas:

$$\forall u.\ I(u) \Rightarrow R(u)$$
$$\forall u, u'.\ R(u) \wedge T(u, u') \Rightarrow R(u')$$

A safety property $P : U \rightarrow bool$ is a state predicate. A safety property P holds on a transition system (I, T) if it holds on all reachable states, i.e., $\forall u.\ R(u) \Rightarrow P(u)$, written as $R \Rightarrow P$ for short. When this is the case, we write $(I, T) \vdash P$. We assume the transition relation has the structure of a top level conjuntion. Given $T(u, u') = T_1(u, u') \wedge \cdots T_n(u, u')$ we will write $T = \wedge_{i=1..n}$ for short. By further abuse of notation, T is identified with the set of its top-level conjuncts. Thus, $T_i \in T$ means that T_i is a top-level conjunct of T, and $S \subseteq T$ means all top level conjuncts of S are top-level conjuncts of T. When a top-level conjunct T_i is removed from T, we write $T \setminus \{T_i\}$

The idea behind finding an IVC for a given property P [**?**] is based on inductive proof methods used in SMT-based model checking, such as k-induction and IC3/PDR [**?**, **?**, **?**]. Generally, an IVC computation technique aims to determine, for any subset $S \subseteq T$, whether P is provable by S. Then, a minimal subset that satisfies P is seen as a minimal proof explanation called a minimal inductive validity core. Ghassabani et al. demonstrate that the minimization process is as hard as model checking [**?**], so finding a minimal inductive validity core may not be possible for some model checking problems.

Definition 1. *Inductive Validity Core (IVC) [?]: $S \subseteq T$ for $(I, T) \vdash P$ is an Inductive Validity Core, denoted by $IVC(P, S)$, iff $(I, S) \vdash P$.*

Definition 2. *Minimal Inductive Validity Core (MIVC) [?]: $S \subseteq T$ is a minimal Inductive Validity Core, denoted by $MIVC(P, S)$, iff $IVC(P, S) \wedge \forall T_i \in S.\ (I, S \setminus \{T_i\}) \nvdash P$.*

A k-induction model checker utilizes parallel SMT-solving engines at each induction step to glean information about the proof of a safety property. The transition formula is translated into clauses such that satisfiability is preserved [**?**]. The translated system, consisting of the constrained formulas of the transition system and the negation of the property, is often called a *constraint system*. The `All_MIVCs` algorithm collects all *minimal unsatisfiable subsets* (MUSs) of a constraint system generated from a transition system at each induction step. [**?**, **?**].

Definition 3. *A Minimal Unsatisfiable Subset (MUS) [?] M of a constraint system C is a set $M \subseteq C$ such that M is unsatisfiable and $\forall c \in M : M \setminus \{c\}$ is satisfiable.*

The MUSs are the minimal explanation of the infeasibility of this constraint system; equivalently, these are the minimal sets of model elements necessary for proof of the safety property.

Returning to our running example, this can be illustrated by the following. Given the constraint system $C = \{G_p, G_t, G_r, \neg P\}$, a minimal explanation of the infeasability of this system is the set $\{G_p, G_t, G_r, \}$. If all three guarantees hold, then P is provable.

A related set is a *minimal correction set*:

Definition 4. *A Minimal Correction Set (MCS) [?] M of a constraint system C is a subset $M \subseteq C$ such that $C \setminus M$ is satisfiable and $\forall M' \subset M : C \setminus M'$ is unsatisfiable.*

A MCS can be seen to "correct" the infeasability of the constraint system by the removal from C the constraints found in an MCS. In the case of an UNSAT system, we may ask: what will correct this unsatisfiability? Returning to the PWR example, we can find the MCSs of the constraint system C: $MCS_1 = \{G_t\}$, $MCS_2 = \{G_p\}$, $MCS_3 = \{G_r\}$. If any single guarantee is violated, a shut down from that subsystem will not get sent when it should and the safety property P will be violated.

A duality exists between the MUSs of a constraint system and the MCSs as established by Reiter [**?**]. This duality is defined in terms of *Minimal Hitting Sets* (*MHS*).

Definition 5. *A hitting set of a collection of sets A is a set H such that every set in A is "hit" by H; H contains at least one element from every set in A.*

Every MUS of a constraint system is a minimal hitting set of the system's MCSs, and likewise every MCS is a minimal hitting set of the system's MUSs. This is noted in previous work [**?**, **?**] and the proof of such is given by Reiter (Theorem 4.4 and Corollary 4.5) [**?**]. For the PWR top level constraint system, it can be seen that each of the MCSs intersected with the MUS is nonempty. This gives the minimal set of guarantees for which, if violated, will cause P to be violated.

4 Formalization

The set of all nominal guarantees of the system G consists of conjunctive constraints $g \in G$. Given no faults (i.e., nominal system) and a transition relation T consisting of conjunctive constraints T_i as defined in section **??**, each g is one of the transition constraints T_i where:

$$T = g_1 \wedge g_2 \wedge \cdots \wedge g_n \tag{1}$$

We consider an arbitrary layer of analysis of the architecture and assume the property holds of the nominal relation $(I, T) \vdash P$. Given that our focus is on safety analysis in the presence of faults, let the set of all faults in the system be denoted as F. A fault $f \in F$ is a deviation from the normal constraint imposed by a guarantee. Without loss of generality, we associate a single fault and an associated fault probability with a guarantee. Each fault f_i is associated with an *activation literal*, af_i, that determines whether the fault is active or inactive.

To consider the system under the presence of faults, consider a set GF of modified guarantees in the presence of faults and let a mapping be defined from activation literals $af_i \in AF$ to these modified guarantees $gf_i \in GF$.

$$gf_i = \textit{if } af_i \textit{ then } f_i \textit{ else } g_i$$

The transition system is composed of the set of modified guarantees GF and a set of conjunctions assigning each of the activation literals $af_i \in AF$ to false:

$$T' = gf_1 \wedge gf_2 \wedge \cdots \wedge gf_n \wedge \neg af_1 \wedge \neg af_2 \wedge \cdots \wedge \neg af_n \tag{2}$$

Theorem 1. *If $(I, T) \vdash P$ for T defined in equation **??**, then $(I, T') \vdash P$ for T' defined in equation **??**.*

Proof. By the mapping of each constrained activation literal $\neg af_i$ to the associated guarantee g_i and the weakening of the antecedent by introduction of the activation literals, the result is immediate. □

Consider the elements of T as a set $GF \cup AF$, where GF are the potentially faulty guarantees and AF consists of the activation literals that determine whether a guarantee is faulty. This is a set that is considered by an SMT-solver for satisfiability during the model checking engine procedures. The posited problem is thus: $GF \wedge AF \wedge \neg P$ for the safety property in question.

Let us view this in terms of the PWR system example and focus on the temperature sensor subsystem. The safety property to be proved is G_t, the supporting guarantees are found in each of the three temperature sensors, g_{ti}. Faults f_{ti} are defined for each sensor. The transition system is:

$$T = gf_{t1} \wedge gf_{t2} \wedge gf_{t3} \wedge \neg af_{t1} \wedge \neg af_{t2} \wedge \neg af_{t3}$$

The MIVCs for this subsystem layer correspond to all pairwise combinations of constrained activation literals. Intuitively, if any two sensor faults do *not* occur, then two of the three sensor guarantees are not violated and the system responds appropriately to high temperature; therefore, G_t is provable.

The MCSs for this subsystem layer happen to also correspond to all pairwise combinations of constrained activation literals. If any two sensor faults *do* occur, then two of the three sensor guarantees will be violated and the system does not respond to high temperature as required. This would result in the inability to prove G_t. (Note: it is not always the case that the MCSs are the same as the MIVCs – in this case it is due to majority voting on three sensors.)

4.1 Transforming MCS into Minimal Cut Set

The MCSs contain the information needed to find minimal cut sets, but their elements consist of constrained activation literals and guarantees, whereas minimal cut sets contain faults. An activation literal af_i represents a potentially active fault and f_i is the manifestation of the fault in the transition system. For ease of notation, we define a mapping $\sigma : AF \rightarrow F$ where $\sigma(af_i) = f_i$. Let $MCS = \{af_1, \cdots, af_m\}$ and let $\sigma(MCS) = \{\sigma(af_1), \cdots, \sigma(af_m)\} = \{f_1, \ldots, f_m\}$ be a mapping where MCS is a minimal correction set with regard to some property P and $MCS \subseteq AF$.

Lemma 1. *$\sigma(MCS)$ is a cut set of P.*

Proof. Assume towards contradiction that $\sigma(MCS)$ is not a cut set of P. Then $gf_1 \wedge \cdots \wedge gf_n \wedge af_1 \cdots \wedge af_m \wedge \neg af_{k+1} \wedge \neg af_n \wedge \neg P$ is unsatisfiable. Thus, the unconstrained activation literals do not affect the provability of P. This contradicts $C \setminus MCS$ is satisfiable. □

Notice in lemma **??** that the constraint system requires that a subset of the activation literals to be constrained; these faults *do not* occur. This is a subtlety that allows the model checker to choose the assignment for each unconstrained activation literal for each step of the transition system that leads to failure, and may not involve entirely simultaneous failures.

Lemma 2. *$\sigma(MCS)$ is minimal.*

Proof. Assume toward contradiction that $\sigma(MCS)$ is not minimal with regard to P. Then there exists $S \subset MCS$ such that $\sigma(S)$ is a minimal cut set of P. This implies that the corresponding constraint system $C \setminus S$ is satisfiable. This contradicts the minimality of MCS. □

A compositional proof is performed starting from the top hierarchical layer down and the analysis results are composed to form a proof of the system level property. The constraint system for a safety property is given in terms of the properties gf_i of its direct subcomponents. Each of gf_i must in turn be the safety property for a constraint system consisting of the guarantees of its subcomponents. MIVC generation is performed in this way as well; the system properties as well as any intermediate layer guarantees each have their own minimal inductive validity cores, and by extension their own minimal correction sets. The formulation described above shows the validity of this approach for a single layer; now we show that these results may be composed.

Let $C(P)$ be the constraint system associated with a safety property P and $MCS(P)$ be a minimal correction set for $C(P)$:

$$C(P) = \{gf_1, gf_2, \ldots, gf_n, \neg af_1, \neg af_2, \ldots, \neg af_n \neg P\}$$

Assume that there exists a $gf_k \in MCS(P)$ such that gf_k is an intermediate level guarantee. Then there exists a $MCS(gf_k)$. Replace $gf_k \in MCS(P)$ with $MCS(gf_k)$ and call this set $R(P)$. We define an extension to $C(P)$ in the following manner. Let $C(gf_k)$ be the associated constraint system for gf_k:

$$C(gf_k) = \{gf'_1, gf'_2, \ldots, gf'_n, \neg af'_1, \neg af'_2, \ldots, \neg af'_n \neg gf_k\}$$

and let an extended constraint system $C_{ext} = C(P) \cup (C(gf_k) \setminus \{\neg gf_k\})$. Informally, this constraint system is extended by adding in all supporting guarantees for gf_k and their constrained activation literals.

Lemma 3. *$R(P)$ is a correction set for C_{ext}, i.e., $C_{ext} \setminus R(P)$ is satisfiable.*

Proof. Assume that it is not satisfiable. Then $\exists S \subset R(P)$ such that $C_{ext} \setminus S$ is satisfiable. Let $af_i \in R(P)$ where $af_i \notin S$.

Case 1 $af_i \in MCS(P)$: by definition of MCS, af_i contributes directly to the feasibility of $\neg P$. This contradicts the definition of $MCS(P)$.

Case 2 $af_i \in MCS(gf_k)$: Then af_i becomes unconstrained in $C_{ext} \setminus R(P)$. But by definition of $MCS(gf_k)$, the unconstrained af_i contributes to the violation of gf_k and is in fact necessary based on the definition of $MCS(gf_k)$. This contradicts the definition of $MCS(P)$.

□

Lemma 4. *$R(P)$ is a minimal correction set for C_{ext}.*

Proof. Minimality: Let $S \subset R(P)$ and let $af_i \in R(P)$ where $af_i \notin S$.

If $af_i \in MCS(P)$, it is directly required for the feasibility of $\neg P$. If $af_i \in MCS(gf_k)$, it is indirectly required for the feasibility of $\neg P$ by case 2 in Lemma **??**. Since $af_i \notin S$, $C_{ext} \setminus S$ is unsatisfiable. □

The full composition of a safety property's minimal correction sets consists of replacement of every guarantee with its own minimal correction sets. Let $MCS(P) = \{af_m, \ldots, af_n, gf_i, \ldots, gf_j\}$ where $af_k \in MCS(P)$ refers directly to leaf level activation literals and $gf_k \in MCS(P)$ refers to intermediate level guarantees. Let $R(P)$ be the set consisting of all elements of $MCS(P)$ but with intermediate guarantees replaced with their minimal correction sets.

Theorem 2. *$R(P) = \{af_m, \ldots, af_n, MCS(gf_i), \ldots, MCS(gf_j)\}$ is a minimal cut set for P.*

Proof. The proof proceeds by induction on each verification layer.

◊ *Base Case*: At the leaf level of the proof, all guarantees can be directly replaced with their associated activation literals. By Lemmas **??** and **??**, $R(P)$ is a minimal cut set of P.

◊ *Induction Step*: Assume for layer n, $R(gf_k) = \{af'_m, \ldots, af'_n, MCS(gf'_i), \ldots, MCS(gf'_j)\}$ is a minimal cut set for property gf_k.
Let P be the safety property at layer $n+1$; then $MCS(P) = \{af_m, \ldots, af_n, gf_i, \ldots, gf_j\}$. Perform replacement of all $gf_x \in MCS(P)$ with $MCS(gf_x)$. Since $R(gf_k)$ is defined for an arbitrary $gf_k \in MCS(P)$, by inductive assumption and successive applications of Lemmas **??** and **??**, $R(P)$ is a minimal correction set for P. By application of Lemmas **??** and **??**, $R(P)$ is a minimal cut set of P.

□

5 Implementation

The `All_MIVCs` algorithm requires specific equations in the Lustre model to be flagged for consideration in the analysis; these we call *IVC elements*. All equations in the model can be used as IVC elements (through use of the `all_assigned` parameter) or one can specify directly which equations are of import in the analysis. In this implementation, the IVC elements are added differently depending on the layer. In the leaf architectural level, only explicitly defined faults are added to IVC elements. In middle or top layers, supporting guarantees are added. This is shown in Figure **??**.

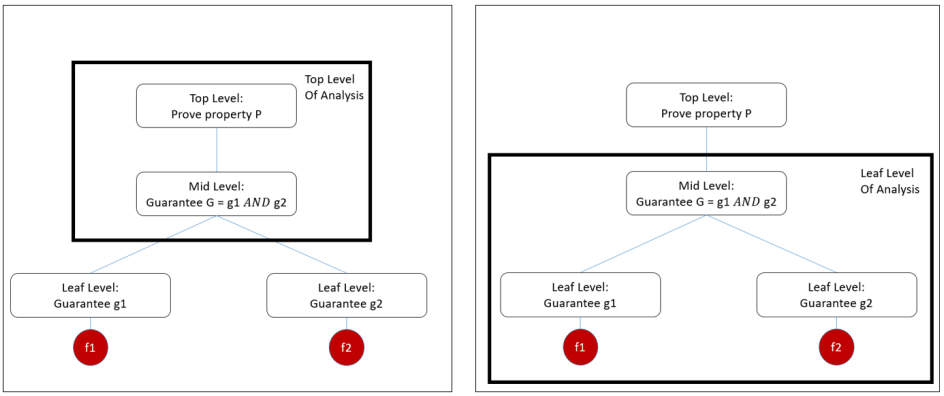

Fig. 5. Illustration of Two Layers of Analysis

The figure shows an arbitrary architecture with two analysis layers: top and leaf. The top layer analysis adds G as IVC element; the leaf layer analysis adds f_1 and f_2 as IVC elements. The first layer of analysis shows that G supports the proof of P, thus is an MIVC. The second layer of analysis shows that if both f_1 and f_2 are constrained to false, a proof is found for G.

Each explicit fault defined in the safety annex is added to the Lustre program as described in safety annex implementation [**?**,**?**] and additionally the constrained faults are added as IVC elements for leaf layer analysis. A requirement of the hitting set algorithm is that to find *all* MCSs, *all* MUSs must be known. Ghassabani et al. [**?**] showed that finding all MIVCs is as hard as model checking; and thus cannot always be easily found. It is a requirement of this analysis that all MIVCs are computed; if at any point in MIVC generation not all MIVCs can be generated, the minimal cut sets likewise cannot be computed. Once the MIVC analysis is complete for a property at a given layer,

a hitting set algorithm is used to generate the related MCSs [**?**]. Depending on the layer of analysis, the MCSs contain either faulty (or violated) guarantees or explicitly defined faults as per the IVC elements.

For a safety property P, the set of all MCSs are understood as $\vee_{i=1}^{n} MCS_i(P)$; intuitively, this means if all constraints found in any single MCS are removed from the constraint system, $\neg P$ is satisfied. For each element $gf_j \in MCS_i$, this is understood as $\wedge_{i=1}^{m} gf_i$ and speaks to the minimality of the correction set. Thus the MCSs form a disjunctive normal formula over the safety property at that layer. As the proof proceeds down the hierarchy, each of the subcomponent guarantees become a property to be proven and thus MIVCs/MCSs are generated. The composition of the MCSs consists of replacing a contract in a higher layer MCS with the disjunctive normal formula of its own MCSs. After all replacements have been made, the system property formula is converted back into disjunctive normal form.

The composition of these results is performed top down and shown in Algorithm **??**. For each guarantee found in an MCS, a replacement must be made with the guarantees own MCSs. This is done recursively until all replacements have been made (line 7, 8 of Algorithm **??**). If on the other hand there are no MCSs for a given guarantee, that guarantee is replaced by its associated fault activation literal (line 10). At the leaf level of analysis, no guarantees have associated MCSs and thus reaches the end of recursion. At that time, the formula must be converted back into disjunctive normal form to finish the translation into minimal cut sets (line 11).

Algorithm 1: Compose Results

```
1  R ← All_MCSs(P) = ∨_{i=1}^{n} MCS_i
2  where MCS_i = ∧_{j=1}^{m} gf_j
3  Function resolve(R):
4      for ∀ OR-node in R do
5          for ∀gf_j in OR-node do
6              if ∃MCS(gf_j) then
7                  R ← replace gf_j in R with All_MCSs(gf_j);
8                  resolve (All_MCSs(gf_j));
9              else
10                 R ← replace gf_j in R with af_j;
11     convert R to DNF
```

The number of replacements r that are made for a single property P are constrained by the number of minimal cut sets there are for each of the α contracts within the initial MCS. We call the set of all minimal cut sets for a contract g: $Cut(g)$. The following formula defines an upper bound on the number of replacements.

Lemma 5. *The number of replacements r is bounded by the following formula:*

$$r \leq \sum_{i=1}^{\alpha} (\prod_{j=1}^{i} |Cut(g_j)|) \tag{3}$$

Proof. Assume there exists a $g_i \in MCS(P)$. The number of replacements made between g_i and its minimal cut sets is at most $|Cut(g_i)|$. We iteratively perform this replacement for all α contracts in $MCS(P)$ and make, in the worst case, $|Cut(g_1)| \times |Cut(g_2)| \times \cdots \times |Cut(g_\alpha)|$ replacements. □

It is also important to note that the algorithm terminates.

Theorem 3. *Algorithm* **??** *terminates*

Proof. No infinite sets are generated by the `All_MIVCs` or minimal hitting set algorithms [**?**, **?**]; therefore, every MCS produced is finite. Thus, every minimal cut set of every contract is finite. Furthermore, a bound exists on the number of additional formulas that are added to $MCS(P)$: $|MCS(P)| \leq r$ by Lemma **??**. □

Given that the growth of the DNF formula can be exponential in the worst case, we implemented strategies to prune the size of the cut sets and hence the growth of these intermediate sets.

5.1 Pruning to Address Scalability

The safety annex provides the capability to specify a type of verification in what is called a *fault hypothesis statement*. These come in two forms: maximum number of faults or probabilistic analysis. Algorithm **??** is the general approach, but the implementation changes slightly depending on which form of analysis is being performed. This pruning improves performance and diminishes the problem of combinatorial explosion in the size of minimal cut sets for larger models.

Guarantees with no associated faults If a guarantee is found in a minimal correction set and no fault has been defined in the model that can violate it, this minimal correction set is pruned from the output. Future work includes collecting these pruned correction sets and presenting them to the user for further safety analysis consideration.

Max n analysis The max n hypothesis statement in the safety annex restricts the number of faults that can be independently active simultaneously and verification can be run with this restriction present. In terms of minimal cut sets, this statement restricts the cardinality of minimal cut sets generated to n. If the number of elements in an MCS exceeds the threshold set in the hypothesis statement, that MCS is eliminated from consideration.

Probabilistic analysis pruning This type of hypothesis statement restricts the cut sets by use of a probabilistic threshold. Any cut sets with combined probability higher than the given probabilistic threshold are removed from consideration. The allowable combinations of faults are calculated before the compositional algorithm begins; this allows for a pruning of minimal correction sets during the transformation. If the faults within an MCS are not a subset of any allowable combination, that MCS is pruned from the formula.

6 Related Work

Minimal cut sets generated by monolithic analysis look only at explicitly defined faults throughout the architecture and attempt through various techniques to find the minimal violating set for a particular property. We outline some of the common monolithic approaches to minimal cut set generation in this section.

The representation of Boolean formulae as Binary Decision Diagrams (BDDs) was first formalized in the mid 1980s [**?**] and were extended to the representation of fault trees not many years later [**?**]. After this formalization, the BDD approach to FTA provided a new approach to safety analysis. The model is constructed using a BDD, then a second BDD - usually slightly restructured - is used to encode MinCutSets [**?**]. Unfortunately, due to the structure of BDDs, the worst case is exponential in size in terms of the number of variables [**?**, **?**, **?**]. In industrial sized systems, this is not realistically useful.

SAT based computation was then introduced to address scalability problems in the BDD approach; initially it was used as a preprocessing step to simplify the decision diagram [**?**], but later extended to allow for all MinCutSet processing and generation without the use of BDDs [**?**]. Since then, numerous safety related research groups have focused on leveraging the power of model checking in the problems of safety assessment [**?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**].

Bozzano et al. formulated a Bounded Model Checking (BMC) approach to the problem by successively approximating the cut set generation and computations to allow for an "anytime approximation" in cases when the cut sets were simply too large and numerous to find [**?**, **?**]. These algorithms are implemented in xSAP [**?**] and COMPASS [**?**].

The model based safety assessment tool AltaRica 3.0 [**?**] performs a series of processing to transform the model into a reachability graph and then compile to Boolean formula in order to compute the MinCutSets [**?**]. Other tools such as HiP-HOPS [**?**] have implemented algorithms that follow the failure propagations in the model and collect information about safety related dependencies and hazards. The Safety Analysis Modeling Language (SAML) [**?**] provides a safety specific modeling language that can be translated into a number of input languages for model checkers in order to provide model checking support for MinCutSet generation.

To our knowledge, a fully compositional approach to calculating minimal cut sets has not been introduced.

7 Conclusion and Future Work

We have developed a way to leverage recent research in model checking techniques in order to generate minimal cut sets in a compositional fashion. Using the idea of Minimal Inductive Validity Cores (MIVCs), which are the minimal model elements necessary for a proof of a safety property, we are able to restate the safety property as a top level event and provide faults of components and their contracts as model elements to the `All_MIVCs` algorithm which provides all minimal IVCs that pertain to this property. These are used to generate minimal cut sets. Future work includes leveraging the system information embedded in this approach to generate hierarchical fault trees as well as perform scalability studies that compare this approach with other non-compositional approaches to minimal cut set generation. To access the algorithm implementation, Safety Annex users manual, or example models, see the repository [**?**].

Acknowledgments. This research was funded by NASA contract NNL16AB07T and the University of Minnesota College of Science and Engineering Graduate Fellowship.