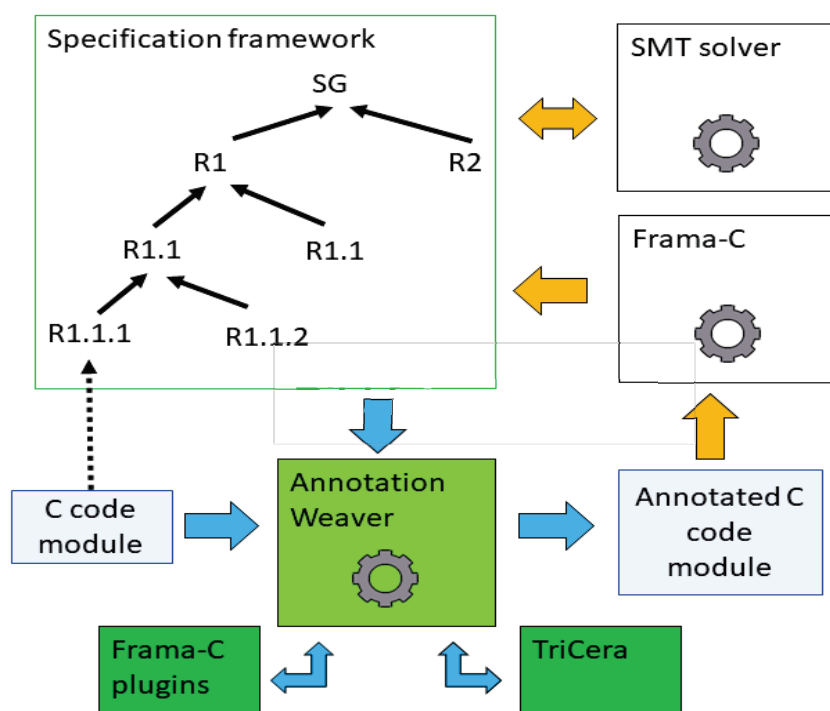


AVerT2

Automatiserad verifiering och testning

Publik rapport



Författare: [Dilian Gurov](#)

Datum: [2024-12-27](#)

Projekt inom [Delprogrammet Elektronik, mjukvara och kommunikation](#)



Fordonsstrategisk
Forskning och
Innovation

Innehållsförteckning

1 Sammanfattning.....	3
2 Executive summary in English.....	3
3 Bakgrund.....	3
4 Syfte, forskningsfrågor och metod.....	3
5 Mål.....	3
6 Resultat och måluppfyllelse.....	3
7 Spridning och publicering.....	4
7.1 Kunskaps- och resultatspridning.....	4
7.2 Publikationer.....	4
8 Slutsatser och fortsatt forskning.....	4
9 Deltagande parter och kontaktpersoner.....	4

Kort om FFI

FFI, Fordonsstrategisk forskning och innovation, är ett samarbetsprogram mellan staten och fordonsindustrin som sedan 2009 finansierar forskning och innovation inom vägtransporter.

Läs mer på www.ffisweden.se

1 Sammanfattning

AVerT2-projektet hade **fem arbetspaket**:

- WP1: Projektkoordinering
- WP2: Kravspecifikation och dekomposition
- WP3: Deduktiv verifiering av C-kod
- WP4: Verifiering och testning av Simulink-modeller
- WP5: Praktisk utvärdering och tekniköverföring

I WP2 levererade vi tre **formella specifikationsspråk**: (1) den temporala logiken LTL för att specificera systemnivåns temporala egenskaper, (2) den reelltidtemporal logiken MITL för att specificera systemnivåns reelltidsegenskaper och (3) specifikationsspråket ACSL för att specificera funktionskontrakt på programnivå i C. Vi levererade också ett **ramverk och verktyg för kravdekomposition** i form av ett bevisystem, som mekaniserades i den interaktiva teorembevisaren HOL4.

I WP3 levererade vi en metod för verifiering av flyttalsaritmetik som involverar deduktiv verifiering, där flyttal abstraheras till reella tal, kombinerat med en ny dynamisk analysmetod som beräknar konfidensnivåer i termer av Chernoff-gränser. Vi levererade också ett verktyg som infogar funktionskontrakt och hjälpannoteringar för Frama-C, i form av **verktyget AutoDeduct**, som släpptes offentligt på GitHub.

Arbetet i WP4 bedömdes vara för svårt att förverkliga inom projektets givna tidsram. Vi valde därför att fokusera på de andra arbetspaketen, eftersom de bedömdes vara viktigare för projektets industriella framgång.

I WP5 levererade vi en **fallstudie i kravspecifikation och dekomposition** baserad på ett bränslenivåvisningssystem, och en **fallstudie i C-kodsverifiering** baserad på en styrmodul som körs i Scania-lastbilar. Den andra fallstudien publicerades i form av en konferensartikel (se nedan vår RE'24-artikel) som bl.a. illustrerar utmaningarna med att formalisera systemnivåkrav och deras efterföljande översättning till funktionskontrakt på programnivå.

Akademiskt sett resulterade projektet i en **doktorsavhandling** (där arbetet hade påbörjats i föregångarprojektet AVerT), ett annat doktorandprojekt halvvägs, och ett antal **examensprojekt** som genomfördes på Scania och examinerades vid KTH. När det gäller **publikationer** var projektet mycket framgångsrikt. En särskild höjdpunkt är CAV'23-artikeln, som tilldelades **CAV Distinguished Paper Award 2023** (se nedan).

AVerT2-projektet hade en mycket god synergi med avancerade kursen **DD2452 Formella metoder**, som Dilian Gurov undervisar vid KTH. Å ena sidan har kursens innehåll påverkats av teknikerna och fallstudierna som utvecklades i projektet. Å andra sidan har flera studenter som tog kursen senare genomfört sina examensprojekt på Scania. Våra två doktorander är bland dem.

Slutligen vill vi lyfta fram den **Scania Open Day 2024**, som ägde rum på Scania den 24 oktober 2024 och med vilken vi avslutade AVerT2-projektet och gick vidare till uppföljningsprojektet FormAI. Vi hade flera framstående forskare från hela Sverige som deltog med föredrag och i en slutpaneldiskussion. Workshopen besöktes också av många ingenjörer från Scania och till och med av några KTH-studenter från kursen i Formella metoder.

2 Executive summary in English

The AVerT2 project had **5 work packages**:

- WP1 Project coordination
- WP2 Requirements Specification and Decomposition
- WP3 Deductive Verification of C Code
- WP4 Verification and Testing of Simulink Models
- WP5 Practical Evaluation and Technology Transfer

In WP2, our tasks were to deliver:

1. A requirements language for specifying C-code modules.
2. A requirements language for specifying Simulink/Stateflow Models.
3. A framework and tool for requirements decomposition.

The first deliverable was realised by means of **three formal specification languages**: (1) the temporal logic LTL for specifying system-level temporal properties, (2) the real-time temporal logic MITL for specifying system-level real-time properties, and (3) the specification language ACSL for specifying program-level C-function contracts. The algorithmic translation of requirements in the first two languages to contracts in the third language could not be fully solved within this project, but will be completed in our follow-up work. The second deliverable was deemed unnecessary. The third deliverable was realised in the form of **a proof system**, which was **mechanised in the HOL4 interactive theorem prover**.

In WP3, our tasks were to deliver:

1. A report on method for verification of floating-point arithmetic.
2. A report on preliminary results on verifying C++ code.
3. A tool that inserts function contracts and auxiliary annotations for Frama-C.

The first deliverable was realised as a research paper. Our method for verification of floating-point arithmetic involves deductive verification, where floating-point numbers are abstracted into real numbers, combined with **a novel dynamic analysis technique** that computes confidence levels in terms of Chernoff bounds. The second deliverable was not realised due to time constraints. The third deliverable was realised in the form of **the AutoDeduct tool**, which was **publicly released on GitHub**. It currently handles almost all constructs of the C programming language. The main renaming construct are stack pointers, which will be implemented in the next version of the tool.

The work in WP4 was deemed too difficult to realise in the given time frame of the project. We therefore decided to focus on the other work packages, since they were deemed more important for the industrial success of the project.

In WP5, our tasks were to deliver:

1. A case study on requirements specification and decomposition.
2. A case study on C code Verification.
3. A case study on Simulink/Stateflow Verification.

The first deliverable was realised as a research paper. The case study was a **fuel level display** system. The second deliverable was realised as a conference publication (see below our RE'24 paper). The case study here was a **steering module** running in Scania trucks. It also illustrated the challenges in formalising system-level requirements and their subsequent translation to

program-level C-function contracts (as described above for WP2). The third deliverable was not realised, for reasons explained above.

Academically, the project resulted in **one PhD thesis defended** (where work had started in the predecessor project AVerT), another PhD project halfway, and **a number of MSc projects** conducted at Scania and defended at KTH. In terms of **publications**, the project was very successful. As a particular highlight one should mention the CAV'23 paper, which was awarded a **CAV Distinguished Paper Award 2023** (see below). The AVerT2 project received considerable visibility in the **scientific community** working on Applied Formal Methods, and was represented at several research venues, including a Dagstuhl Seminar in 2023, a Lorentz center workshop in 2024, and the IsoLA symposia in 2022 and 2024.

The AVerT2 project had a very good mutual synergy with the advanced-level course **DD2452 Formal Methods** that Dilian Gurov teaches at KTH. On one hand, its contents have been influenced by the techniques and case studies developed in the project. On the other hand, several students who took the course later conducted their MSc project at Scania. Our two PhD students are among them.

Finally, we should highlight the **Scania Open Day 2024**, which took place at Scania on October 24, 2024, with which we concluded the AVerT2 project and ventured into the follow-up FormAI project. We had several distinguished scientists from around Sweden who participated with talks and in a final panel discussion. The workshop was also attended by numerous engineers from Scania, and even by some KTH students from the Formal Methods course.

3 Bakgrund

AverT var ett tidigare projekt med syftet att skapa formella metoder och verktyg för industriell utveckling av säkerhetskritiska system. AVerT2 har varit ett fortsättningsprojekt på projektet AverT. Som FFI:s färdplansdokument påpekar, går vi mot en transportmodell där autonoma fordon, sammankopplade med varandra och med transportinfrastrukturen, kommer att bilda ett komplext system-av-system, där fordonsmjukvaran behandlar ett kontinuerligt flöde av inkommande information, är utbyggbart och självanpassande och garanterar höga nivåer av säkerhet. Alla dessa utvecklingar resulterar i en ökning av mjukvarans och systemens säkerhet. Särskilt, färdplanerna pekar på behovet av att hantera olika systemnivåer, att ge stöd för agil mjukvaruutveckling och kontinuerlig integration, samt uppfylla högre krav på funktionell säkerhet och tillhandahålla verktygsstöd för automatisering av själva V&V-processen.

4 Syfte, forskningsfrågor och metod

Syftet med AVerT2 har varit att lösa tre huvudproblem som uppstod från det tidigare AverT projektet: (1) problemet med att härleda från systemkraven på högsta nivån och på ett halvautomatiskt sätt krav på de individuella mjukvarumodulerna, (2) problemet med att helt automatisera V&V-processen och (3) problemet med att integrera den nya V&V-tekniken i de befintliga processerna för mjukvaruutveckling. Det första problemet skulle lösas genom att utveckla kravnedbrytningsalgoritmer, med hjälp av en formell ramverk erhållet i AVerT för hierarkiska programvaruarkitekturbeskrivningar som omfattar kraven. Det andra problemet skulle byggas på resultat som erhållits i AVerT på automatiserad generering av kommentarer från modulkrav och utöka dessa till att omfatta mer avancerade språkkonstruktioner (som loopar) och datatyper (som flyttal). Det tredje problemet skulle hanteras genom att genomföra fallstudier i nära samarbete med ingenjörer på Scania.

5 Mål

Projektets mål var att automatisera till en hög grad V&V-processen genom att skapa verktyg som:

1. tar systemkrav, översätter dem till krav på mjukvarumodulnivå, och
2. verifierar mjukvaran helt automatiskt.

Utvecklade verktygen skulle sedan:

1. utvärderas på tre stora fallstudier och
2. integreras i de befintliga processerna för mjukvaruutveckling.

6 Resultat och måluppfyllelse

De viktigaste resultaten från AVerT2 är:

1. Formella specifikationsspråk för att specificera systemnivåns temporala egenskaper och funktionskontrakt på programnivå i C.
2. Ett ramverk och verktyg för kravdekomposition.
3. Verktyget AutoDeduct som infogar funktionskontrakt och hjälpannoteringar för Frama-C och därmed automatiserar mjukvaruverifikationsprocessen nästan helt.
4. En fallstudie i kravspecifikation och dekomposition baserad på ett bränslenivåvisningssystem, och en fallstudie i C-kodsverifiering baserad på en styrmodul som körs i Scania-lastbilar.
5. En doktorsavhandling, ett antal publikationer, och en öppen workshop.

Projektet har utvecklat **kompetens och teknologi** för automatiserad formell verifiering av säkerhetskritisk inbyggd programvara. Tekniken är tillräckligt mogen för att integreras i mjukvaruutvecklingsprocessen. För närvarande verkar det främsta hindret för detta vara bristen på välskrivna krav. Framtida arbete med fokus på utveckling av en disciplin för att ta fram sådana krav kommer att vara den möjliggörande faktorn för vår teknik, men detta kommer också att behöva backas upp med beslut på ledningsnivå i fordonsföretag som Scania.

En andra, långtgående effekt av vårt projekt är att våra tekniker och verktyg för automatiserad formell verifiering av mjukvara gör att de kan kombineras med tekniker för **generativ AI**. Den förväntade effekten av detta är möjligheten att producera korrekt programvara med mycket mindre mänsklig insats än i nuvarande praxis. Denna idé ledde till ett uppföljningsprojekt av AVerT2, kallat FormAI, som också finansieras av Vinnova.

En tredje effekt av vårt projekt är den påverkan det har haft på utvecklingen av verifieringsverktyg av andra forskargrupper, framför allt verktyget TriCera, utvecklat vid Uppsala universitet.

Sammanfattningsvis är vi nöjda med resultaten av AVerT2-projektet, både teoretiskt och praktiskt. Men vi inser också att fordonsföretag som Scania måste fatta beslut på ledningsnivå om att ta fram högkvalitativa krav på fordons- och mjukvarunivå. Det kanske mest anmärkningsvärda praktiska resultatet av projektet är att vi lyckades öka medvetenheten inom Scania om fördelarna med att använda formella metoder för att höja förtroendenivån för det korrekta beteendet hos inbyggd mjukvara. Detta är viktigt eftersom det kan leda till beslut på ledningsnivå om processen för att ta fram krav och deras användning för testning och formell verifiering.

A. Bidrag till FFI:s mål:

Öka den svenska kapaciteten för forskning och innovation och därigenom säkerställa konkurrenskraft och arbetstillfällen inom fordonsindustrin. Verktøygen som levererades av AVerT2-projektet är öppen källkod och är tillgängliga för den svenska fordonsindustrin. Detta gör det möjligt för fordonstillverkare, som Scania och Volvo Cars, att införliva dessa verktyg i sina verktygskedjor och V&V-processer, och därmed stärka sina fordons konkurrenskraft genom att ge ännu högre garantier för funktionssäkerhet.

Utveckla internationellt sammanlänkade och konkurrenskraftiga forsknings- och innovationsmiljöer i Sverige. AVerT2-projektet har genomförts inom ramen för flera andra relaterade projekt, varav tre är europeiska. Detta har främjat detta mål.

Främja tvärindustriellt samarbete. Medan verktygsutveckling och fallstudier har varit till övervägande del utförda på Scania, omfattade spridningen öppna workshops och demodagar som var öppna för svensk fordonsindustri. I synnerhet har vi bjudit in utvecklare från Volvo Cars som redan hade uttryckt intresse för de potentiella resultaten av AVerT2. Detta, tillsammans med verktygens öppna källkodskaraktär, bidrar till ett tvärindustriellt samarbete.

Främja samarbetet mellan industri, universitet och högskolor. AVerT2 är unikt genom att det omfattar ett antal forskare som är anställda både på Scania och KTH. Projektledaren, prof. Mattias Nyberg från Scania, är adjungerad professor på KTH, medan alla återstående deltagare från KTH har tillbringat tid på Scania i samband med olika föregångarprojekt till AVerT2. Projektet har också främjat samarbetet genom deltagarnas industriella och akademiska nätverk. Särskilt kan vi nämna KTH Innovativa Centrum för inbyggda system (ICES) (www.ices.kth.se/), som involverar medlemmar från flera forskargrupper vid KTH och har ett nära samarbete med en lång rad industriella partners.

B. Bidrag till delprogrammets mål:

Simulering och validering på olika systemnivåer. Hela konceptet med AVerT2 var att möjliggöra den formella verifieringen i kombination med testning på olika systemnivåer, fångad av en hierarkisk arkitekturmodell. Nyckeln till att koppla ihop de olika nivåerna är ramverket för kompositionsresonemang som har utvecklats inom AVerT2. Alla aspekter av detta stöds av automatiserade verktyg som syftar till att dölja det mesta av komplexiteten bakom tillvägagångssättet.

Stöd för tidig och kontinuerlig integration. Kompositionsaspekten av vårt tillvägagångssätt tillåter modularisering av V&V (i vårt fall, formell verifiering och testning). Detta innebär att V&V kan utföras inkrementellt, och V&V-resultat (d.v.s. feedback från verifiering och testning) kan lagras, spåras och återanvändas. Detta är nyckeln till agil mjukvaruutveckling med integrerad V&V, eftersom det gör att små ändringar i kodbasen snabbt och fullständigt kan verifieras.

Funktionell säkerhet och certifiering. Målet för AVerT2 är funktionssäkerhet för enskilda fordon såväl som för fordon inom en transportinfrastruktur. Vi har inriktat oss på ISO 26262-standarderna, både med avseende på nedbrytningen av säkerhetskrav och till det faktiska sättet för deras V&V. Standarderna rekommenderar formell verifiering (bland andra tekniker). AVerT2 bidrog till att öka mognaden för formell verifiering inom fordonsindustrin. De kommande standarderna för autonoma fordon förväntas följa denna linje.

Mjukvarubaserad V&V-metodik. Den långsiktiga visionen för AVerT2 är en helautomatiserad V&V-metodik som är integrerad i mjukvaruutvecklingsprocessen. AVerT2 tåg ett synligt steg mot detta mål genom att utveckla verktyg som automatiserar några av de mest arbetsintensiva och kompetenskrävande uppgifterna för traditionell formell verifiering av programvara.

7 Spridning och publicering

7.1 Kunskaps- och resultatspridning

Hur har/planeras projektresultatet att användas och spridas?	Markera med X	Kommentar
Öka kunskapen inom området	x	Genom forskning och industriella fallstudier.
Föras vidare till andra avancerade tekniska utvecklingsprojekt	x	Förs vidare till fortsättningsprojektet FormAI och därigenom också till Scania och annan fordonsindustri.
Föras vidare till produktutvecklingsprojekt	x	Den stora fallstudien i AVerT2 visar att vårt nya verktyg AutoDeduct har potential att föras i produktion.
Introduceras på marknaden	x	Ja, se ovan.
Användas i utredningar/regelverk/tillståndsärenden/ politiska beslut		Inte direkt, men genom att AVerT2 visar att formell verifiering fungerar bra industriellt, kan standarder inom t.ex. funktionell säkerhet komma att påverkas, eftersom sådana standarder löpande uppdateras för ta in nya tekniska framsteg.

7.2 Publikationer

Automated Deductive Verification of Safety-Critical Embedded Software

Christian Lidström

KTH PhD Thesis, 2024-03-15

AutoDeduct: Automated Deductive Verification

Jesper Amilon, Christian Lidström, Gustav Ung, Dilian Gurov, and Mattias Nyberg

<https://github.com/rse-verification/auto-deduct-toolchain>

Tool release, 2024

Post-Hoc Formal Verification of Automotive Software with Informal Requirements: An Experience Report

Gustav Ung, Jesper Amilon, Dilian Gurov, Christian Lidström, Mattias Nyberg, and Karl Palmeskog

In Proceedings of: *Requirements Engineering 2024*

IEEE, pp. 287-298, 2024

Deductively Verified Program Models for Software Model Checking

Jesper Amilon, and Dilian Gurov

In Proceedings of: *Leveraging Applications of Formal Methods, Verification and Validation 2024*

Lecture Notes in Computer Science, vol. 15221, pp. 8-25, 2024

An Exercise in Mind Reading: Automatic Contract Inference for Frama-C

Jesper Amilon, Zafer Esen, Dilian Gurov, Christian Lidström, and Philipp Rümmer

In book: *Guide to Software Verification with Frama-C*, 2024

Automatic Program Instrumentation for Automatic Verification

Jesper Amilon, Zafer Esen, Dilian Gurov, Christian Lidström, and Philipp Rümmer

In Proceedings of: *Computer Aided Verification 2023*

Lecture Notes in Computer Science, vol. 13966, pp. 281-304, 2023

Recipient of a **CAV Distinguished Paper Award 2023**

Trace-based Deductive Verification

Richard Bubel, Dilian Gurov, Reiner Hähnle, and Marco Scaletta
In Proceedings of: *Logic Programming and Automated Reasoning 2023*
EPIc Series in Computing, vol. 94, pp. 7395, 2023

Contract Based Embedded Software Design

Christian Lidström, and Dilian Gurov
In Proceedings of: *Theoretical Aspects of Software Engineering 2023*
Lecture Notes in Computer Science, vol. 13931, pp. 77-94, 2023

Deductive Verification Based Abstraction for Software Model Checking

Jesper Amilon, Christian Lidström, and Dilian Gurov
In Proceedings of: *Leveraging Applications of Formal Methods, Verification and Validation 2022*
Lecture Notes in Computer Science, vol. 13701, pp. 7-28, 2022

Bounded Invariant Checking for Stateflow

Predrag Filipovikj, Gustav Ung, Dilian Gurov, and Mattias Nyberg
In Proceedings of: *Formal Methods for Autonomous Systems 2022*
EPTCS 371, pp. 38–52, 2022

Alice in Wineland: A Fairy Tale with Contracts

Dilian Gurov, Christian Lidström, and Philipp Rümmer
Lecture Notes in Computer Science, vol. 13360, pp. 229-242, 2022

8 Slutsatser och fortsatt forskning

Projektet har utvecklat **kompetens och teknologi** för automatiserad formell verifiering av säkerhetskritisk inbyggd programvara. Tekniken är tillräckligt mogen för att integreras i mjukvaruutvecklingsprocesser. För närvarande verkar det främsta hindret för detta vara bristen på välskrivna krav inom industrin. Framtida arbete med fokus på utveckling av en disciplin för att ta fram sådana krav kommer att vara den möjliggörande faktorn för vår teknik, men detta kommer också att behöva backas upp med beslut på ledningsnivå i företagen.

En andra, långtgående effekt av vårt projekt är att våra tekniker och verktyg för automatiserad formell verifiering av mjukvara möjliggör att de kan kombineras med tekniker för generativ AI. Den förväntade effekten av detta är möjligheten att producera korrekt programvara med mycket mindre mänsklig insats än i nuvarande praxis. Samarbetet mellan KTH och Scania fortsätter nu i denna riktning inom det **nya projektet FormAI**, som också finansieras av Vinnova.

9 Deltagande parter och kontaktpersoner

Koordinerande projektpart har varit **Kungliga Tekniska Högskolan**, Skolan för elektroteknik och datavetenskap, med projektledare och kontaktperson **Prof. Dilian Gurov**.



Projektparter har varit **Scania CV AB**, med kontaktperson **Adj. Prof. Mattias Nyberg**.

