C2AADL_Reverse: A Model-Driven Reverse Engineering Approach to Development and Verification of Safety-critical Software

Zhibin Yang¹, Zhikai Qiu¹, Yong Zhou¹, Zhiqiu Huang¹, Jean-Paul Bodeveix², Mamoun Filali²
¹Nanjing University of Aeronautics and Astronautics (NUAA), China ²IRIT-Université de Toulouse, France

June 17, 2022





Agenda

- Background & Motivation
- Overview of the Main Contributions
- C2AADL_Reverse Approach
- Validation and Verification Approach for C2AADL_Reverse
- Prototype Tool
- Case Studies
- Conclusion and Future Work

✓ RE for Safety-critical software

- Long-term maintenance (20-30 years or more)
- Complex challenge: The SCS communities have been struggling to manage and maintain their legacy software.
- FAA: Reverse Engineering (RE) has been increasingly used.

DOT/FAA/TC-15/27 Federal Aviation Administration William J. Hughes Technical Center Aviation Research Division Altarric City International Airport New Jersey 08405	Reverse Engineering for Software and Digital Systems	Certification Authorities Software Team (CAST)
		Position Paper CAST-18
		Reverse Engineering in Certification Projects
		Completed June 2003
		(Rev 1)
	February 2016 Final Report	
	This document is available to the U.S. public through the National Technical Information Service (NTIS), Springfield, Virginia 22161. This document is also available from the Federal Aviation Administration William J. Hughes Technical Center at actilibrary.Ic.faa.gov.	<i>NOTE:</i> This position paper has been coordinated among the software specialists of certification authorities from the United States, Europe, and Canada. However, it does not constitute official policy or guidance from any of the authorities. This document is provided for educational and
	U.S. Department of Transportation Federal Aviation Administration	informational purpose: only and thould be discussed with the appropriate certification authority when considering for actual projects.

✓ Reverse Engineering (RE)

 A process to build more abstract representations (such as architectural models, or use cases, etc) from a lowlevel representation of a (software) system (such as source code, or execution traces)

✓ The main objective of RE:

 Provide a better understanding of the software system's current state, which can be used to correct (e.g. fix bugs), update (e.g. alignment with updated user requirements), upgrade (e.g. add new capabilities), or even completely re-engineer the system under study.

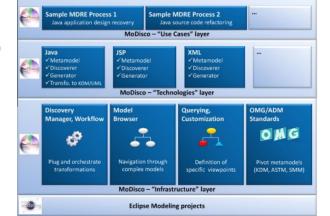
4

Generally, RE is a time-consuming and error-prone process.

- ✓ Model-driven Reverse Engineering (MDRE) [Spencer Rugaber, 2004]
 - The application of model driven engineering (MDE) principles and techniques to RE
 - Meta-model, model-based views on legacy systems
 - Raising the degree of automatic process through model transformations

✓ Related work [Claudia Raibulet,2017][André Pascal, 2019][Hugo Brunelière 2014]

- General solutions
 - MoDisco (model discovery and model understanding, JAVA/JSP/XML -> UML2)
- Specific solutions (desktop/business/...)
 - Src2MOF (Java -> UML)
 - BREX (Java -> business rules)
 - ITACG (C-> UML)
 - Wang et al. , STOOD (C -> AADL)



- The characteristics of MDRE in desktop or business domains. [Hugo Brunelière 2014]
 - Genericity
 - Extensibility
 - Partial/Full coverage
 - Direct (re) use and integration
 - Automation

- The characteristics of MDRE in the safetycritical domain.
 - Genericity
 - Extensibility
 - Partial/Full coverage
 - Architecture
 - Functional Behavior
 - Runtime
 - Direct (re) use and integration
 - Automation
 - validation of RE process

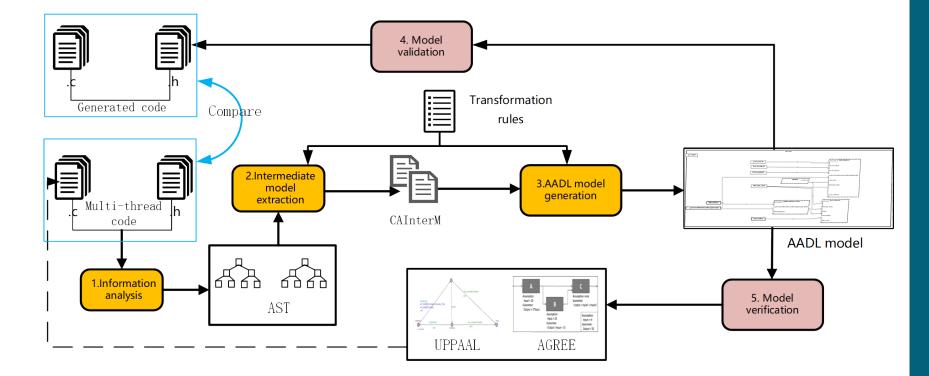
6

 Verification of resulted models

Main contributions

- We propose C2AADL_Reverse, a MDRE approach for safetycritical software development and verification:
 - **Domain**: SCSs need the modeling of architecture, functional behaviors and runtime.
 - Source artifacts: multi-task C source code conforms with the "coding rules" in the aerospace industry.
 - Target models: compared with the modeling languages used in the existing works of MDRE such as UML, AADL (Architecture Analysis and Design Language) is a powerful modeling language for complex embedded system, which provides a unified formalism for the modeling of architecture, functional behaviors, and runtime.
- ✓ Validation and verification approach for C2AADL_Reverse
- ✓ Prototype tool
- ✓ Industry case studies

Main contributions



C2AADL_Reverse approach: step 1- step 3 V&V of C2AADL_Reverse: step 4, step 5

✓ The features of the source code

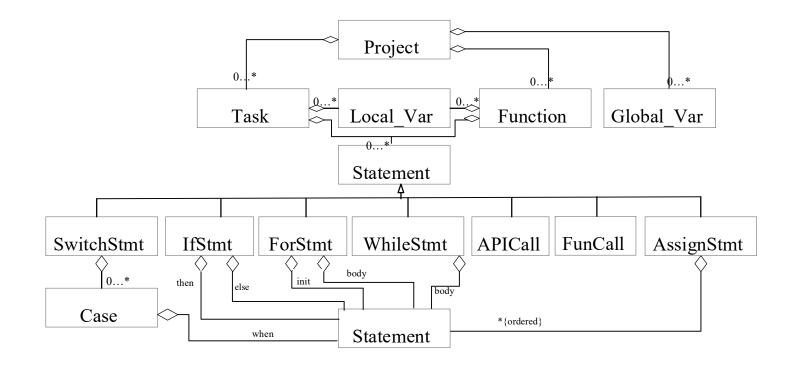
- Our Case: the code is structured, i.e., conforms with the coding/programming rules in aerospace industry
 - Multi-tasks
 - Strict development patterns, for example with clear separation of communications, data types, components types, etc.
 - Safety programming: Cyclomatic complexity <10, LOC of each function <100, ...
- The code is not structured, then it needs pre-processing (code annotations written manually)

It makes that the RE from C to AADL is feasible



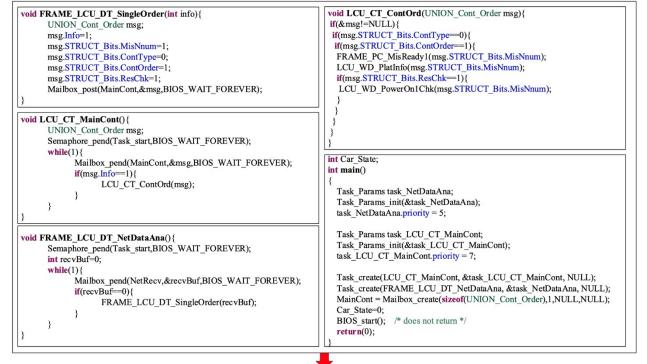
Code analysis to build code structure model

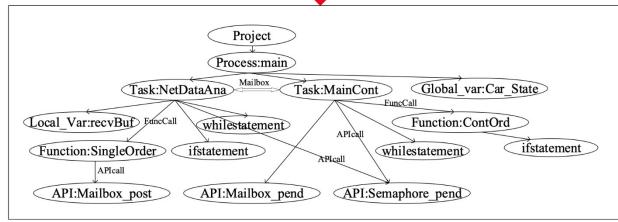
Simplified meta-mode of multi-task C code structure



*Statement is duplicated for readability

Example:





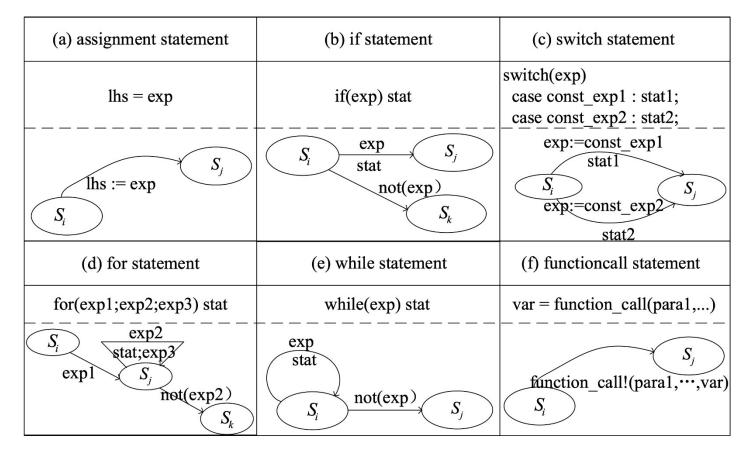
Structure transformation

- Plain code -> namespaces, source code files -> create hierarchy
- Data types -> data components
- function definition -> subprogram component
- Task definition -> thread component

C language	AADL
int a; int *a	Base_Types::Integer;
	requires data access Base_Types:Integer
char a; char *a	Base_Types::Character
	requires data access Base_Types: Character
bool a; bool *a	Base_Types::Boolean
	requires data access Base_Types:Boolean
float a; float *a;	Base_Types::Float
	requires data access Base_Types:Float
struct_name a;	User_Defined::struct_name.impl;
struct_name *a;	requires data access User_Defined::struct_name.impl;
	Base_Types::Integer_32;
	Base_Types::Unsigned_32;
	data Integer_32 extends Integer
	properties
signed int a;	Data Model::Number Representation ⇒ Signed;
unsigned int b;	end Integer 32;
	data Unsigned 32 extends Integer
	properties
	Data_Model::Number_Representation => Unsigned;
	end Unsigned_32;
struct dataname{	data dataname
type_spec var_name;	properties
};	Data_Model::Data_Representation⇒
	(Struct/Union/Enum);
enum;	end dataname
	data implementation dataname.impl
union;	subcomponents
	var_name: data package_name::type_spec
	end dataname.impl
function definition	subprogram component
task structure	thread component

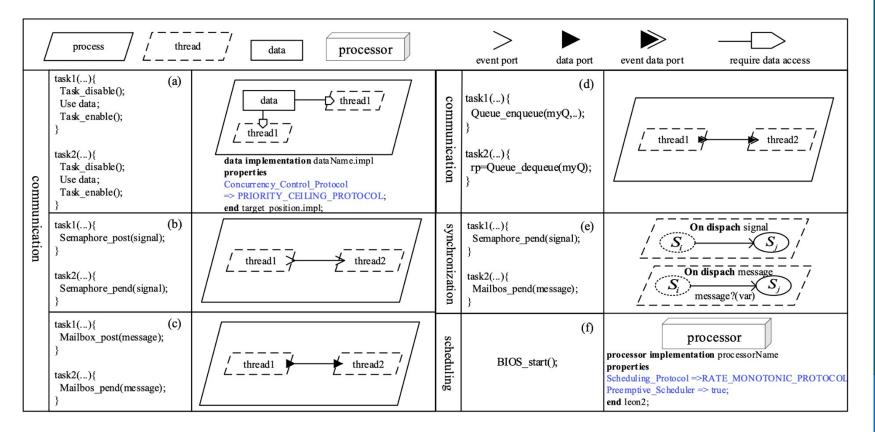
Behavior annex transformation

 Modeling and verification feasibility: cyclomatic complexity <10, LOC of each function <100, ...



✓ Run-time information transformation

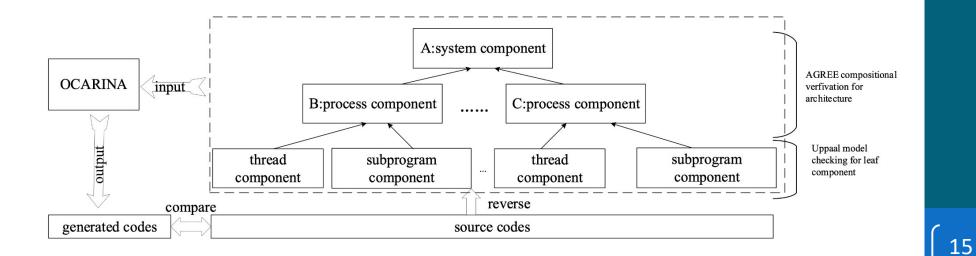
• The APIs of OS or runtime execution platform



*Without loss of generality, we consider TI SYS/BIOS Real-time Operating System (SYS/BIOS) which is broadly used in the aerospace domain.

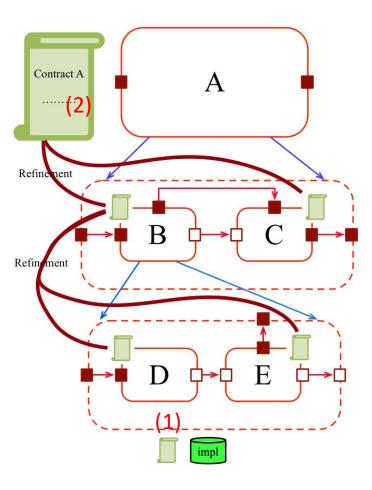
✓ Global view

- Validation of the RE process by using a comparison between twoversions code (see case study)
- Compositional verification of the architecture model
- Verification of the leaf components



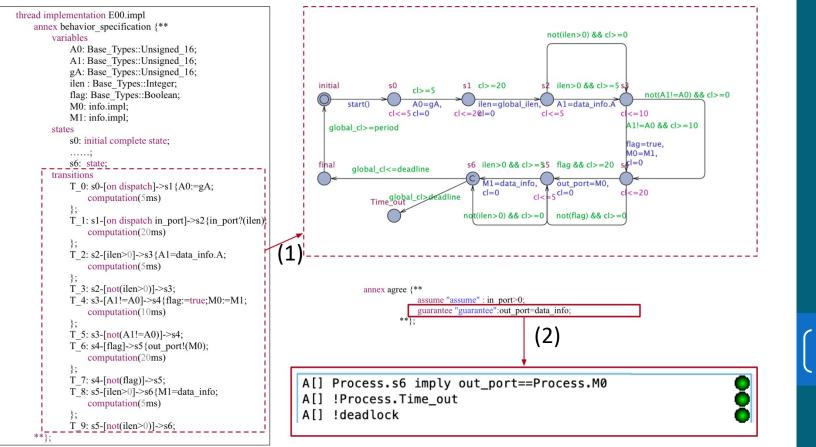
The principle of compositional verification

- The state-explosion problem
- The verification of a composite system is reduced to the verification of its parts.
- Requirements are decomposed and formalized into contracts and subcontracts: <Assume, Guarantee>
- AADL AGREE annex and tool



Verification of leaf components with UPPAAL

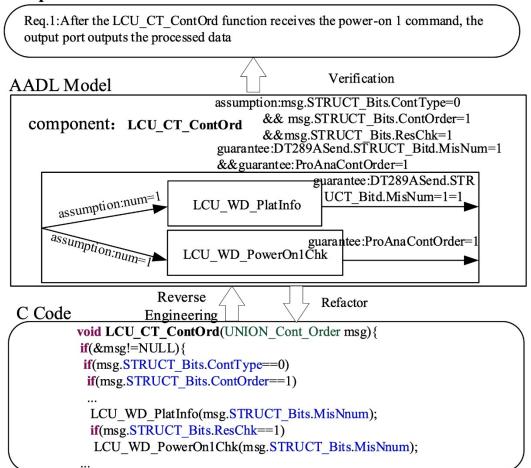
- Why we use UPPAAL? It has been used in industry
- Properties: safety, liveness, no deadlock (Component-level contracts: Assume-> initialization function, Guarantee -> TCTL)



Compositional verification of AADL architecture model

System-level properties (system-level contracts)

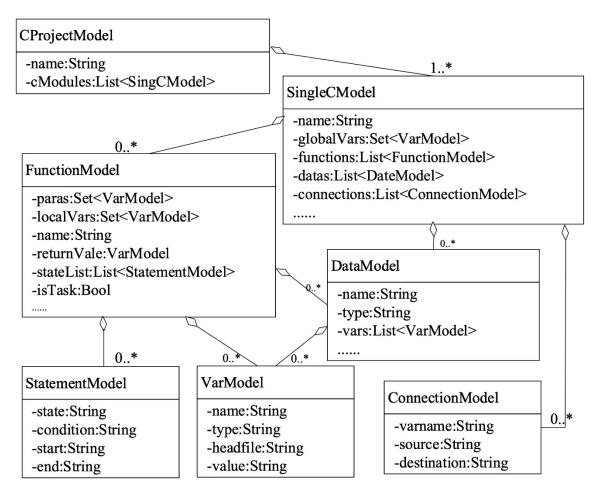
Specification



T3: Prototype tool

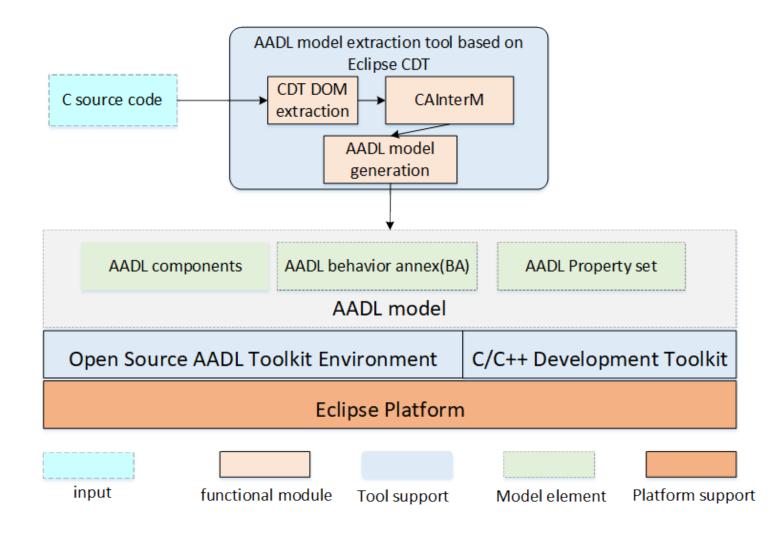
✓Intermediate model

• Consider the extensibility



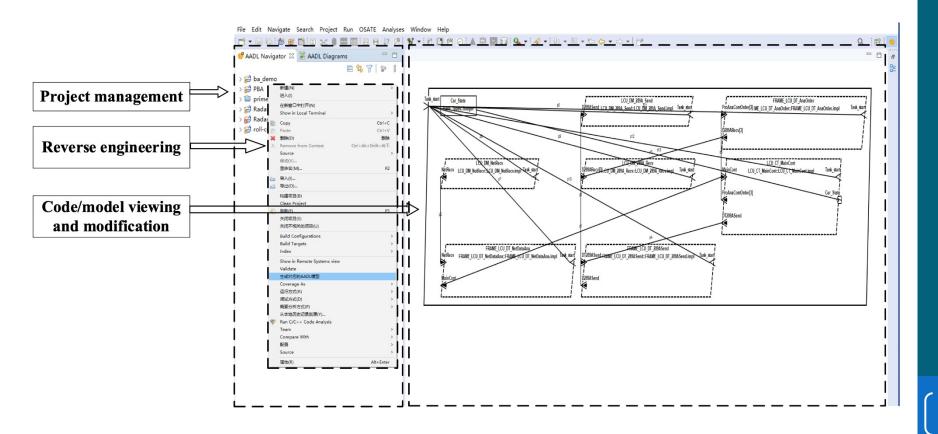
T3: Prototype tool

✓ Implementation of the tool

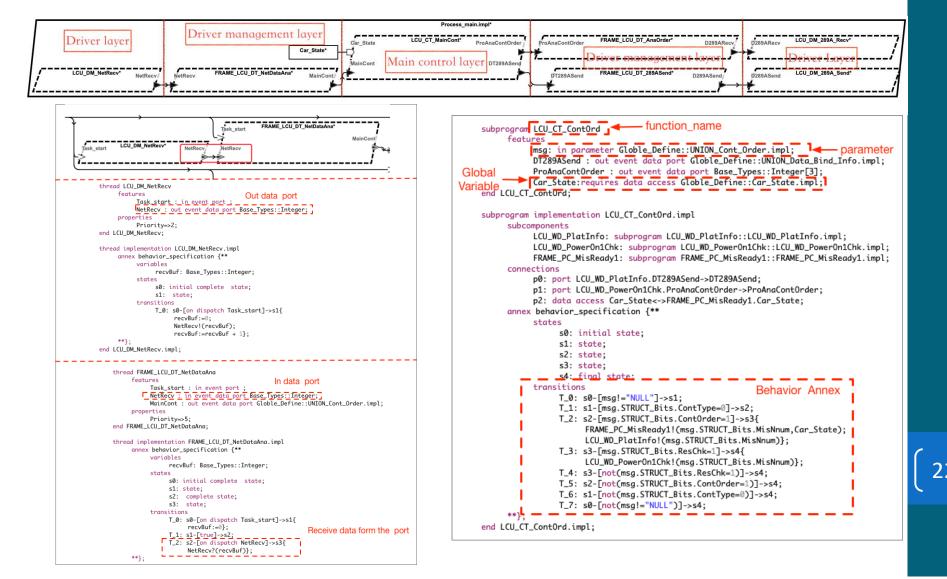


T3: Prototype tool

✓Implementation of the tool



✓ Generated AADL models



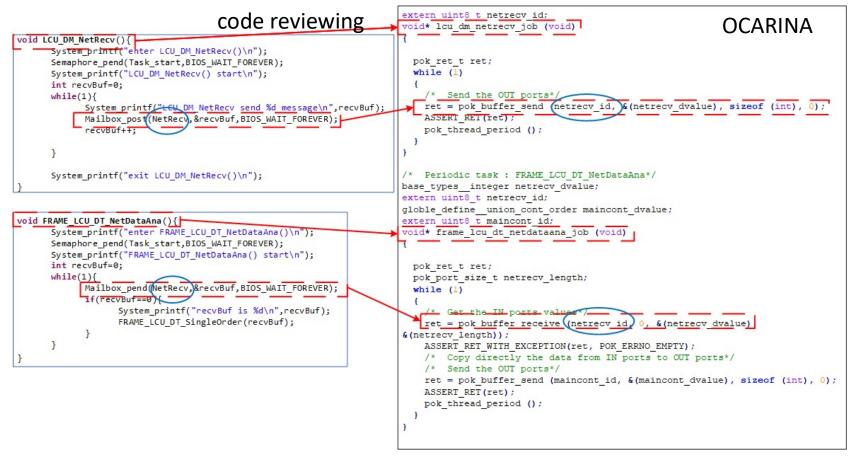
Generated AADL models

	AADL model (line)	Threads	Subps	Coverage
Exhaust cover control	1800+	4	14	93%
Rocket hatch control	1600+	3	12	92%
Control of rocket launch preparation/cancellation	1600+	4	11	94%
Rocket power-on control	4000+	18	32	93%
Control of rocket hatch unlock/lock	2000+	6	13	95%
Thermal battery activation control	1800+	6	15	93%
Control of safety mechanism unlock/lock	1900+	6	15	95%
Rocket ignition control	1700+	4	13	93%
Rocket power-off control	1200+	4	9	94%
Rocket launch control system	17600+	55	134	94%

*The reason why the coverage rate of the generated model does not reach 100% is that some codes are not easily expressed in the behavior annex, such as bit operation and type mandatory conversion, etc. It allows us to complement and refine the models.

✓ Validation of the C2AADL RE process

Comparison between two-versions source code

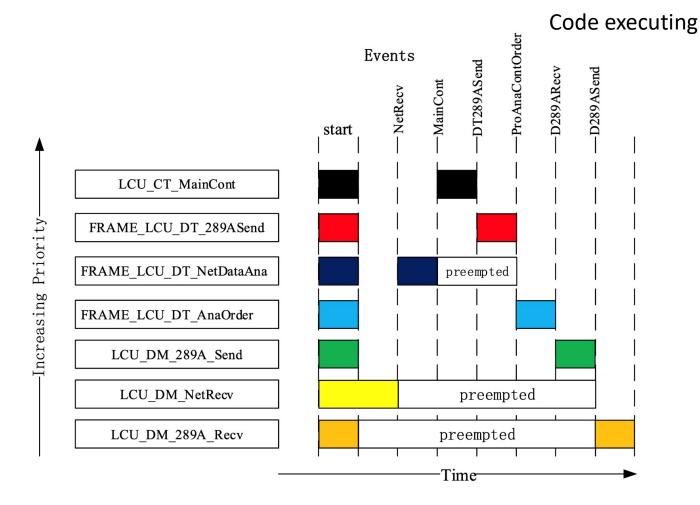


SYS/BIOS: Mailbox_post, Mailbox_pend

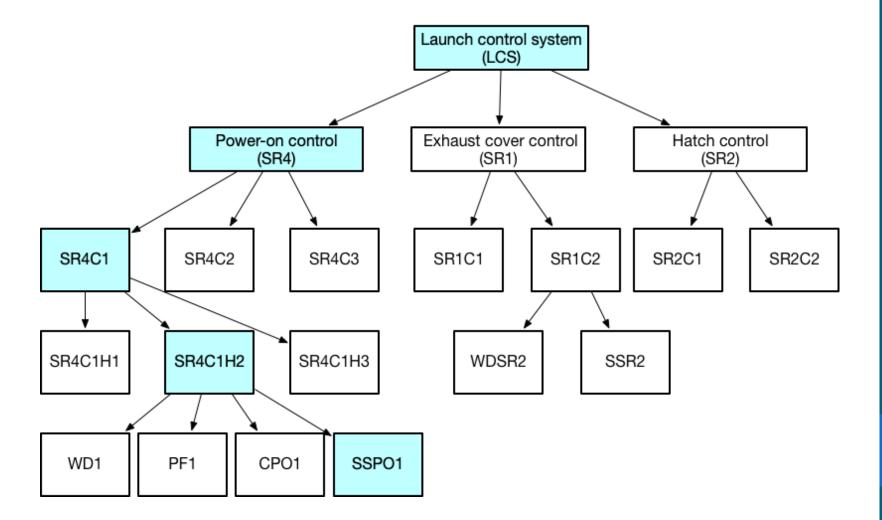
POK: pok_buffer_send, pok_buffer_receive

✓ Validation of the C2AADL RE process

Comparison between two-versions source code



Compositional Verification of the generated AADL models



Compositional Verification of the generated AADL models

thread FRAME_LCU_DT_NetDataAna			
features			
Task_start : in event port ;			
NetRecv : in event data port Base_Types::Intege			
MainCont : out event data port Globle_Define::U	JNION_CONC_Order. Impl;		
<pre>properties Priority=>5;</pre>			
annex agree {**			
property judge_MainCont=			
MainCont.Info=1 and MainCont.STRUCT_Bits.MisNnum=1	and MainCont STRUCT Rits Continuo-0.		
assume "A:FRAME_LCU_DT_NetDataAna receive data from			
guarantee "G:FRAME_LCU_DT_NetDataAna send data to N			
**};	atheone. Judge_Matheone,		
end FRAME_LCU_DT_NetDataAna;	thread LCU_CT_MainCont		
	features		
	Task_start : in event port ; MainCont : in event data port Globle_Define	:::UNION_Cont_Order.impl:	
	DT289ASend : out event data port Globle_Def	ine::UNION_Data_Bind_Info.impl;	
	ProAnaContOrder : out event data port Base_ Car_State:requires data access Globle_Defin		
	properties	ecur_state.tmpt,	
	Priority=>7;		
	annex agree {** property judge_MainCont=		
		s.MisNnum=1 and MainCont.STRUCT_Bits.ContType=0;	
	property judge_DT289ASend=		
	DT289ASend.STRUCT_Bits.MisNnum=1 ;		
thread FRAME_LCU_DT_AnaOrder	assume "A:LCU_CT_MainCont receive cont from augrantee "G:LCU_CT_MainCont send order to	<pre>FRAME_LCU_DT_289ASend" : ProAnaContOrder=1;</pre>	
features	guarantee "G:LCU_CT_MainCont send data to F		
Task_start : in event port ;	<pre>**}; end LCU_CT_MainCont;</pre>		(
<pre>ProAnaContOrder : in event data port Base_Types::Integer; D289ARecv : out event data port Base_Types::Integer;</pre>	end LCO_CT_Mathcone,		27
properties			
Priority=>4;			
annex agree {**	Contractor Contractor 1		
assume "A:FRAME_LCU_DT_AnaOrder receive data from LCU_CT_Ma guarantee "G:FRAME_LCU_DT_AnaOrder send data to LCU_DM_289A			
<pre>**};</pre>			
end FRAME_LCU_DT_AnaOrder;			

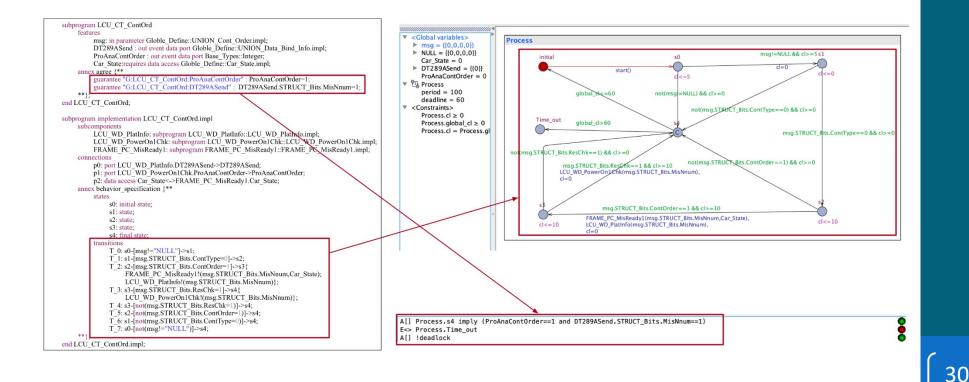
Compositional Verification of the generated AADL models

Property	Result
Verification for Process_main.impl	70 Valid
🔻 🗸 Contract Guarantees	8 Valid
LCU_DM_NetRecv assume: A:LCU_DM_NetRecv	Valid (Os)
✓ FRAME_LCU_DT_NetDataAna assume: A:FRAME_LCU_DT_NetDataAna receive data from NetRecv	Valid (Os)
LCU_DM_289A_Send assume: A:LCU_DM_289A_Send	Valid (Os)
LCU_DM_289A_Recv assume: A:LCU_DM_289A_Recv	Valid (Os)
FRAME_LCU_DT_289ASend assume: A:FRAME_LCU_DT_289ASend	Valid (Os)
FRAME_LCU_DT_AnaOrder assume: A:FRAME_LCU_DT_AnaOrder	Valid (Os)
LCU_CT_MainCont assume: A:LCU_CT_MainCont receive cont from NetDataAna	Valid (Os)
Subcomponent Assumptions	Valid (Os)
This component consistent	1 Valid
LCU_DM_NetRecv consistent	1 Valid
FRAME_LCU_DT_NetDataAna consistent	1 Valid
LCU_DM_289A_Send consistent	1 Valid
LCU_DM_289A_Recv consistent	1 Valid
FRAME_LCU_DT_289ASend consistent	1 Valid
FRAME_LCU_DT_AnaOrder consistent	1 Valid
LCU_CT_MainCont consistent	1 Valid
Component composition consistent	1 Valid
Verification for LCU_DM_NetRecv	4 Valid
Verification for FRAME_LCU_DT_NetDataAna	9 Valid
Verification for LCU_DM_289A_Send	3 Valid
Verification for LCU_DM_289A_Recv	3 Valid
Verification for FRAME_LCU_DT_289ASend	4 Valid
Verification for FRAME_LCU_DT_AnaOrder	4 Valid
Verification for LCU_CT_MainCont	26 Valid

Compositional Verification of the generated AADL models

🟦 Problems 🔲 Properties 🏥 AADL Property Values 🖷 Progress 🥃 AGREE Results 🛛 🖳 Console	
Property	Result
Verification for RocketExhaustCoverTopLevelSys	13 Valid
 Contract Guarantees 	3 Valid
Subcomponent Assumptions	Valid (0s)
✓ Enter the rocket exhaust cover switch cover command and return the execution result of the command	Valid (0s)
✓ Feedback status of exhaust cover after execution	Valid (0s)
> 🗸 This component consistent	1 Valid
> 🗸 Rocket_ExhaustCover_input consistent	1 Valid
> 🗸 SelfCenter consistent	1 Valid
> 🗸 Component composition consistent	1 Valid
🗸 🗸 Verification for SelfCenter	6 Valid
🗸 🛹 Contract Guarantees	2 Valid
Subcomponent Assumptions	Valid (0s)
Self-Check whether the structure is normal	Valid (0s)
> 🗸 This component consistent	1 Valid
> 🗸 watchdog_sequence consistent	1 Valid
> 🗸 SelfCenter1_sequence consistent	1 Valid
> 🗸 Component composition consistent	1 Valid

Model checking of the leaf components



✓ Effectiveness

- Three industrial case studies
- Two OS platforms

✓ Comparison with other MDRE tools

Tools	Parameters					
	Source Artifact	Target Model	Structural	Behavioral	Runtime	V&V
MoDisco	Java, JSP	UML	Y	N	N	N
fREX	Java	UML.	N	Y	N	N
RE-CMS	PHP	AST of PHP	Y	N	N	N
Src2MoF	Java	UML	Y	Y	N	N
srcYUML	C++	UML	Y	N	N	N
Wang [13]	C	AADL	Y	N	N	N
C2AADL reverse	C	AADL	Y	Y	Y	Y

Comparisons with a part of MDRE tools.

Conclusion and Future work

- Conclusion:
 - An MDRE approach named C2AADL_Reverse: the transformation from multi-task C source code to AADL includes three parts: Structural, Behavioral and Run-time transformations.
 - Validation and verification approach of C2AADL_Reverse
 - The prototype tool
 - Industrial case studies
- Future work
 - We are currently working on the semantics preservation proof of C2AADL within Coq
 - Coq semantics of AADL synchronous fragments
 - Coq semantics of specific C multi-thread libraries
 - Semantic-preserving transformation from C to AADL
 - Compositional verification of the AADL asynchronous execution model (X-AGREE)

Thank you very much!