From high-integrity embedded systems to resiliant autonomous cyber-physical systems

Daniela Cancila

École doctorale
Sciences et Technologies de l'Information et de la Communication
(ED STIC)

Mémoire présenté pour l'obtention du

Diplôme d'Habilitation à Diriger des Recherches

Université Paris Sud

Spécialité « Informatique »

présentée et soutenue publiquement par

Daniela CANCILA

3 Févier 2022

From High-Integrity Embedded Systems to Resiliant Autonomous Cyber-Physical Systems

Composition du jury

Frédéric BOULANGER, parrain	Professeur à CentraleSupélec
Frédéric MALLET, rapporteur	Professeur à l'Université Côte d'Azur
Amel MAMMAR, rapporteur	Professeur à Télécom SudParis
Peter MARWEDEL, rapporteur	Professeur i.R, Dortmund University
Jérôme HUGUES, examinateur	Senior Researcher à Carnegie Mellon University
Laurent PAUTET, examinateur	Professeur à Télécom ParisTech.

Caminante, no hay camino,
se hace camino al andar.

Antonio Machado
Proverbios y cantares XXIX
Campos de Castilla

Contents

Acknowledgements

I have many to thank for having helped me to this achievement:

the University of Paris Saclay for accepting my candidature for the *Habilitation à Diriger des Recherchers* diploma: I am proud to defend my HDR thesis in this university's vibrant environment;

all colleagues I have worked with during my research trajectory, in academia, in industry, and at the CEA: thank you all!

Nicole Bidoit-Tollu, for her valuable comments on my application procedures;

Frédéric Boulanger for willing to be my *parrain* and for giving me his precious advice to improving my *Mémoire*;

the members of my HDR Exam Board for their constructive recommendations to improve the manuscript and the thorough discussion of it;

the head of my Department at CEA, Fabien Clermidy, for being so strongly supportive of my HDR project;

my MSc supervisor, Corrado Böhm and and my PhD supervisors, Furio Honsell and Marina Lenisa, whom I shall never forget, for having educated me to formal methods, their applications and to team's coordination;

my family and my friends for their strong and constant support.

Chapter 1

Introduction

This document constitutes my *Mémoire* in order to obtain the French national diploma *Habilitation à Diriger des Recherchers* (HDR).

This *Mémoire* reflects my research activities, which I have carried out, coordinated and promoted over 15 years after the attainment of the *Doctor of Philosophy* (PhD) in Computer Science (Univ. of Udine - Italy).

Along my research trajectory, I started from studying embedded systems before expanding my interest to cyber-physical systems (CPS), thus focusing on the integration of digital embeddedness with a dynamically-changing physical world where the system is deployed, via a variety of sensors and actuators. The scientific community had it very clear that the CPS realm constituted a disruptively new discipline of systems engineering, whose development would require new, highly interdisciplinary, methods, design and analysis tools. As a consequence of that, CPS require new pedagogical methods for teaching and training.

In the passage from embedded to cyber-physical systems, I have especially investigated critical system properties, my research activities contributing with to a solid, rich and fruitful structural collaboration with KTH (Sweden), the University of Trento (Italy), the University of Berkeley (USA), and my workplace, CEA. In that trajectory, I have witnessed a paradigm shift from high-integrity embedded systems to safety for cyber-physical systems, to resilience for mobile and autonomous cyber-physical systems.

The structure of the *Mémoire* reflects my research trajectory from high-integrity embedded systems to resilience for autonomous cyber-physical systems (ACPS), and outlines a medium and long-term vision in the field of critical systems. The remainder of this document is organized as follows.

Chapter 2 concerns High-Integrity Embedded Systems, and, more specifically, how to apply and automate correctness by construction in the design phase in order to reduce the cost during the verification and validation phase and how to prove coherence and data consistency of the system over a wealth of flourishing languages and heterogeneous tools.

I investigated the aforementioned research topics, by integrating, between 2004 and 2009, high-quality international research teams in different cities and countries. With prof. Tullio Vardanega (Padua, Italy), we studied how to apply the "correctness by construction" principle to the software modeling of high-integrity embedded systems. A couple of years later, I transferred similar scientific results to a French small industry, by working with the research team driven by Vincent David (Paris, France). With François Terrier (CEA, France), we devised safety analysis techniques to apply to the modeling of critical systems, hence using model-based techniques and languages. Finally, I contributed to the investigation in the European current efforts for using meta-modeling and tools in the integrated

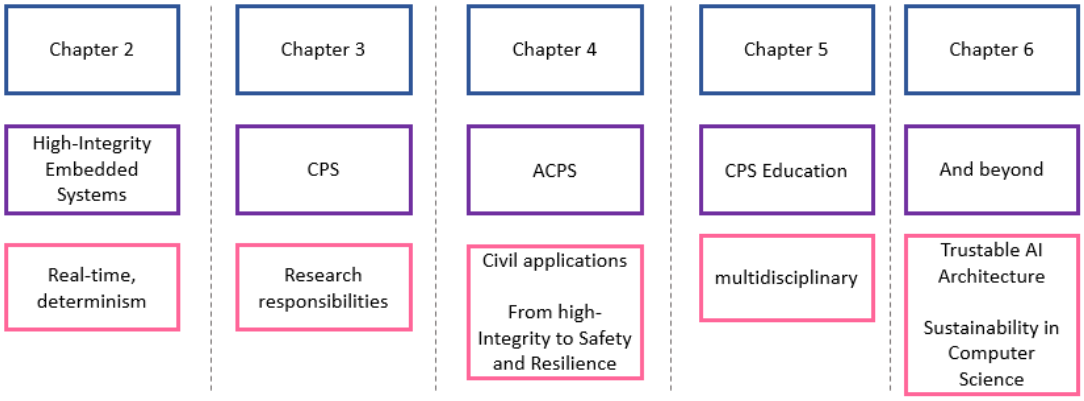


Figure 1.1 – Overview of the Structure of this *Mémoire*

development of critical systems. The study provided an overview of the current European situation. It was conducted in parallel by two research teams, one in USA and one in Europe (which I was part of), and driven by Alberto Sangiovanni-Vincentelli.

My results are published in IEEE Transactions on Industrial Informatics (IEEE TII), IEEE Design & Test (IEEE D&T) and in several conferences and workshops [CP08; Can+10; CPV08; Pas+09; Can+09; Esp+09].

Chapter 3 addresses safety for Cyber-Physical Systems. It reflects my responsibilities at international level, based on the direct involvement in the European roadmaps on CPS, and outlines the lessons that I have learnt across them. Moreover, the chapter discusses my technical contribution to the safety and certification of critical CPS, with a special focus on proving the compliance of the design of a critical CPS in the railway domain to the corresponding safety norms.

After my research formation (2004-2009), between 2010 and 2021, I took direct responsibility in the construction of research proposals under European and French national research funding, in operating as Principal Investigator for the ensuing projects, in supervising MSc and PhD students, and in contributing to the Electronic Components & Systems (ECS) Strategic Research and Innovation Agenda (SRIA) for the ECSEL program, where in 2020 I became the co-leader of the "Quality, Reliability, Safety and Cybersecurity" core group, directly in charge of the "Safety and Resilience" and "Cybersecurity and Privacy" Challenges. My involvement in the ECS-SRIA contributed to the setup of the ECSEL (`https://www.ecsel.eu/` [ECS]) and PENTA/Eureka (`https://penta-eureka.eu/` [Pen]) and more recently Xecs (`https://eureka-xecs.com/` [Xec]) industrial European research framework programs. Those programs aim to "provide the Electronic Systems & Components community with a specific support capability to address the huge challenges created by the rapid development of the global digital economy. Our vision is to offer a program that is open to all the elements of the ECS Community; Large Enterprise, SME's, RTO's and Universities" [Pen].

The results from this strand of work have been published in [ECS21; Sou+16; Dal+10] and given via keynotes in [Can15a; Can12].

Chapter 4 concerns resilience for (mobile) critical Autonomous Cyber-Physical Systems. The chapter discusses several research problems on (mobile) critical ACPS, with/without AI, via several civil applications. Whilst its sections investigate the civilian application under study together with the addressed research problem, my analysis and reflection considers ACPS applications as a whole (e.g. autonomous vehicles, drones, robots, trains, nuclear reactor) and is grouped at the end of the chapter.

The majority of my research work at CEA is situated within this space - thus contributing to the vision *"From Research To Industry"* brought by the *Technological Research Direction* (*Direction de la Recherche Technologique*) of CEA [CEAa] to which my department, (*Integrated Circuits and Digital System Division*), belongs.

Results are published in IEEE Design & Test, IEEE Access and Ada User Journal and in several workshop and conferences [May+19][CCC21] [MCR17][LCC17] [Pas+19][Mou+19][Can+19] [Can+14][CSP14][CLB17].

Chapter 5 reflects my experiences on *the CPS Education Challenge*. This activity is not my main and priority activity at CEA. Therefore, the results and the analysis in this chapter are not "original" in the sense of the those discussed in the other chapters of this *Mémoire*. Most of them are firstly introduced by my colleagues in Europe at KTH, TU Dortmund, Université de Toulouse - ISAE, and in the USA at Berkeley, Halmstad University, Rice University. However, as I am part of the WESE community, I actively contributed to the WESE intent by applying these methods in my working space. The discussion I present in this chapter is the reward of several years of my reflection on this topic and actions that I am carrying out on it.

Some results are published in ACM, WESE series [Can+16; CZP15].

Chapter 6 outlines the way forward that I envision for my future research in critical ACPS to address methods and analysis techniques for the design and development of sustenability and trustable architectures of AI-driven mobile and autonomous critical cyber-physical systems operating in highly dynamic environments.

Chapter 2

High-Integrity Embedded Systems

Contents

In the first years of the 21st century, embedded systems, i.e. information processing systems embedded into enclosing products [Mar03], became more and more complex. Because of the increasing complexity, several errors can occur at different levels of the design phase, for example in the component's specification, in the software development and in the software and hardware integration. As a result, the complexity of embedded systems rises with the quest for better functionality and quality. In face to this landscape, the scientific community devotes an important effort to the development of methodologies and tools in order to take under control the Design and the Development (D&D) and the Verification and Validation (V&V) phases.

My formation to research occurs in this context. I contribute to understand how we can anticipate system verification, how we can enforce and prove correctness by construction of an automatic code generation and how we can realize it. In doing so, I adopted contract-based approaches [BCP07] and exploited the interfaces of components [AH01] to specify real-time properties. The result shows also an unexpected consequence: we arrive at the same conclusions of J. Sifakis, which he expressed in [Sif05a], in a different and independent way [Can+10]. Then, I address how we can guarantee correctness and safety analysis under heterogeneity of languages and tools in a model-based design approach [Pas+09; Can+09; Esp+09].

2.1 Research Context

Not only 2001 is the emblematic year chosen by Stanley Kubrick for the space odyssey, it is also the year in which the article *Interface Automata* by Luca de Alfaro and Tom Henzinger has been published [AH01]. Before discussing that article, let us analyse the research context, which was predominant in the first decade of the new century.

High-Integrity Embedded Systems are embedded systems, generally performing few dedicated and predefined tasks in a given platform under a set of temporal requirements. Those systems are mainly applied to the traditional critical application domains, such as railway, avionic, space or nuclear, and then the compliance to the related safety regulations is mandatory. The guarantee of the determinism

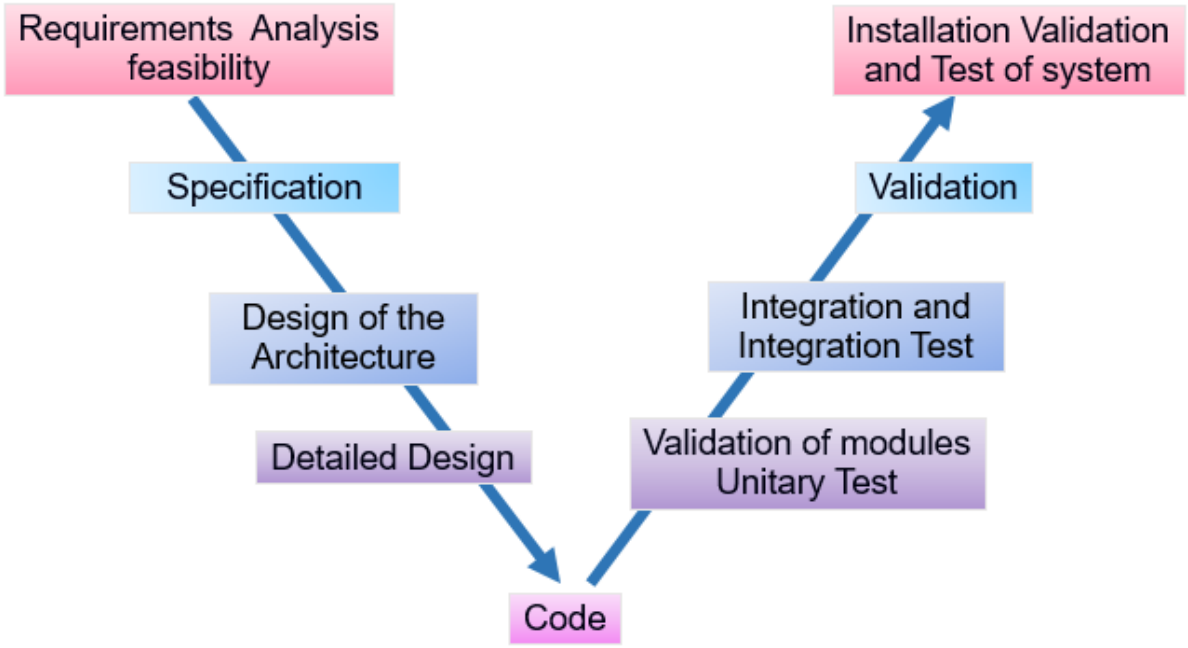


Figure 2.1 – A Simplistic Overview of the V&V Model

property, i.e. given the same set of inputs and the same state of the system, the system provides the same output (including the control-commands) in the same execution time, is highly demanded and required by the safety regulations. As a result, in that period, the control of real-time properties acquired a crucial and predominant role in many research works and directions (e.g. [BB04; DMT00; DB05; Bal+98; GLN01]).

Although the safety norms mainly differ from the required safety level that a system should meet, from the applied countermeasures to detect an error or to control an accident, all of them introduce a similar V-schema for the D&D and for the V&V phases of these systems. Figure 2.1 is an illustration of the V&V Model. The descent axe addresses the specification of the requirements until the design and the development of software and to software and hardware integration. The ascendant axe concerns the validation and verification phases.

The increasing complexity of high-integrity embedded systems involves a higher cost in the ascendant phase, mainly because of errors, which are discovered too late in the D&D phases. Moreover, because the V&V phase could impact a revisiting of the D&D phase, it is performed too late with respect to the ambition to reduce the overall cost: an error at the V&V stage has an important impact on the cost of the system. Figure 2.2 [Hal07] relates at which stage an error is found in the life cycle of a system with the associated cost. Hall's study [Hal07] highlights how the system and software cost is increasing during the development, with a peak on the cost if the error is found in the maintenance phase. As a consequence, two main research questions clearly emerge. On the one hand, how to achieve a better quality of the software. On the other hand, how to obtain a cost reduction both in the descending and ascending phases of the V&V model, in an *esprit* of adhering to the 'faster, better and cheaper' principle to the D&D.

To solve these issues the scientific community is looking forward to new methods and tools. One of them is the layered top-down approach, inherently suggested by the application of the V-Schema. The quest for understanding which parameters belong to a given layer carries out the so-called *separation of concerns* principle, which specially focuses on the separation between functional and non-functional (or structural) properties. The latter concerns parameters related to the execution

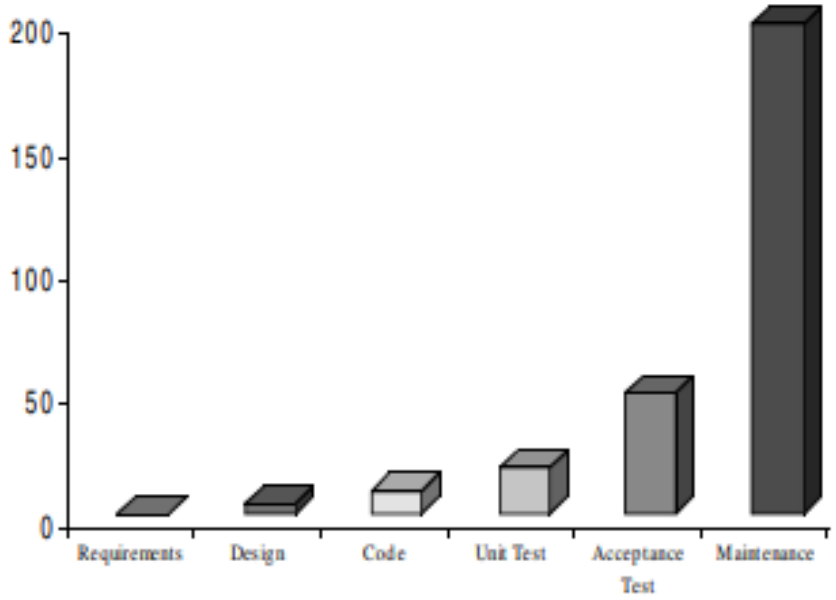


Figure 2.2 – The impact of errors in the D&D phases. Figure extracted from [HALL]

platform, included performance, concurrency management to memory, real-time and safety guarantees. The separation of concerns principle allows designers to develop specifications for function, timing, architecture and structure which can later be reused or adapted to more quickly achieve the desired result. This separation is expected to provide an additional increase in software reuse: the more a functional specification of a system is free from structural aspects, the more it can be reused and/or integrated in another functional specification - thus improving software and system reusability.

However, despite the separation of concerns and an good organization of the system and software development, the strictly top-down approach shows some limitations, among which obtaining a high reuse of components.

In this regards, the scientific community develops a wealth of flourishing formalism and tools for the software and system development of High-Integrity Embedded Systems. In face to that heterogeneity, in which non-functional aspects, such as real-time, safety, performance, are addressed in dedicated tools and different methods, two main idea rise. First, the need to 'revise the most basic computing paradigm and methods' [Sif05b] in order to enforce 'an abstraction level able to encompass heterogeneity of components' [Sif05b]. Second, the need to anticipate system verification and the capability to infer properties from components to the systems. This aspect is fundamental to achieve a modular certification of an evaluative system (certification is mandatory for High-Integrity Embedded Systems).

Nowadays, most of those emerging formalisms have been accepted, albeit not without live debates, other have disappeared. But all have contributed to understanding and reasoning on the D&D and V&V phases for High-Integrity Embedded Systems. The methodologies that prospered in that period can be classified by the dual taxonomy of abstract vs. a constructive approaches.

The abstract approach defines a formalism comprised of uninterpreted operations on abstract components. A component in this context is often viewed as an interface and the focus is placed on specifying the conditions under which that interface can be substituted or refined by another interface. The abstract formalism can be axiomatic, algebraic, based on graph theories as well as grammars.

The constructive approach is dual to its abstract alternative. It centers around a given language to express components, a given platform to host and execute those components at run time, and a set of positive (shall) or negative (shan't) constraints on what those components can do in terms of

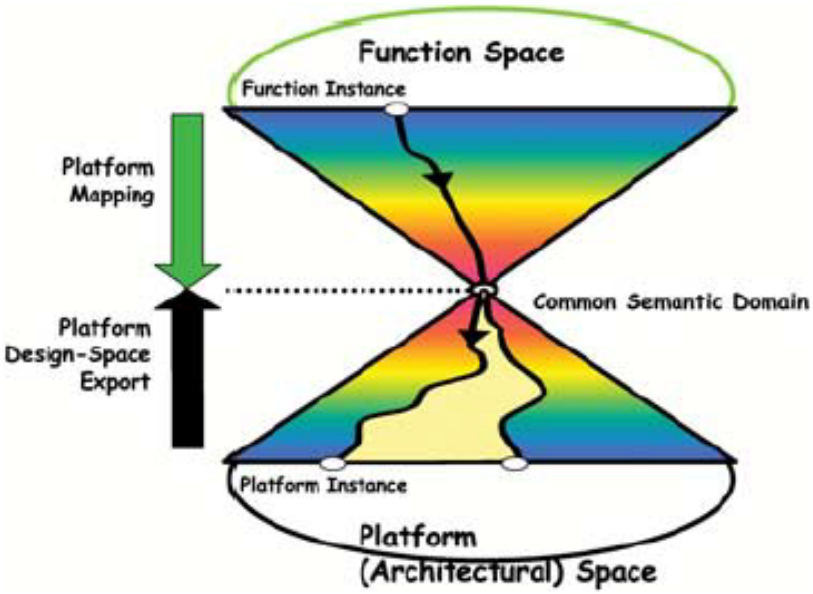


Figure 2.3 – CAPTION and extracted from [San07]

run-time behavior and interaction.

The abstract approaches have the advantage of being able to capture a large set of application domains. However, the moment a specific application needs to be addressed, the abstract theory must be instantiated to it (to permit concrete implementation) and the consistency between the implementation and the original theory must be proven, for example axiomatically. The instantiation process may thus become problematic. It may even require an inordinate amount of effort, without the guarantee of success, depending on the distance between what the theory requires and what the implementation technology actually provides.

The constructive approaches have the dual problem: they often originate from a given application domain and attempt to generalize beyond its frontier by augmenting the availed expressive power. There is no a priori guarantee however that the generalization can actually succeed: the requirements from a given domain, normally expressed in terms of attributes and constraints, may turn out to be incompatible with those from another domain.

Methodological approaches, such as platform-based design [Pin+06], model-driven engineering [Sel07; Sch06], Platform-Independent Model (PIM) and Platform-Specific Model (PSM) layers [Sch06] try to overcome this duality, by separating the space of specification with the space of the implementation - thus simplifying the design and increasing the reuse. As a result, the concept of platform clearly emerges as separated from the functional one.

Figure 2.3, firstly introduced in [San07], shows two different spaces, one devoted to the functional specification and another one to the platform execution, in a layered approach. Both layers converge in a middle point, which represents a mapping and optimization, i.e. the selection of a dedicated execution platform, which is able to execute the functional code by guarantying that the properties specified and *wished* at the functional level are feasible during the execution.

To improve reuse, the scientific community addresses research topics, such as determining if two components are compatibles, if a component can be refined by other components, or if a component can be substituted by another component. In this line, in 2005, J. Sifakis spread the use of two terms, which have had enormous success in the literature: *composability* and *compositionality*. Composability allows inferring that a component's properties are not affected when its structure is modified [Sif05a].

Compositionality allows inferring global system properties from component properties [Sif05a].

However, the increased adoption of software reuse methodologies was gradually shifting the cost linked to software development from the design phase to the verification and validation (V&V) phase. This occurs because components are employed in contexts that are potentially different from those they were initially intended for. Thus, particular care must be taken in ensuring that the interaction between the reused components yields the correct results. The validation of this interaction is typically more complex than that of behavior alone, because of the potential large number of different synchronizations that may take place. Indeed, errors in software systems are most commonly the result of some unforeseen interaction.

To overcome the increasing cost in the V&V phases, in 2002, one year after the publication of *Interface Automata* [AH01], some researchers promoted a correctness by construction approach. The main objective of this approach is 'to deliver durable software that is resilient to change throughout its lifecycle' [Cha05]. Correctness by construction demands a rigorous design methodology, which can help establish a correct implementation path early in the design flow. This approach is described by Chapman as an "economical method to develop security and safety-critical applications" [Cha05] which does not involve zero defects. Rather zero tolerance of defects in fact [HC02]. Therefore, the underlying idea in a correct by construction approach is to make it hard to introduce errors early in the development and, otherwise, to detect them and remove them as soon as they are introduced using static analysis techniques. The techniques promoted by Sifakis on compositionality and composability are a means to ensure correctness by construction [Sif05b].

It is now the time to come back to *Interface Automata* [AH01]. In the article, the authors introduce a lightweight formalism to enforce temporal aspects of software component interfaces. This formalism stems from a contract-based approach, and reinforces it by identifying the interfaces of components as key-elements (to specify *guarantees* and *assumptions*) in order to achieve a modular and safer design. As an inherent result, engineers can apply the *divide et impera* paradigm for understanding, specifying and reasoning on single modules. The article had a significant impact on our research community and constituted a cornerstone over the years of my research formation.

2.1.1 Summary: Solved and Open Questions

During the first decade of the new century, in face of the increasing complexity of High-Integrity Embedded Systems, the scientific community devoted an important effort to the introduction of new methods, languages and tools in order to reduce the cost of the system and software development. This scientific direction fixed several issues, mainly devoted to the reuse of components, by introducing:

- a clear notion of execution platform, which is independent and distinct from the functional design;
- the identification of interfaces as key-elements in software modularity and, then, their reuse;
- the separation of concerns;
- composability and compositionality;
- the introduction and the evangelization of several approaches, such as platform-based, model-based, model-driven, component-based and contract-based approaches;
- and the correctness by construction principle.

The main open-questions still concerns how to apply and automate correctness by construction in the design phase in order to reduce the cost during the verification and validation phase and how to prove coherence and data consistency of the system over a wealth of flurishing languages and heterogeneous tools.

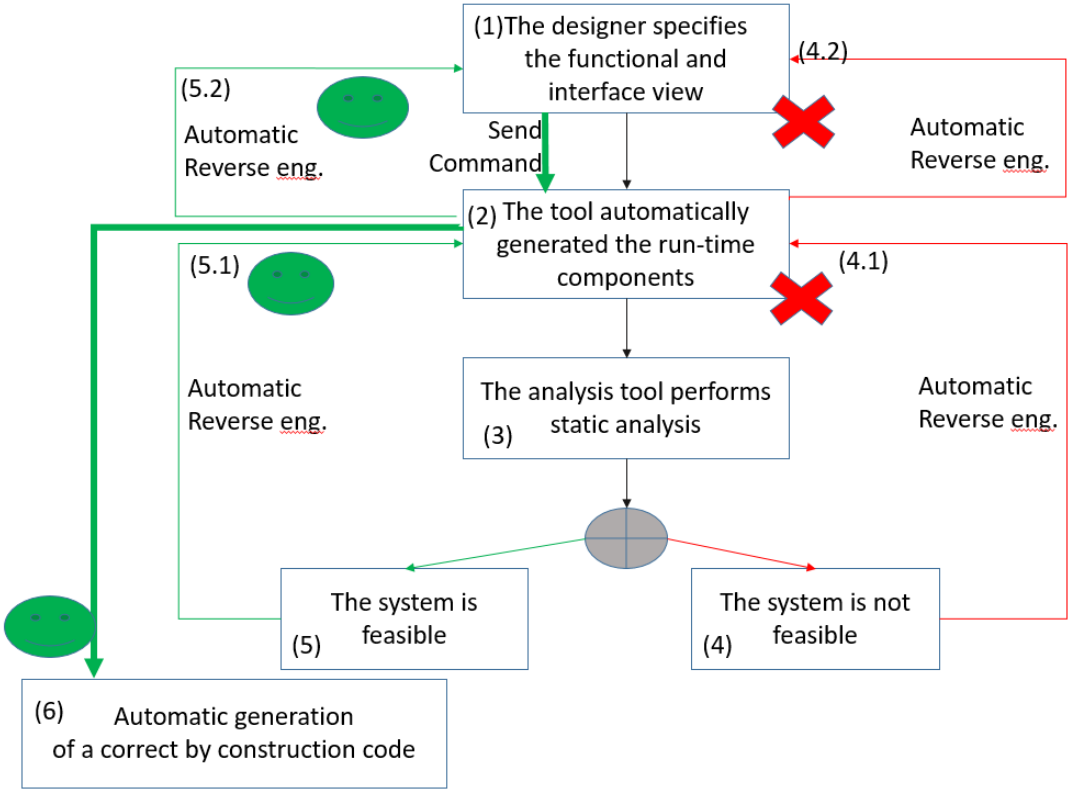


Figure 2.4 – Overall Strategy: from the design space to the automatically generated correct code

2.2 Contribution: Foundation of Correctness-by-Construction

My main contribution concerns the formal methods underlying the application of a correctness by construction approach in a modeling tool for High-Integrity Embedded Systems. The main objective of this work was to specify a modeling tool, which is able to automatically provide model transformations and automatically generated the code, without any semantic distortion. Moreover, the tool has to integrate the correctness by construction principle and capitalizes on the results of the scientific community related to methods for increasing software and systems reuse. The investigation and the results have been provided for the Ravenscar Computational Model (RCM). RCM is a collection of concurrent components designed to satisfy the restrictions of the Ada Ravenscar profile, which is a subset of the Ada programming language tailored to high-integrity real-time embedded systems [Var06]. The Ravenscar Profile is adopted by the European Space Agency and in the railway domain. RCM is aimed at addressing two crucial questions for high-integrity real-time systems: how to manage concurrency through trustworthy architectural choices, and how to guarantee static analyzability of a system. In order to warrant determinism and predictability, RCM prohibits task synchronization, any form of recursion, and dynamic creation or allocation of resources. Moreover, RCM requires tasks to be nonterminating "to mitigate the hazard that may be caused by tasks terminating silently" [Var06].

The way forward follows two main trajectories. First, anticipating temporal analysis early in the design phase (See Figure 2.4).

In this step, the temporal analysis provides a first evaluation of the feasibility of temporal constraints specified by designers on the components, for example if the system specification meets all deadline. At the tool level, we need to create a reverse engineering process from the analysis tool and the design space, which clearly indicates to designers the values that have to be changed. Only once the system meets all temporal constraints, and is feasible at execution time, one can automatically

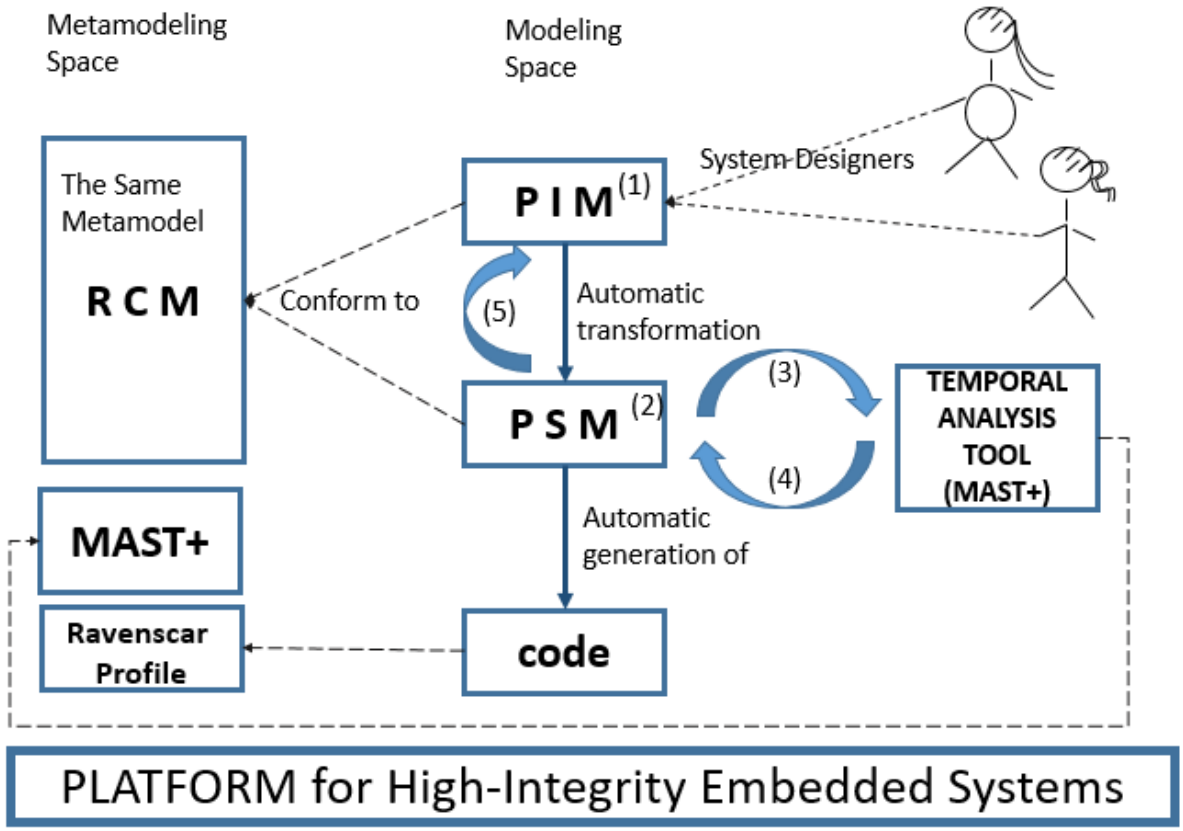


Figure 2.5 – PIM, PSM, model transformation and automatic generation of code

generate the corresponding (structural) code, which is required to being correct by construction. This implies that we need to prove that the tool guarantees and preserves correctness by construction in a seamless process.

Second, we have to separate functional space from platform space, as shown in Figure 2.5. The design space in Figure 2.5 is refined in two layers, PIM (Platform-Independent Model) and PSM (Platform-Specific Model). In most cases, the difference between PIM and PSM reflects the separation between functional and structural models. The functional model abstracts from the requirements of the implementation. The structural model, instead, addresses the requirements placed on the run-time behavior to be exhibited on the target platform and it must also undergo the verification and validation activities that precede deployment. The PSM components are automatically obtained by model transformation, from (a view of) PIM to (a view of) PSM.

To guarantee that model transformations do not introduce any semantic distortion, PIM and PSM shall have the same metamodel. This solution provides (i) *consistency* in the information by using the same syntax and semantics and (ii) *propagation of information* from PSM to PIM by using constraints in the metamodel. In addition, a PSM component shall be correctly realized on the target platform and can execute in a property-preserving manner [CPV08]. In our case, PIM and PSM share the same metamodel (RCM) and the code is compliant to the Ravenscar Profile [Can+10].

In an unrestricted general model space, problems may arise from the composition of components, which may cause unexpected, unwanted or erroneous behavior to emerge at run time. This is the reason why the generated run-time components at PSM cannot be directly changed by the designer, otherwise the seamless and automatic process with preservation of guarantees is broken. As introduced in the discussion of Figure 2.4, while the structure of PSM is generated to guarantee certain properties, the temporal properties are statically and automatically analyzed using MAST+ [Gon+01]. The results

of the analysis are propagated back and made available to the designer in PIM. The designer can then change the timing attributes and iteratively reanalyze the system until all temporal requirements are met. Once the system meets all temporal constraints, and then it is feasible at the execution time, a designer can automatically generate the corresponding (structural) Ravenscar code, which is correct by construction. The code is compliant to the Ravenscar Profile.

The entire work from PIM to PSM and Ravenscar code has required a controlled separation of concerns related to the RCM attributes, which is based on a solid mathematical formalization. I would highlight here that when we deal with High-Integrity Embedded Systems, tools must not only provide some static analysis and automatic generation of models or code, but also provide a mathematical foundation, or proof, that what they are doing is correct and guarantees some wished properties. To achieve correctness by construction at the formal level, I adopt different formalisms.

Table 2.1 summarizes the process and explains the different views. Table 2.2 associates the different views with the adopted formal method for the verification. I adopt three main methods:

1. *RCM Interface Grammar* to model transformations from PIM components to PSM run-time components;
2. *hypergraphs* to capture the functional semantics including the attributes risen from the separation of concerns;
3. a *Contract-Based Approach* to set properties on the component's interface.

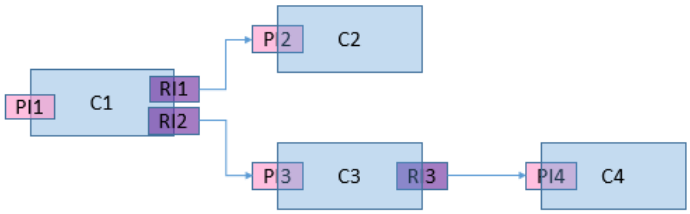
VIEW	MEANING
Functional views	the designer specifies the functionality of the system as a collection of classes and interfaces using a UML-like notation. Interactions between components are made explicit by identifying functional dependencies between methods. Provided and required interfaces explicitly specified
Interface view	Specification of the runtime attributes on the interfaces, such as the worst-case execution time and the period of an operation
Implementation view	automatically generated from the interface view, the run-time components are compliant to RCM

Table 2.1 – Main views in the design process

VIEW	VERIFICATION
Functional, Interface and Implementation views	Separation of concerns between functional and structural attributes based on RCM grammars
From Functional view to Interface view	Preservation of functional semantics by hypergraph fully-faithful extension
Interface view	Correct task behaviour and interactions by non-functional attributes setting
From Interface view to Implementation view	Full semantic preservation by RCM Interface Grammar
From Implementation view to Interface view	Static feasibility and sensitivity analysis and round-trip

Table 2.2 – Connecting modeling views and verification steps

Provided and Required Interfaces as Guarantee/Assumption pairs (See Example 1) In a contract-based approach, a contract is defined by a guarantee and a finite set of assumptions [BCP07],



The figure shows four components, $C_1, \ldots, C_4$. Component C_1 has one Provided Interface, PI_1, and two Required Interfaces, RI_1, RI_2. The figure is represented by the following contracts:
Contract$_1 = (PI_1, RI_1, RI_2)$ where PI_1 is the guarantee and RI_1, RI_2 are the assumptions.
Contract$_2 = (PI_2, RI_{2.0})$ where PI_2 is the guarantee and the assumptions are empty.
Contract$_3 = (PI_3, RI_3)$ where PI_3 is the guarantee and RI_3 the assumption.
Contract$_4 = (PI_4, RI_{4.0})$ where PI_4 is the guarantee and and the assumptions are empty.

Example 1 – Guarantee and Assumptions; Provided and Required Interfaces

Contract= (Guarantee, Assumption$_1$ ... Assumption$_n$)

An interface of a component can be *provided* or *required*. It is decorated by one and only one method and some predefined attributes with respect to the applied method (for example, attribute deadline to the *cyclic* method). A Provided Interface of a component represents the guarantee, ensured by the component, and the invoked Required Interfaces -if any- are the assumption. The contract is declared on the component. The adoption of a contract-based approach allows designers to specify formally the interfaces of a component and its properties.

Functional and Interface views as Hypergraph (Examples 2 and 3) The Functional view is characterized by a weighted directed hypergraph, $HG_F = (V_F, E_F)$, where V_F (vertices) is a finite set of services and E_F (hyperedges) are relations over services.

The set of vertices V_F is obtained as the union of the set I_P of provided and the set I_R of required interfaces in the system, i.e., $V_F = I_P \cup I_R$. Each $PI \in I_P$ and $RI \in I_R$ has a weight specified by a combination of attributes which collectively characterize the assumptions and guarantees.

The set of edges E_F is made of two disjoint subsets: hyperedges Δ and simple edges X. A hyperedge $\delta \in \Delta$ relates a single PI to a (possibly empty) list of RIs. We denote a hyperedge δ that links vertex PI with vertices $RI_1, \ldots, RI_k$ as

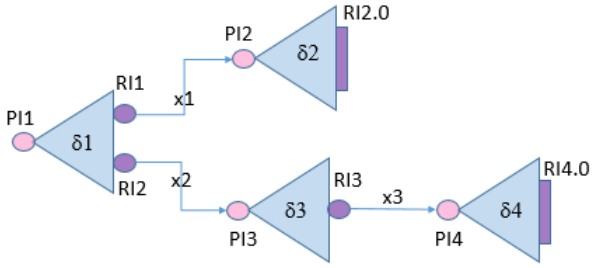
$$\delta = \{PI, \langle RI_1, \ldots, RI_k \rangle\}.$$

The hyperedge identifies the dependency between a provided service and its required services. Conversely, an edge $x \in X$ links a single RI to a single PI and represents the binding of a required interface (function call) to its corresponding provided service. An edge is denoted simply as

$$x = \langle RI, PI \rangle.$$

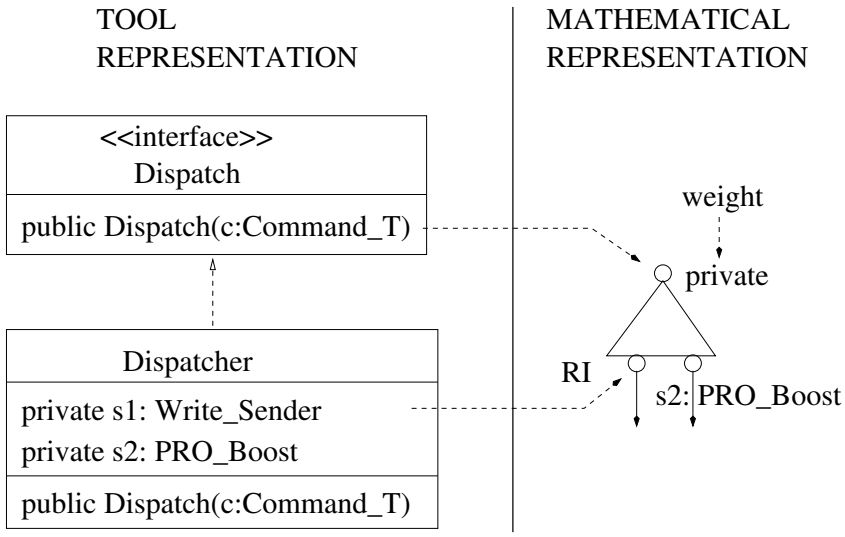
The interface view is characterized by a weighted directed hypergraph, such that $HG_I = (V_I, E_I)$ where $V_I = V_F \cup \{v_0\}$ and $E_I = E_F \cup X' \cup \Psi$.

Vertex v_0 represents the underlying interrupt service infrastructure of the run-time platform, which invokes all $PI \in I_P$ whose activation depends on external events. New set of hyperedges ψ in Ψ (which links the nominal PI to its set of possible modifiers) and additional attributes on the structure and on the time (the temporal attributes, described below), are not shown on the figure to avoid clutter.



The figure represents ones in Example 1, where δ is a hyperedge, which relates a single PI to a (possibly empty) list of RIs, and x is an edge, which related a single PI to a single RI.

Example 2 – Hypergraph of a Functional view



On the left, the figure shows the modeling tool representation of a component in the Functional view. On the right, the same component is shown in its formal representation. It is not required that a designer should learn how to manipulate hypergraphs. Thanks to this representation, we can exploit the associated formal methods to prove properties for high-integrity embedded systems and demonstrate that no semantic distortion occurs during the model transformation and that the automatically generated code is correct by construction.

Example 3 – Tool and Mathematical Representation

From Functional to Interface views By definition, hypergraph HG_I is a fully-faithful extension (in the categorical sense) of hypergraph HG_F by an inclusion preserving homomorphism (an embedding), i.e., HG_I preserves the vertices (with their weights), the hyperedges and the edges of hypergraph HG_F.

Thus, the only additional information conveyed by HG_I consists essentially of attributes on the structure and on the time of the vertices. Despite this, the distinction between HG_F and its extension HG_I is important. In particular, the same Functional view HG_F may be extended by multiple and distinct Interface views, each specifying alternative communication and synchronization architectures. Different Interface views may also be useful to express different deployments for design space exploration.

From Interface to Implementation views The implementation view is automatically generated from the interface view by model transformations. Correctness of the model transformations is proved according to the RCM Grammar and production rules.

The Interface view is formalized by the *Specification Language*, L_I. The implementation view is formalized by the *Implementation Language*, L_C. Both languages share the same set of non-ambiguous attributes. Intuitively, a token in languages L_I and L_C represents a type of PI. Language L_I only includes terminal token, while L_C includes terminal and non-terminal tokens. The production and semantics rules in L_C are the form:

$$\text{non-terminal token} \rightarrow \text{set of terminal tokens.}$$

The RCM Interface Grammar is deterministic and non ambiguous, so that given a token in L_I we can always find one (and only one) production rule which identifies the run-time component able to realize the method specified by the designer in the Interface view. Model transformations are fully described by the inclusion of two formal languages and the production and semantic rules.

Correctness by Construction The proof of correctness is based on several steps and reported in the appendix. The hypergraph formalism allows us to prove the following proposition:

Proposition 1 *If a cycle is present in a thread graph, then it is possible to detect it.*

Therefore, the generated code is by construction guaranteed to be free of harmful cycles.

2.2.1 Impact of the Theorem

The theorem plays an important role because it anticipates the ascending phase in the descending phase of a design, by guaranteeing high-integrity properties. Figure 2.6 shows the applied process in the V-Schema. We expect to achieve both a reduction of the cost and an increase of the quality of the software, thanks to an automatic and correct by the construction process of some insidious steps in the software development.

Moreover, the theorem highlights the importance of combining formal methods, which are necessary for formally reasoning on the system and mathematically prove some mandatory properties for High-Integrity Embedded Systems, to system's modeling specification, where the designer cannot be hindered either by the implementation details nor by the knowledge of formal methods or the target language. To achieve this goal, it is necessary to automatically manage what is *superfluous* in the design phase, through a greater level of both usability of the tool and accessibility to the information about the components.

This last point is all the more relevant in the industrial environment, where the new generation of designers is less prepared to manipulate abstract and formal methods than previous generations, and where it is highly demanded to be immediately productive and efficient. In other words, engineers have

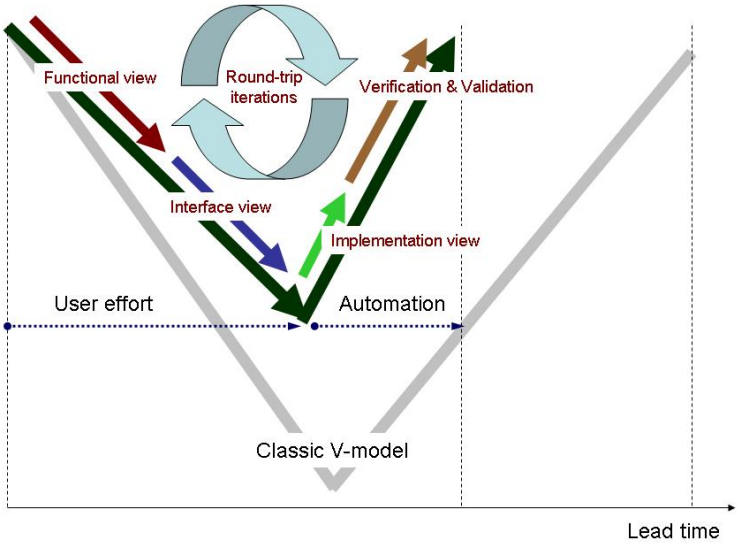


Figure 2.6 – Anticipation of the V&V Phase in the design phase

thought to use modeling techniques for the design of High-Integrity Embedded Systems, while not directly manipulating formal methods to validate whether the system meets all expected properties (e.g. WCET, deadline, safety, etc.). The theorem states that this approach is now available, because the tool is based on formal methods and on RCM.

Although the investigation has been firstly provided on Ravenscar Computational Model (RCM), the intellectual process can be easily generalized and applied to other languages and target platforms.

This is what I did several years later in the *industrial partnership* between CEA and Krono-Safe (*common labs*). Krono-Safe is a French small enterprise which develops the ASTERIOS Tool Suite for the design of avionics and automotive applications. The tool suite "guarantees built-in determinism and correct-by-design software application development" [KRO]. In this context, the target platform was Asterios and the target programming language was PsyC [AD03].

Unlike the Ravenscar Profile, which is a subset of the Ada Language tailored to High-Integrity Embedded Systems, PsyC is a synchronous programming language based on C. PsyC was created to respond to industrial needs related to the nuclear domain [DD00]. Like the Ravenscar Profile, which has a computational model, called RCM, PsyC is supported by OASIS, a deterministic multitask framework for nuclear safety-critical real-time systems. OASIS is industrialized by Areva NP (now acquired by EDF) for the French Nuclear Power Plants. Among the main characteristics, OASIS provides:

1. a clear mathematical model and a supporting formal tools, based on graphs theory,
2. automatic static temporal analysis of real-time properties,
3. the property of determinism, which is essential when we design control-commands for nuclear applications.

The ASTERIOS platform is an evolution of OASIS for different applications domains, where some strong requirements can be relapsed. For example, ASTERIOS exploits multi-cores and advanced techniques on mixed-criticality.

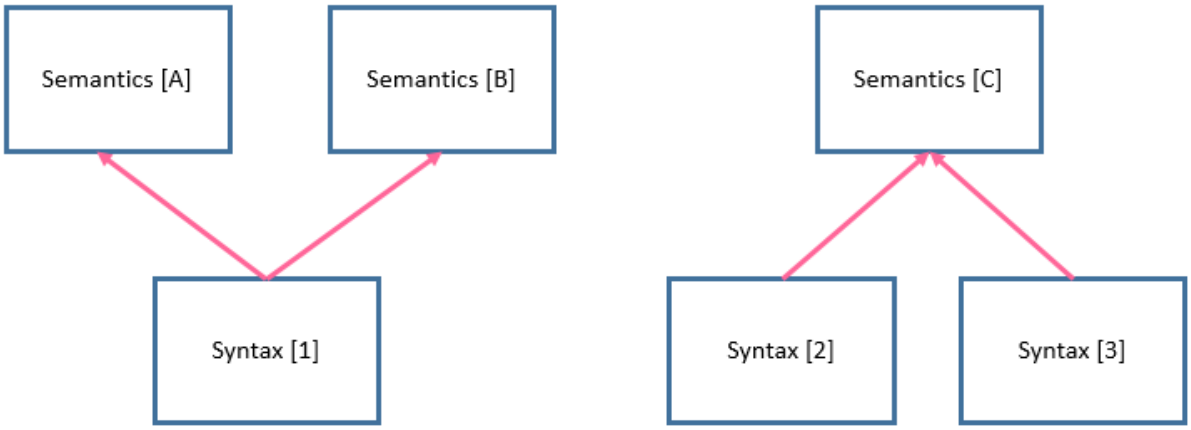


Figure 2.7 – homonyms and synonyms under the composition of languages

In 2013, my research activity concentrated on adding an abstract high-level modeling space to the existing industrial technology and guarantee a seamless process from this space and the existing technology - thus achieving an overall seamless chain from the design of an abstract software component to its generated programming component (based on PsyC), its analysis, compilation and finer control of the micro-kernel.

Results have been provided under confidential deliverable.

2.3 Contribution: Guaranteeing Correctness Under Heterogeneity

The increasing complexity of high-integrity embedded systems requires an expressive language, which is able to specify the system and its properties. In this context, however, a common error is to force a given language to capture things that the language cannot inherently express. A popular example is the C language, which has no native support to express real-time properties. This is the reason why Vincent David introduced the PsyC language [DD00; AD03] for the control-command in the nuclear domain application. The PsyC language enriches the C language with temporal properties and, then, can be used to guarantee determinism. This is possible thanks to a formal mathematical foundation and a supporting compiler, which ensure determinism and correctness from the specification of the software components to their execution on the microkernel. A similar case is represented by UML [OMGc], which initially targeted software development, without considering time. Several years after its introduction, the OMG [OMGa] standardized MARTE [OMGb], the UML profile addressing real-time designs and concurrency.

The above examples show that a single language is not sufficient to cover all concerns involved in the design and analysis of a system and its software. We have then witnessed a proliferation of languages and tools during the first decade of the 21st century.

A consequent difficulty concerns the guarantee of the system's coherence and formal proofs under the composition of the languages. The problem of the composition of languages has relevant repercussions at the tool level, in the case where several profiles or modeling languages are applied to the same model. One of the major difficulties is constituted by homonyms and by synonyms (See Figure 2.7). Homonyms have the same linguistic form, or syntax, but different semantics with respect to the target language. Synonyms have different syntaxes but share a common semantics.

To complicate the situation further, different teams with different cultures and know-how are needed in the D&D and V&V phases. Moreover, all safety norms require that the safety engineers shall be different from the system and software designers. This situation produces an immediate

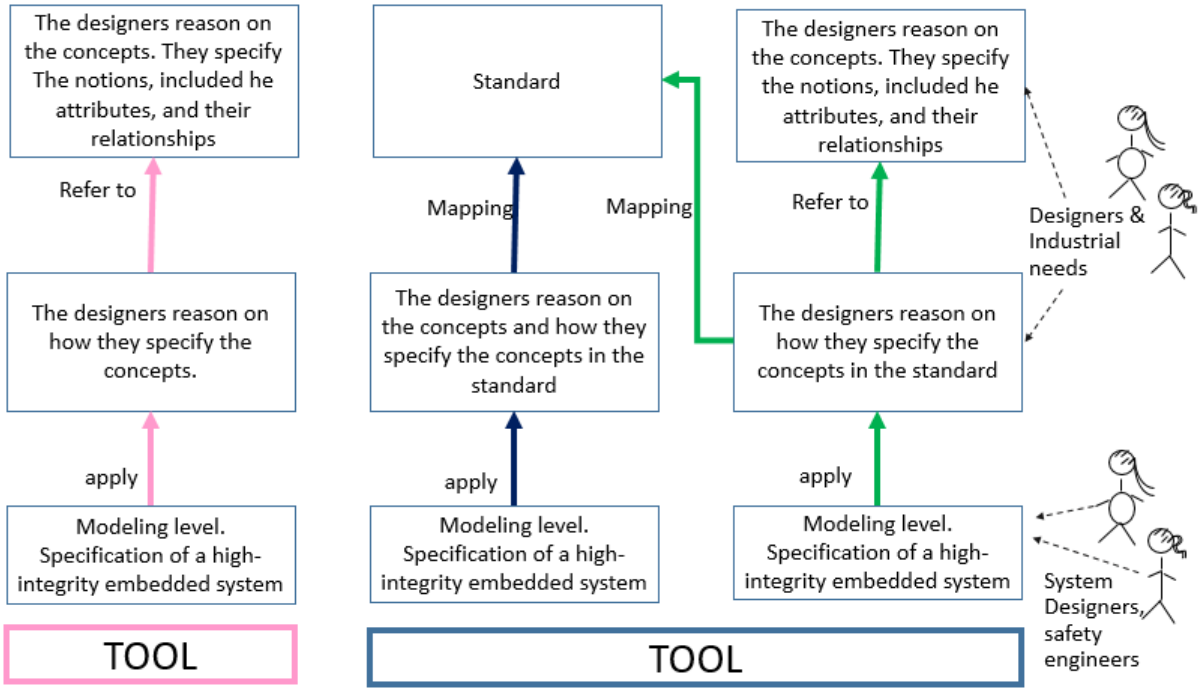


Figure 2.8 – In the Pink color, strategy (A). In the Blue color, strategy (B) and in the Green color strategy (C)

loss of control over which notion of a language is being used, generating a superfluous confusion in the analysis and verification of the model. For example, in a view of the model, a designer uses Stereotype A_{L_1} with meaning B_{L_1} in Language L_1 and in another view a designer applies stereotype A_{L_2} with meaning B_{L_2} in Language L_2 and the syntax of A_{L_1} is equal to that of A_{L_2}. Although a tool tracks the stereotype or notion applied by a designer, it is highly probable that human errors may occur in reading and interpreting the model. Errors can be generated by the fact that the designer's focus is not the language itself, but giving a correct specification of a complex and High-Integrity Embedded System under development. In doing their work, designers cannot afford to be 'distracted' by a useless complexity of the tool or by the subtleties of the language's representation in the tool. This type of human errors is precisely what a correct by construction approach wishes to avoid.

The above situation is all the more critical when we analyze the system. What is the notion that the modeling platform is considering for analysis or automatic code generation? Is it possible to guarantee the coherence and consistency of the system in the face of homonyms and synonyms? If it were possible, how much does it cost in terms of time and effort?

These kinds of questions were analyzed in my article [Esp+09]. The emphasis was placed on SysML and MARTE standards, arguing a needed common agenda between SysML and MARTE, because designers have to apply both in the system design.

In 2009, in face of this landscape, two studies, driven by Alberto Sangiovanni-Vincentelli, were conducted in parallel: one in USA and one in Europe. The two parallel studies have involved several researchers. I was part of the European team and results are published in [Pas+09]. The questions we have to fix were: *how many ways exist to merge languages together? How many ways are used to integrate tools together?*

We found that only two ways are widespread in Europe (see Figure 2.8):

(A) Specify an ad-hoc language (also known as a domain-specific language) and a tool, which are dedicated to a single concern

(B) Specify a language as an extension (or profile) of a pre-existing one - thus taking advantage of a pre-existing tool by only adding a plug-in.

If the first strategy fully meets the needs, a drawback consists in overcoming the interface between tools by ensuring that no semantic distortions occur.

The second strategy has the considerable advantage of using an existing tool and hence an expected easier verification. However, it could happen that the new language is semantically limited by the language of which it is an extension.

In that period, a third strategy emerged (in Figure 2.8, it is represented by the green color). It exploits different levels of abstraction. The notions, or concepts, their attributes and the relationships between attributes and concepts, are specified in a specific space, called metamodel or ontology (with respect to the context in which the new language is placed). Then, the language is defined as a profile, returning to the second strategy. More rigorously, a metamodel is the result of capturing concepts and rules of a specific modeling language via more or less formal means. A profile is a means to interface the new metamodel with a preexisting one. In this context, a model conforms to a metamodel/profile if the model respects the set of modeling rules defined in the metamodel and implemented in the profile [Pas+09].

This last strategy overcomes the drawbacks of strategies (A) and (B), because the concepts are defined independently of the mother language, and, thereby, gain the benefits of a domain-specific approach. Moreover, this approach exploits tool interoperability and facilitates the interface and traceability between different modelling aspects of the same system.

2.4 Conclusion

During the first years of the 21st century, the scientific community devoted an important effort to the study of High-Integrity Embedded Systems (modeling and programming languages, specification and verification of real-time and deterministic properties, methodologies and principles - such as correction by construction, separation of concerns, platform-based ...) The correct automation of “details" (whether they are linked to programming languages, to formal methods, to real-time properties ...) and the resulting new abstraction levels emerge as necessary keys to face the growing complexity of such systems.

Chapter 3

Safety for Cyber-Physical Systems

Contents

In the second decade of the 21st century, Cyber-Physical Systems (CPS) assumed an increasingly significant role in a number of disciplines, especially in Computer Science, becoming a paramount object of study in the domain of dynamic, heterogeneous systems. CPS integrate digital embeddedness with a dynamically-changing physical context of deployment, by means of a variety of sensors and actuators. The scientific community had it very clear that the CPS realm constituted a disruptively new discipline of systems engineering, whose development would require new, highly interdisciplinary development methods. As a consequence, CPS require new pedagogical learning and training methods. The CPS Education Challenge will be addressed in Chapter 5.

Precisely while all that happened, I took direct responsibility in several areas of CPS: (1) the construction of research proposals on (critical) CPS under European and French national research funding, (2) operating as Principal Investigator for the ensuing projects, in supervising MSc and PhD students on (critical) CPS, and (3) contributing to the Electronic Components & Systems (ECS) Strategic Research and Innovation Agenda (SRIA) for the ECSEL program, where in 2020 I became the co-leader of the "Quality, Reliability, Safety and Cybersecurity" core group, directly in charge of the "Safety and Resilience" and "Cybersecurity and Privacy" Challenges. All my mentioned activities address CPS under different perspectives.

In the transition from embedded to cyber-physical systems, I have especially investigated critical system properties. My research activities lead me to build a solid, rich and fruitful structural collaboration with KTH (Sweden), the University of Trento (Italy), the University of Berkeley (USA), and my workplace, CEA. In that trajectory, I have witnessed a paradigm shift from high-integrity embedded systems to safety for cyber-physical systems.

This chapter reflects these experiences and outlines the lessons that I have learnt across them, based on the direct involvement in the European roadmaps on CPS, and my technical contribution to the safety and certification of critical CPS, with a special focus on proving the compliance of the design of a critical CPS in the railway domain to the corresponding safety norms.

3.1 Context

The term 'cyber-physical systems' was first introduced by Helen Gill to broadly capture a similar meaning of the term 'cyberspace' and 'cybernetics' [Pto14]. The CPS concept arises from the cradle of Silicon Valley, and more precisely at the University of Berkeley, which subsequently became the CPS flag bearer. Although CPS have attracted the international scientific community from the beginning, it is only around 2010 that we have witnessed the CPS topic to exert maximal attraction. As a result, all research and industrial institutes are now familiar with CPS and host active research teams on CPS projects. In Europe, for example, the following research centers and universities are distinguished (non-exhaustive list): Fraunhofer, TU Dortmund, KTH, INRIA, CEA, ISAE-SUPAERO (Institute for Space and Aeronautics Engineering), Verimag, University of Trento, University of Padua. A capital example of this rise of interest happened to my institution. In 2013, during my talk "*Critical Cyber-Physical System: a contract-based approach*" [Can15a], I brought the CPS topic to the attention of my institution. At that time, it was not uncommon to explain what CPS meant and why it was necessary to pay a greater interest to the topic. Today, after 8 years, not only I do not need to explain what CPS are, but my department has created a whole new research team and tasked it to address the CPS domain.

The scientific community often makes reference to the following definition:

> *Cyber–physical systems (CPS) are integrations of computation and physical processes* [Lee07]

However, from the beginning we have witnessed CPS to have a much larger scope and scale than traditional embedded systems, as well as to exhibit a vast diversification of their constituents, which range from networking properties to connect two CPS (e.g. two semi-autonomous machines) to the data fusion of a heterogeneous sensors architecture (e.g. of a mobile robot). In 2012, the National Institute of Standards and Technology (NIST) Workshop on CPS provided a first taxonomy of what CPS involve and require (see Figure 3.1). In 2021, the NIST includes under CPS the following topics [Nat21]:

- Internet of Things (IoT);
- Industrial Internet; Smart Cities;
- Smart Grids;
- "Smart" Anything (e.g., Cars, Buildings, Homes, Manufacturing, Hospitals, Appliances)

The large scale and diversification of what is meant by CPS becomes so large that M. Di Natale and Alberto Sangiovanni-Vincentelli state:

> *There seems to be no general consensus on the definition of a CPS* [NS14]

and provocatively title the article "Are we Losing Focus on the Cyber-Physical Aspects in the CPS Research Agenda?"

In Computer Science, a clear, precise and shared vision of what is excluded from the CPS scope seems to not exist yet. However, an unclear delineated CPS boundary led to a great variety of researches and, interestingly, to a lively and multi-disciplinary interaction which does not involve any longer a single, stand-alone technology, as often occurred in the embedded system discipline.

It was in these years that I believe I achieved my fullest scientific maturation, as attested by the contributions that I have been able to make. I actively contributed to building a shared and multidisciplinary vision on CPS; I studied whether high-integrity approaches are still suitable for

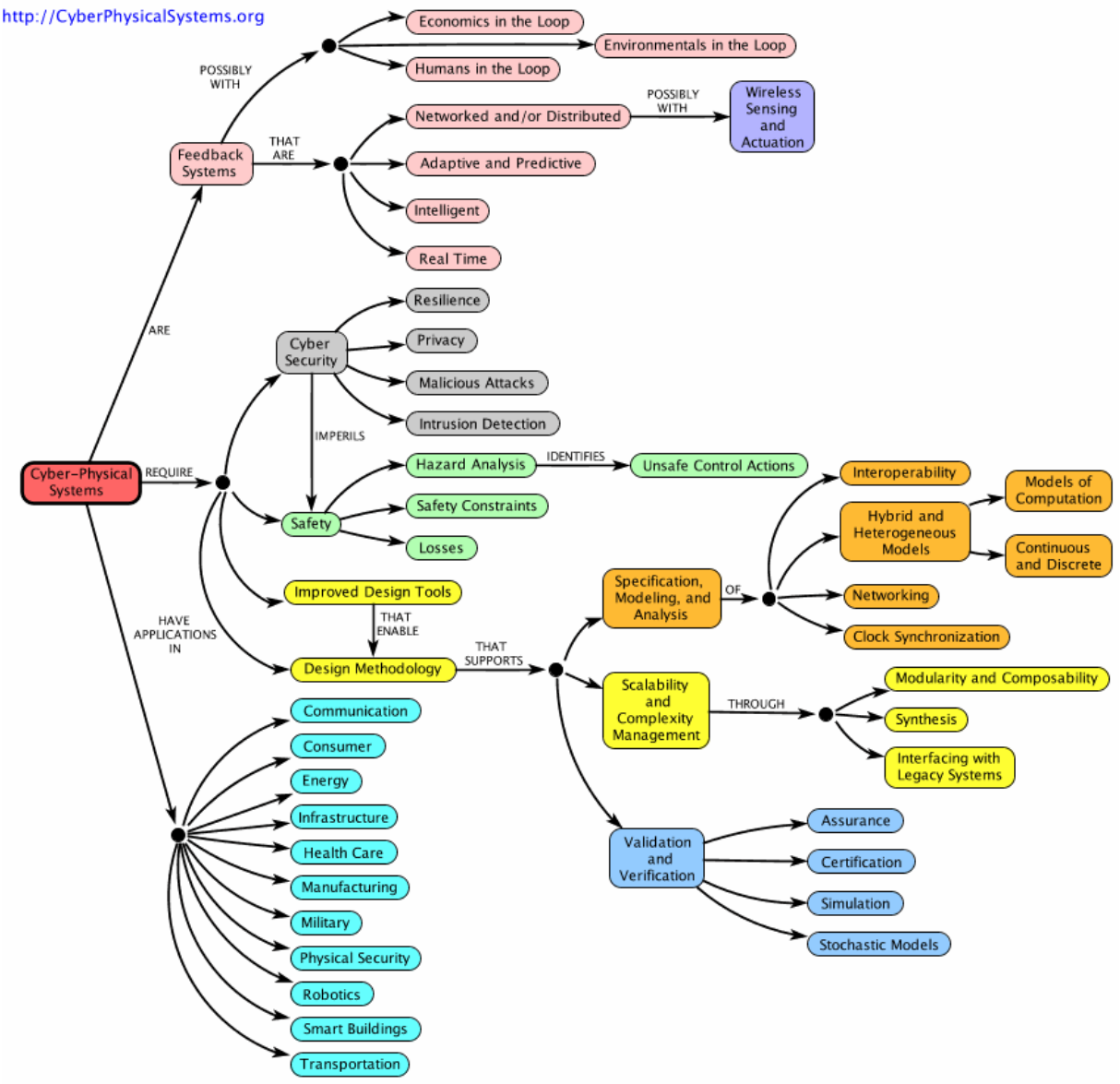


Figure 3.1 – CPS a concept map. Figure drawn from [Nat12]. Authors: Philip Asare (Bucknell University), Georgios Bakirtzis (University of Virginia), Ray Bernard (Ray Bernard Consulting Services), David Broman and Martin Torngren (KTH), Edward A. Lee (UC Berkeley), Gerro Prinsloo (Stellenbosch University) and S. Shyam Sunder (NIST).

ciritical CPS or whether new safety-related methods are needed to embrace the multidisciplinary nature of critical CPS; I introduced several research projects on CPS; and finally I established a solid, rich and fruitful structural collaboration with KTH (Sweden), the University of Trento (Italy), the University of Berkeley (USA), and my workplace, CEA.

I share with Alberto Sangiovanni-Vincentelli, Roberto Passerone and Damm the identification of the following main challenges [SDP12]:

- uncertainty, which becomes a central concern for autonomous CPS systems with AI;
- overall design flows for heterogeneous systems;
- formal requirement engineering and design space exploration;
- the verification of CPS, particularly at system integration.

In the cited article, the authors highlight the importance of a research on "design methodologies that seamlessly and coherently combine the various dimensions of the multi-scale design space" [SDP12]. The proposed way forward is based on a layered and contract-based approach. A design layered approach allows us to reason about the system with respect to different levels of abstraction and concerns, preserving all the properties discussed in Chapter 2 while a contract-based approach helps us to define the interfaces of the system's components and their properties, thus facilitating not only the reuse of components but also a seamless approach to the CPS design.

Martin Torngren and Edward Lee in [Der+13] extend a contract-based approach to study and specify real-time and safety-related properties on components. Moreover, the authors point out an important difference between a requirement-based and a contract-based specification: the first have a unidirectional nature, while the seconds have a bidirectional one. Their analysis singles out an extra dimension of contract-based approaches that best fits CPS.

In this context, my research activities contributed to solve the following open questions:

- *What is the difference, if any, between safety for CPS and for high-integrity embedded systems?*
- *Assuming a contract-based approach to the design of CPS, what will be the implication on the critical side of CPS?*
- *Is the process inherently involved by the safety norms, adapted to embrace the design methods for CPS?*
- *How can we reduce the cost of certification for CPS?*
- *How can we build joint action on CPS for a stronger Europe, at the cutting-edge of industrial technology and scientifically competitive?*

3.2 Contribution: European RoadMap on CPS

The analysis in this section reflects my experience drawn from a direct involvement within the European road-map on CPS and as senior expert for the European Commission in the selection and evaluation of European Research Projects for a stronger Europe, at the cutting-edge of industrial technology and scientifically competitive.

In 2017, William Bonvillian, lecturer at MIT and former director of the MIT Washington Office, published in [Bon17] the article *"The rise of advanced manufacturing institutes in the Unites States"*.

The article discusses the Great Recession in the United States over the period 2007-2008 in relation to the "Advanced Manufacturing" sector of industry. In the article, the author delves into the causes of the great recession in 2007-2008. His analysis shows that it was not only the result of an economic crisis, but of a structural change resulting from the inexorable decline in traditional manufacturing. It is then in that context that B. Obama, who was elected President of the United States in 2008, found himself facing the worst economic crisis since the 1930s.

Under his first presidency, the main actions target a firm establishment of an R&D in support to manufacturing technologies, through a political agenda, new legislation, funding of research and innovation programs, and the creation of ad-hoc institutes, such as MIIs (Manufacturing Innovation Institutes).

Already in the early years of the Obama presidency, we have witnessed a shift of the paradigm, no longer centered on a 'single research award' but on a large, complex and mixed collaboration between small, medium and large enterprises, universities and research centers. The main objective of these actions was the creation of new and disruptive technologies in manufacturing as well as 'build workforce skills' [Bon17].

CPS originated in this period as Silicon Valley's response to the above-mentioned political demand. The first examples of CPS applications concern manufacturing, also called Industry 4.0 and are referred as "the fourth industrial revolution". In **2009**, CPS were formally recognised [Nat].

In Europe, the first political interest towards CPS occurs in 2012 with "The German Agenda CPS" [Gei+12]. In this period, the German and European interest concerns the automobile market with the transition from a car with advanced driver-assistance systems (ADAS) functionality to autonomous cars. At the same time, the political emphasis is on Industry 4.0. For example, many National European governments and the European Commission support the Productive 4.0 project, piloted by Infineon (Munich, Germany), with a consortium of more than 100 partners. In Productive 4.0, I am first the Principal Investigator on behalf of CEA and then, once the project has been accepted, the CEA project manager by leading three research teams of my department.

If from the beginning, the CPS constitute a disruptively new discipline with respect to embedded systems, what are the main differentiating elements among CPS and embedded systems? If we pay attention to critical CPS, then we can reformulate the above question as follows:

What is the difference, if any, between safety CPS and high-integrity embedded systems?

We could find an answer by investigating the technological context, which characterizes the two disciplines. It is briefly summarized in Table 3.1 (the table does not intend to be exhaustive).

In the high-integrity embedded systems, the processors are very often mono or dual-core and a critical functionality is deployed on a dedicated core. It involves a simplification in demonstrating that high-level safety properties are correctly preserved during the execution at the platform level. Within CPS, the technological landscape changes significantly. Many and multi-core processors are also adopted for critical applications. In doing do, the scientific and industrial community devote an important effort to mixed-criticality, i.e. functions with different level of safety deployed on the same

High-Integrity Embedded Systems	Critical CPS
Mono/dual core + multi/many core	Multi to Many core
Safety critical functionality on dedicated core	Mixed criticality
Correctness-by-Construction	Correctness-by-Design
Stand-Alone Technology	Connected and Interoperability Technologies
System Engineering	Collaborative System Engineering

Table 3.1 – HIgh-Integrity Embedded Systems and Critical CPS

processor, which becomes an emerging and attractive topic. In this context, between 2013 and 2016, the European Commission supports three projects:

- Design of embedded mixed-criticality CONTRol systems under consideration of EXtra-functional properties (CONTREX) [CON];
- Distributed REal-time Architecture for Mixed Criticality Systems (DREAMS) [DRE];
- Probabilistic real-time control of mixed-criticality multicore and manycore systems (PROXIMA) [PRO].

CONTREX, DREAMS and PROXIMA create the European Mixed-Criticality Cluster (MCC) [Tho16], in which I participated as an expert on behalf of the European Commission. Moreover, I introduced the EMC2 European project (Embedded multi-core systems for mixed criticality applications in dynamic and changeable real-time environments) [EMC], with my colleague Marc Duranton.

Over the years, the growing attractiveness and expectation towards CPS led the European Commission to fund several road-map projects, including (non exhaustive list):

- Cyber-Physical European Roadmap & Strategy (CyPhERS), 2014 2015 [CyP14];
- ToWards Cyber-Physical Systems Engineering Tools, Interoperability Standardization (CP-Setis), 2015-2017 [CPSa]; and
- Platform4CPS [Pla], 2016-2018.

I have had the opportunity to integrate the aforementioned road-maps as external expert or senior expert on behalf of the European Commission.

CyPhERS (Cyber-Physical European Roadmap & Strategy) aims to:

- define the main properties of CPS; and
- identify the main challenges and expected innovations.

Table 3.2 summarizes some results of Deliverable 5.1 of CyPhERS [CyP13]. The project team identifies some properties which characterize CPS and, maybe, will play a leading role in the coming period. Among the CPS-related properties, CyPhERS identifies artificial intelligence, which becomes the real protagonist today. Among the challenges identified, the project extends the analysis not only to the scientific, technological and economic fields, but also to education, legal and social ones. Liability plays a central role in bringing autonomous AI products to market. Therefore, it is necessary to establish traceability of responsibility as well as a structural coordination between the different

CPS Property	Meaning
Context awareness	perception of its own operational context in terms of environmental situations and internal resource conditions for determining the best courses of actions in complex scenarios.
Cognitive computation	functional properties of a system in regard to the reasoning of its own status and thereby the planning for its upcoming behaviours ranging from high level missions to lower level tasks and actions.
Autonomy	the system's property of being sufficiently independent in controlling its own structural and behavioural properties.

Table 3.2 – The content of the table is extracted from [CyP13]

entities that intervene in the design, development and production of an autonomous CPS and their maintenance over years.

Finally, the analysis and vision brought to the CyPhERS project still constitute a cornerstone of this new discipline and either have found a solution or are still topical.

Among the great achievements of CP-Setis (ToWards Cyber-Physical Systems Engineering Tools, Interoperability Standardization), I would like to mention the "Strategic Agenda on Standardization for Cyber-Physical Systems" [CPS17]. The CP-Setis project argues that "there is not a coherent approach to CPS Standardization internationally". The statement arises from a detailed analysis of the landscape of safety regulations and their relationships, which is represented in Figure 3.2.

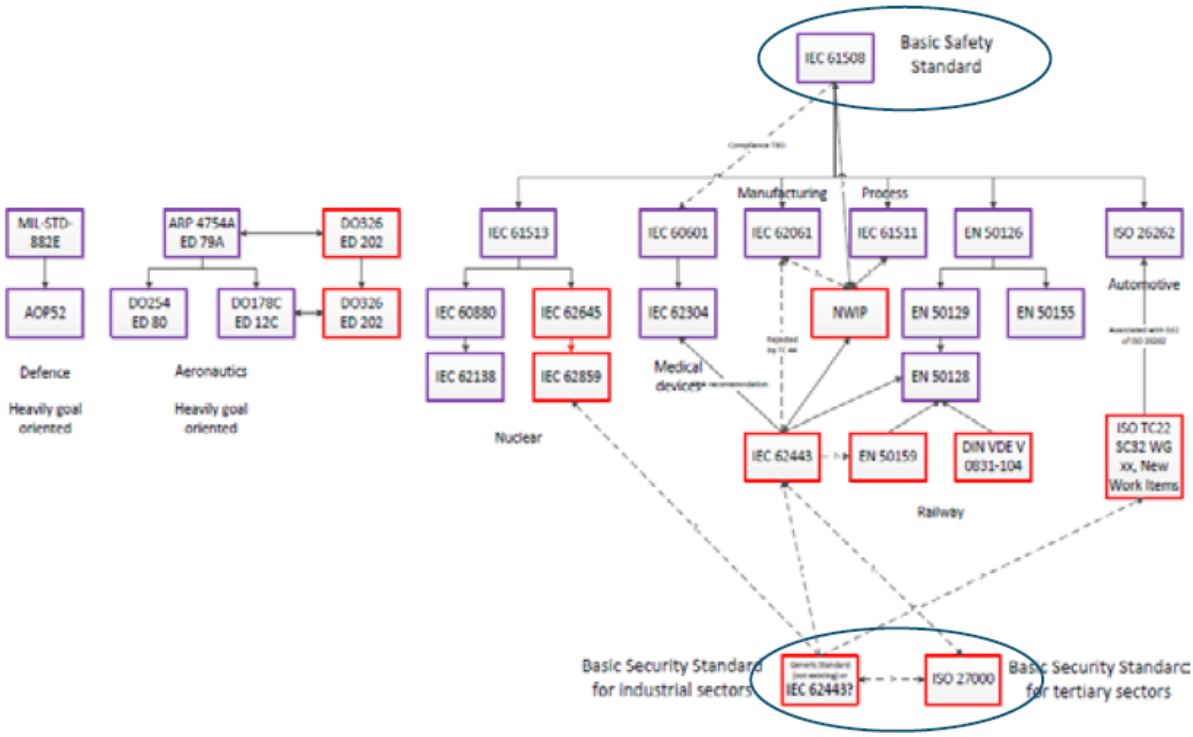


Figure 3.2 – Relationship between safety regulations. Figure extracted from [CPS17]

Their analysis concludes that the landscape is almost fragmented. This is the reason why the project provides strategic directions to further develop the safety-related standards by maintaining "in a consistent, harmonized manner a multi-standards, specification and guidelines". In doing so, the

CP-Setis project identifies and promotes policies based on an "onion leaf" involvement of experts and stakeholders, as shown in Figure 3.3. The expected outcome of the promoted approach is the sharing of consensual safety-related international norms during the design, development, production and maintenance of products on the market.

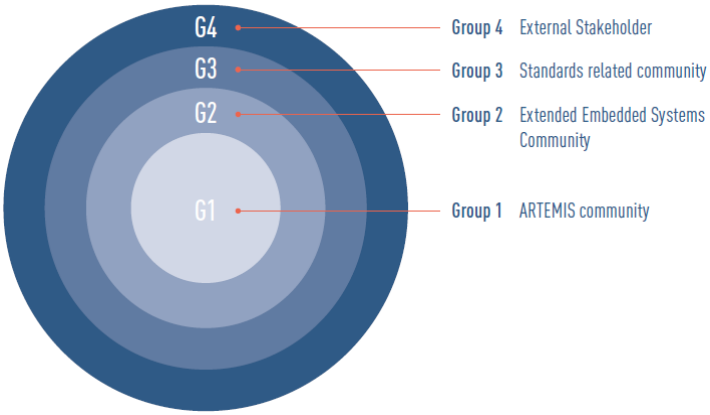


Figure 3.3 – Involvement of experts and stakeholders. Figure extracted from [CPS17]

Platform4CPS [Pla] is among the most recent roadmaps for CPS and then the most mature one, because it has benefited from the vision carried out by the CPS roadmaps, and the research and industrial achievements on CPS, which have flourished in the preceding years.

In Platform4CPS, topics such as the impact on society and the legal aspects of AI, firstly introduced in CyPhERS [CyP13], are here more developed. Moreover, the Platform4CPS roadmap highlights safety and the ethical aspects due to the intrinsic nature of AI. In this regard, the roadmap states: "Ethical issues of AI considering transparency end the need for ethical training for engineers" [Tho+18].

If the first years of the second decade of the 21st century see the European political agenda oriented towards a joint development of industrial competitiveness and innovation, which can be able to compete with the United States and China on the evolution of the international market, the last years are marked by the orientation towards the European sovereignty as a major concern. Then, all issues related to ensuring cybersecurity, confidentiality and privacy play a decisive role in agendas such as Platform4CPS and ECS-SRIA (Electronic Components & Systems (ECS) Strategic Research and Innovation Agenda (SRIA)) [ECS21], where since 2020 I became the co-leader of the "Quality, Reliability, Safety and Cybersecurity" core group, directly in charge of the "Safety and Resilience" and "Cybersecurity and Privacy" Challenges (see Figure 3.4)

My involvement in the ECS-SRIA contributed to the setup of the ECSEL, PENTA/Eureka and Xecs industrial European research framework programs [ECS; Pen; Xec]. Those programs aim to "provide the Electronic Systems & Components community with a specific support capability to address the huge challenges created by the rapid development of the global digital economy. Our vision is to offer a program that is open to all the elements of the ECS Community; large enterprises, small and medium enterprises, research and technology organizations (RTOs), and universities" [Pen].

The ECS-SRIA 2021 edition [ECS21] was produced during the year 2020. Project Proposals that address the 2021 call, such as PENTA [Pen] and ECSEL [ECS], should respond at least in part to challenges that are described in ECS-SRIA 2021. Challenge "Safety and Resilience" is part of the "Quality, Reliability, Safety and Cybersecurity" chapter. The challenge mainly concerns "the development of safe and resilient autonomous cyber-physical systems in dynamic environments, with a continuous chain-of-trust from the hardware level up to the applications that is involved in the accomplishment of the system's mission, including AI. Our vision takes into account physical

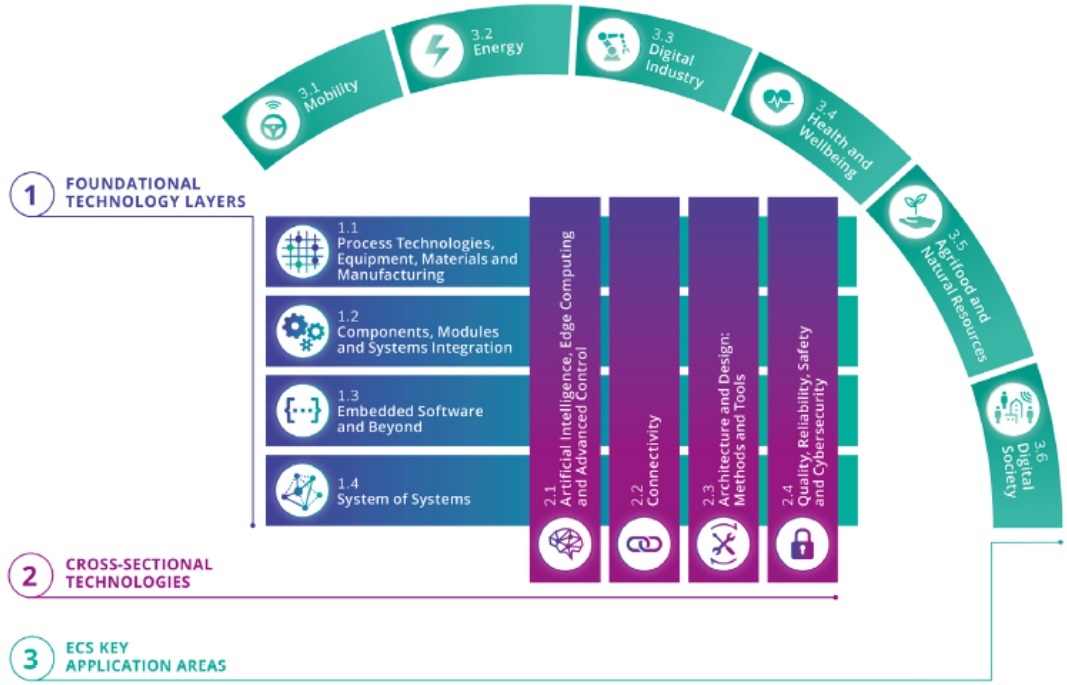


Figure 3.4 – ECS-SRIA overall vision. Figure drawn from [ECS21]

limitations (battery capacity, quality of sensors used in the system, hardware processing power needed for autonomous navigation features, etc)" [ECS21]. The challenge mainly highlights trustability, i.e. the human ability to trust an AI-driven product, which is involved by inscrutability of AI.

More precisely, the main objectives of Challenge "Safety and Resilience" are:

- Safety and resilience of (autonomous AI) systems in dynamic environments, which embraces concepts and principles for trustable integration and the V&V of intelligent functions in systems/products under uncertain and/or dynamic environments (non exhaustive list);
- Modular certification of trustable systems and liability, where the expected outcome is a clear traceability of liability in the case of damage or accident;
- Dynamic adaptation and configuration, self-repair capabilities, (decentralised instrumentation and control for) resilience of complex and heterogeneous systems;
- Safety aspects related to the human/system interaction, where the expected outcome is to ensure safety for the human, system and environment during the nominal and degraded operations in the working environment; and
- Ensuring both safety and security properties, where the main expected outcome is to ensure compatibility, adequacy and coherence in the joint use of the promoted security solutions, and the safety levels required by the system or its components.

The last goal in the above list is introduced in the "Cybersecurity and Privacy" challenge. It is introduced here because is strictly linked to safety.

A capital example of the problems we are facing concerns ensuring the safety mechanisms under both privacy and cybersecurity properties of a full-autonomous vehicle. We created a scenario where the full-autonomous vehicle is maliciously hacked and potentially used for an attack against civilians. To have a control-command in the vehicle which can be used to an external authority in the case of an emergency could potentially violate the privacy property, because the driver is always traced.

In ECS-SRIA 2022, which is under elaboration, Challenge "Safety and Resilience" embraces solutions based on low environmental impact. In the 2021 version, it was decided to include this topic in only one place, the AI chapter. In the new edition, however, the sustainable impact is emphasized and could then be introduced also in the challenge that I led.

3.3 Contribution: Safety and Collaborative Engineering

The pervasiveness of information technologies has had recently an increased impact on several aspects of our life. This trend has been facilitated by the possibility of integrating more and more functionalities into the same device and/or system, and by making them cooperate in a network to provide sophisticated services. This section reflects my experience on ensuring the safety of critical cyber-physical systems. The section addresses the following question:

Is the process inherently involved by the safety norms adapted to embrace the design methods for CPS?

In this section, there are two main contributions. The first contribution concerns how we can ensure safety to a critical cyber-physical systems in the railway application domain, by guarantying compliance to the related safety regulations. The second contribution embraces a collaborative engineering approach to the design and development of critical CPS and studies how ensuring safety-related properties with respect to compliance with safety railway regulations. The main difficulty to overcome is that safety norms are structured in sequential top-down steps, as the first contribution shows. In the second contribution, we study which safety normative steps can be made iterative and how, in order to reduce the cost of certification. Both contributions assume a contract-based approach to the design of CPS. Finally, both contributions took advantage of industrial collaborations. Results are then tested on industrial case studies in the railway domain, although the underlying methodology could be generalized to other application domains.

How safety regulations work Before introducing the contributions in detail, let me briefly explain the safety process of a critical CPS. The development of critical CPS involves the interplay of many different disciplines and, therefore, becomes particularly complex.

In this context, the industrial and academic communities are paying increasing attention to safety issues. Because safety concerns may involve profound changes in the architecture of a system, the values of the safety attributes must be calculated as soon as possible during design and development. This is the case, for example, for the *Safety Integrity Level* (SIL) [IEC00], which is related to the degree of failures that the system must be able to tolerate.

The industrial development of a critical CPS involves its successive refinement down to the program code. A refinement produces logical sub-systems. Then, safety analysis is performed iteratively on the subsystems, by following the norms related to the considered application domain. Such techniques exploit both quantitative analyses, i.e., objectively measurable analyses, and qualitative analyses, which, for the most part, depend on the safety engineers experience. The result is in the form of documentation which includes the values of the safety attributes. Finally, the whole process is certified by a third authority.

Even excluding certification, this process presents many difficulties. For example, it must guarantee that the produced documentation on the architecture and the safety analysis is coherent, and that the

safety analysis documentation is correct with respect to the architecture. In addition, safety norms clearly point out the adoption of semi-formal languages as a *means* to improve the safety analysis.

How to apply the CENELEC safety regulation The first contribution of this section concerns a contract-based methodology on the *Event Recorder* system, a real industrial case study in the railway application domain. Its main functionality is to periodically memorize the state of the system, so that, in case of an accident, the potential causes could be extracted and analyzed.

Table 3.3 introduces the main steps of the promoted contract-based methodology. It is based on the IEC61508 and EN50126 standards [IEC00; CEN99] and can be only used in "simple or self-evident" systems [CEN99]. In Table 3.3, the first two columns, Architecture and Safety Analysis, represent the two domains, design and safety, which must be addressed by different teams in compliance with the rules described by the corresponding standards. The second column specifies the safety analysis with respect to the architecture level. The third column, Achieved Safety Values, describes which safety values we have identified for each level of architecture. The values of these parameters are set with respect to the EN50126 standard [CEN99]. Finally, the fourth column provides the constraints given by the safety objectives that the level should meet.

Architecture	**Safety Analysis**	**Achieved Safety Values**	**Constraints given by Safety Objectives**
(a1) (Safety-Related) Event-Recorder System	(a2) Preliminary Hazard Analysis	(a3) SIL	(a4) The architecture should be redundant to at least one fault
(b1) (Safety-Related) Event-Recorder System with Architecture 1oo2 (see Figures 3.5 and 3.6)	(b2) Reliability Analysis	(b3) PFH, MTBF	(b4) the PFH and MTBF values should correspond to the SIL value; the architectures (e.g. (b1) and (c1)) should always meet the MTBF value
(c1) (Safety-Related) Event-Recorder Subsystems: Power subsystem and Event-Recorder Functional Subsystems	(c2) Reliability Analysis	(c3) MTBF for each analysed subsystem (e.g. Power and Event-Recorder Functional subsystems)	(c4) the pair of the MTBF values should meet the MTBF value of Event-Recorder system

Table 3.3 – Overall Methodology

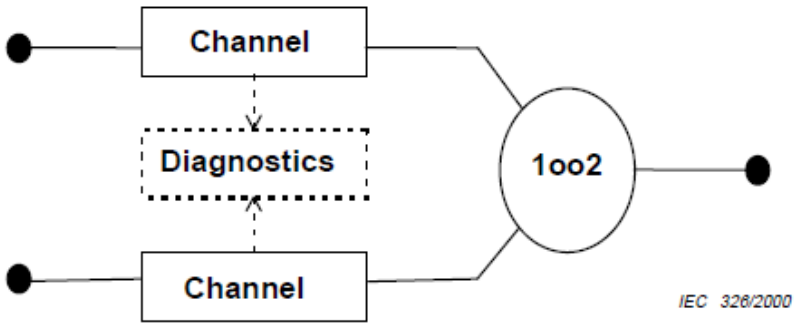


Figure 3.5 – The 1oo2 physical block diagram. Architecture 1oo2 resolves Equation 3.1. Figure extracted from IEC 61508 [IEC00])

The first column shows the refinement of the architecture: from the event-recorder to its subsystem. Once calculated the SIL that the systems shall meet, the designers must provide an architecture redundant to at least one fault among these architectures proposed by the norms. In our use case, we adopt Architecture 1oo2, represented in Figures 3.5 and 3.6, for three main reasons. First of all, it is

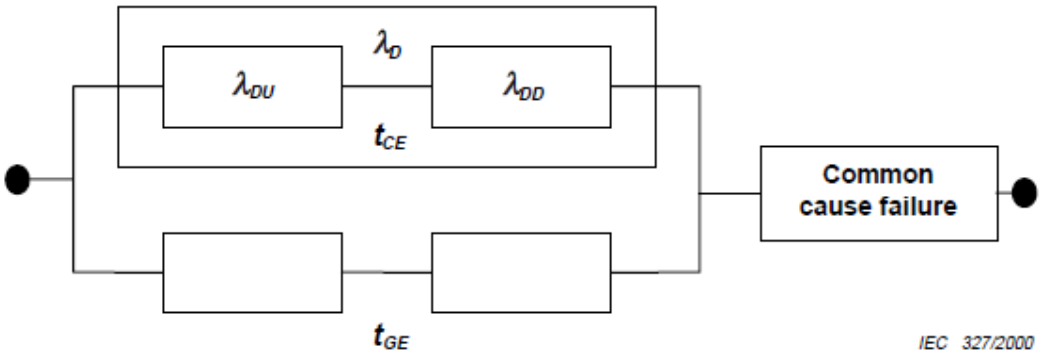


Figure 3.6 – The 1oo2 reliability block diagram. Architecture 1oo2 resolves Equation 3.1. Figure extracted from IEC 61508 [IEC00])

easy to realize. Secondly, we skip the well-known problems due to the critical circuits when three or more units are deployed on the same system. Finally, Architecture 1oo2 is an economically viable solution.

The IEC61508 standard [IEC00] introduces the mappings between *Probability of Failure per Hour* (PFH) and SIL: The system satisfies the quantitative requirements of SIL level if the MTBF value belongs to the PFH real interval corresponding to SIL 4 [IEC00].

Equation (3.1) defines PFH in accordance to the IEC61508 standard [IEC00]. Architecture 1oo2, in Figures 3.5 and 3.6, resolves Equation 3.1 [IEC00]. The meaning of the acronyms of the equation is extracted by the IEC61508 standard [IEC00] and introduced in Table 3.4:

$$PFH = 2((1-\beta_D)\lambda_{DD} + (1-\beta)\lambda_{DU})^2 t_{CE} + \beta_D \lambda_{DD} + \beta \lambda_{DU} \tag{3.1}$$

PFH is given by the sum of three factors (from the left to right):

- the fraction of undetected failures that have a common cause (β) times the undetected dangerous failure rate (λ_{DU});
- the fraction of those failures that have a common fault and that are detected by the diagnostic tests (β_D) times the detected dangerous failure rate (λ_{DD});
- average time in the state of non detected failure times occurrences of detected and undetected failures $2((1-\beta_D)\lambda_{DD} + (1-\beta)\lambda_{DU})^2$

The last column of Table 3.4 shows the safety values, which the architecture must meet. They represent our industrial safety requirements. In the table, the first four values are determined on the basis of the safety engineer experience. They are qualitative values that range in a fixed interval, given by the IEC61508 standard [IEC00]. The last four values depend on the value of the MTBF, which is equal to $\frac{1}{\lambda}$ in most cases.

Figure 3.7 shows the MTBF value that the Safety-Related Event Recorder system must meet to be safe. The MTBF value is obtained for the values in Table 3.4 (last column), and it is given from the intersection of the curve line with the threshold between the SIL3 and SIL4 [IEC00]; that is, 1,100,000. The MTBF value and the values in Table 3.4 are used in Equation 3.1 - thus calculating the PFH value. Finally, as described above, the system satisfies the quantitative requirements of SIL level if the MTBF value belongs to the PFH real interval corresponding to SIL 4 [IEC00].

Param.	Meaning	Value
PFA	Probability of Failure per Hour	
β	The fraction of undetected failures that have a common cause	5%
β_D	The fraction of those failures that are detected by the diagnostic tests, the fraction that have a common cause	2%
DC	Diagnostic coverage	95%
$MTTR$	Mean time to restoration (hour)	$1\ h$
$MTBF$	Mean Time Between Failure	
λ	Failure rate (per hour) of a channel in a subsystem	$\frac{1}{MTBF}$
λ_{DD}	Detected dangerous failure rate (per hour) of a channel in a subsystem	$\frac{\lambda}{2}DC$
λ_{DU}	Undetected dangerous failure rate (per hour) of a channel in a subsystem	$\frac{\lambda}{2}(1-DC)$
λ_D	Dangerous failure rate (per hour) of a channel in a subsystem,	$\lambda_{DU}+\lambda_{DD}$

Table 3.4 – The IEC61508 standard parameters [IEC00] and their industrial values

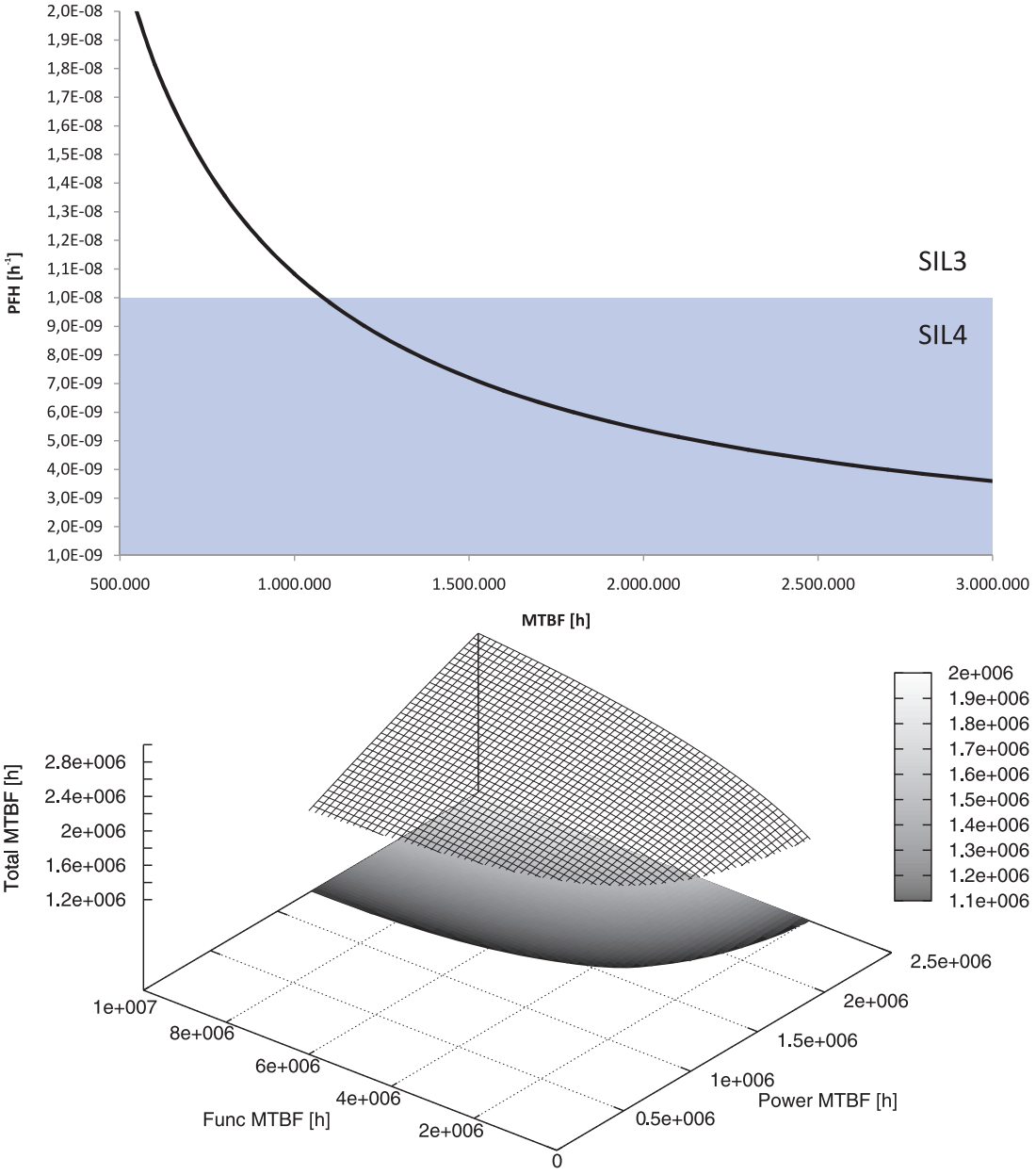


Figure 3.7 – On the top, the MTBF value for the Safety-Related Event Record System. On the bottom, relationship between the MTBF value for the system and the MTBF values for its subsystems

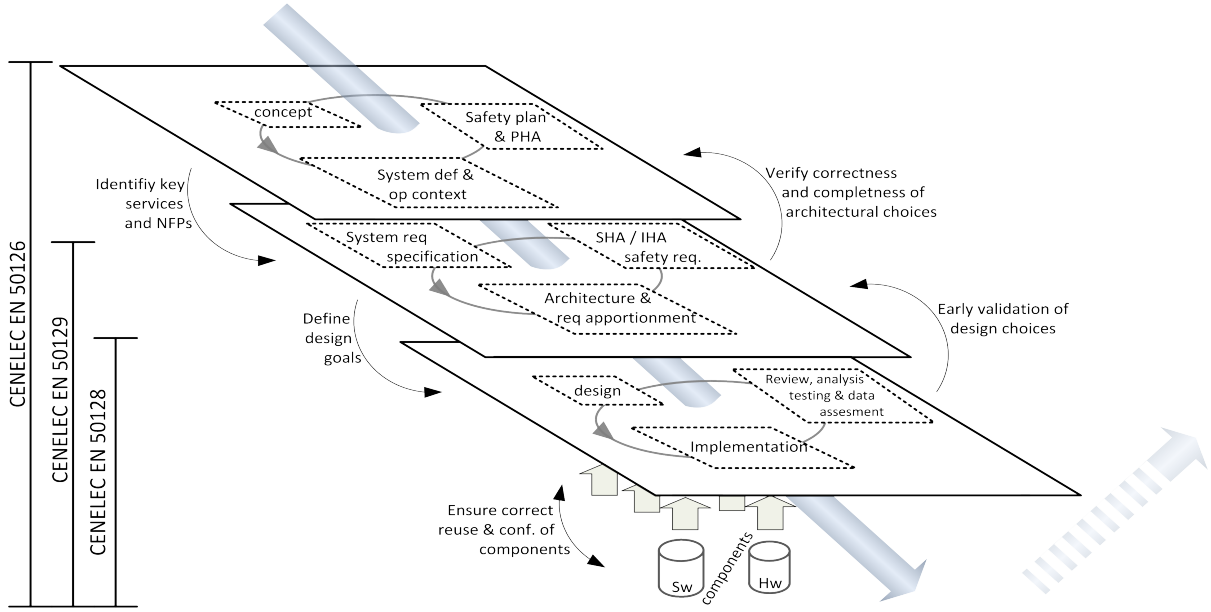


Figure 3.8 – Combining Collaborative Engineering and the CENELEC process

Limits of the sequential approach when considering a CPS collaborative engineering design The proposed approach is strictly top-down: the system is refined down to its subsystems and components. Each level provides safety constraints, which the related subsystem must meet. The values are calculated with a bottom-up procedure and must meet the expected safety constraints. However, as we analyze later, a strictly top-down approach shows some limits when considering collaborative engineering, a suitable approach for the design and development of CPS, where multi-disciplinarity is predominant and has a leading role.

Collaborative Engineering and Safety Regulations: from sequential to iterative steps The Observatory of Automated Metros (UITP) states that "in 2013 there are 674 km of automated metro in operation". This number is expected to triple in the next 10 years to reach 1,800km in 2025 [Obs13]. Still, outside the mass urban transportation segment, the sector is heavily challenged in terms of competitiveness and attractiveness by other transportation industries. This is mostly visible to the general public through the travelers' demands of further improvements regarding speed and comfort, but also, punctuality, availability and reliability. From a business perspective, this means a heavier and heavier pressure from the market for a leap forward in terms of performance, infrastructure and of course, safety. To this end, the railway industry invests almost 1000 Million euros each year in research and development [1] and there are then strong expectations for the stakeholders to deliver cost-effective services for intermediate and final users: more intelligent, integrated and autonomous systems, safe access to vehicles, reduction of the CO_2 footprint, etc.

Figure 3.8 illustrates the overall proposed methodology and process, which is compliant with the first six phases of the CENELEC safety standards [CEN99; CEN11; CEN03]. The proposed methodology and process are divided into three major stages:

1. System Definition: Capture and structure the *user needs* (i.e. the system's operators and its passengers) into a set of *services* that the system shall offer under safety and performance constraints. Perform preliminary hazard analysis.

[1] http://www.unife.org/research/overview.html

2. System Requirement and Architecture: Define the *functional and system architecture* that fulfills the previously defined *services*. Perform system hazard analysis.

3. System Design and Implementation: Define the *physical architecture* with a breakdown into hardware and software components that will implement the *functions* previously defined. Perform testing and data assessment.

An important aspect of the process is its inherently iterative nature. As shown in Figure 3.8, each of the three major stages have micro iteration cycles between their respective engineering activities. The process also supports macro iteration cycles across the major stages. This iterative nature reflects the interactions that happen in the real world between the engineering activities. It also acts as an indicator of two underlying key points:

1. The collaborative nature of the process across engineering activities.

2. The continuous verification of the engineering choices.

These two points are not directly captured by the process and yet they are instrumental to the success of designing a cyber-physical system such as the one presented in our case study.

Collaborative engineering At each stage, the involved teams share a common objective. They produce engineering artifacts that hold different information about the system under study. For example, the overall objective of the first stage is to define services that the system shall offer to end-users (operator and passengers). These services are constrained by non-functional properties, including beyond others, safety and performance ones. Still, the produced artifacts are not entities that exist in a vacuum independently from the rest of the world. All of them have relations with the others. For example in the first phase, there must be traceability relations between the preliminary hazard analysis and the identified services. With more and more complex systems, it is critical to capture and manage those relations.

Continuous verification The stages and activities presented in the process above can have some concurrent overlaps and therefore have continuous feedback loops. The activities in a stage can shed new light on the decisions taken in its upper stage, leading to the early detection of errors in the design or in the implementation. For example, a performance requirement that must be satisfied by an operational scenario is derived at the end into constraints on the performance of some actuators or software functions. Particular implementation choices can lead to the infeasibility of earlier design decision, these kinds of inconsistencies shall be detected and reported. In an iterative process supported by the appropriate tools, the continuous verification of the engineering choices precisely consists in the examination of the input artifacts at all stages and activities, as well as the maintenance of consistent relations between them.

Relating engineering artifacts The critical element for an efficient collaborative engineering and its continuous verification is the management of the relations between the engineering artifacts. To support it, the proposed methodology adopts a technique able to semantically relate engineering data which are exposed by the tools and used by the engineers across technological spaces. This technique is called "federation" and relies here on the Semantic Web [W3C] and Linked Data [Lin]. *In the use case, we exploit a web interface as a means to access data from requirements, safety hazards, functional architecture, etc... The federation enables a controlled acquisition of the data by preserving the full property (of data).* The data shared by all the engineering teams and their relations can be leveraged to perform business-wide analysis and trade-offs. This capability reinforces the engineering and business decisions, ultimately leading to more value for the end users.

Comparison between a strict top-down and a collaborative engineering approach Today, railway (critical) artifacts are submitted to a rigorous process to ensure the expected safety integrity levels. The high cost, related to this process, ought engineers to a strict control in the introduction of mechatronic innovations. In many cases, then, two artifacts present small differences.

At the same time, the high competition between constructors demands innovative functionality and a whole cost-reduction of the artifact. A top-down approach, as described in the CENELEC norms, is not competitive anymore. Often, constructors combine top-down with bottom-up approaches, for example in the reuse of software, system design, electronic or mechanical devices. Interfaces between components acquires a relevant role to fix the objectives of performance, innovation, cost and safety. A change in the interface could impact other components and *de facto* prohibit a correct reuse, thus increasing the overall cost. The current practice synchronizes information between teams at precise time in order to avoid this.

The promoted methodology does not aim to change the adopted tools or the know-how of each company, but only to suggest a possible solution to combine collaborative engineering with the CENELEC rigorous process by reducing the go-to-market timing. The proposed approach can be extended to other application domains. First preliminary results, showed on the railway domain, received a positive and encouraging feedback by an avionic company.

3.4 Conclusion

In the second decade of the 21st century, the scientific community address Cyber-Physical Systems (CPS) and defines the US and European road-map on CPS at the cutting-edge of industrial technology and scientifically competitive. During that period, I took direct research responsibility.

Chapter 4

Resilience for Autonomous Cyber-Physical Systems

Contents

The embedding of AI into Autonomous Cyber-Physical Systems demands new design and development technologies to better support the learning phase, the adaptive maintenance of the learned model, and the traceability of output to requirements.

4.1 Context

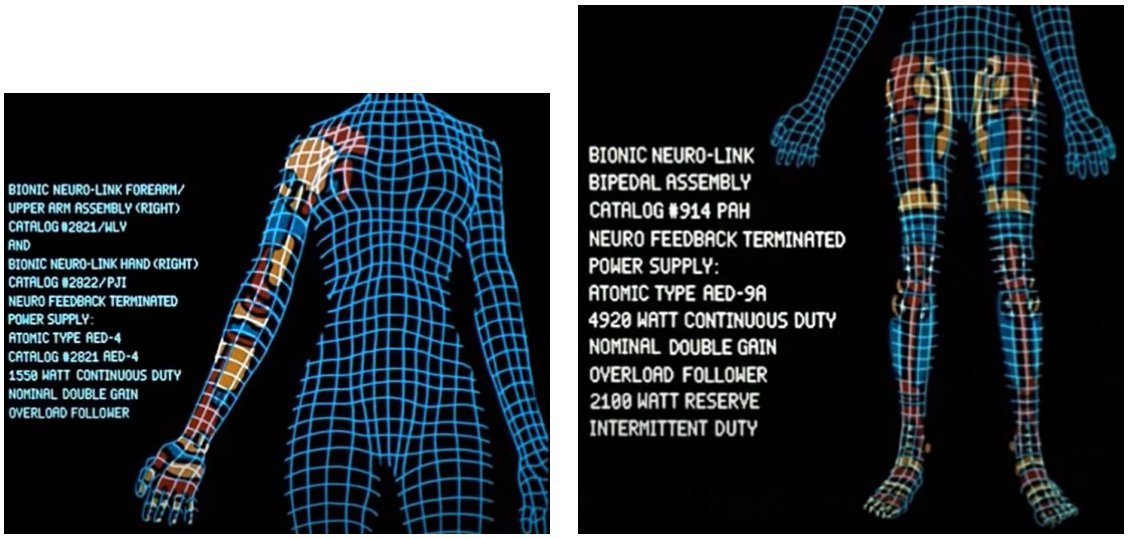
> *"MARVIN: Here I am, brain the size of a planet and they tell me to take you up the Bridge. Call that job satisfaction? 'Cause I don't."* [Dou78]

In the seventies, Douglas Adams wrote *The Hichhiker's Giude To The Galaxy* [Dou78], for the BBC radio broadcast. He imagines Marvin, a humanoid robot full of sensors, actuators, AI, capable of feeling emotions and making judgments.

> *"You watch this door" - continues Marvin - "It's about to open again. I can tell by the intolerable air of smugness it suddenly generates."* [Dou78]

In 2021, have we reached this level of technology?
Almost, but not quite.

In the late seventies, almost all European and American families owned a television set and TV series are prospering. Generation X grows up seeing *"The Bionic Woman"* (first season in 1976), a woman to whom, following a very serious paragliding accident, a bionic arm is implanted, both bionic legs and a bionic ear. Ten years do not pass and Generation X can watch in 1982 *"Knight Rider"*, also known as *K2020*, where the protagonist leads a futuristic machine, named KITT (acronym for "Knight Industries Two Thousand"), having embedded processors and AI, and the series *"Murder, She Wrote"* where during the first season, in 1984, explicit reference is made to autonomous vehicles, with the usual over-optimistic American enthusiasm:



(a) *Bionic arm* (b) *Bionic legs*

Figure 4.1 – Excerpt from the original soundtrack of *The Bionic Woman* series `https://www.youtube.com/watch?v=v7sqvc-k300`

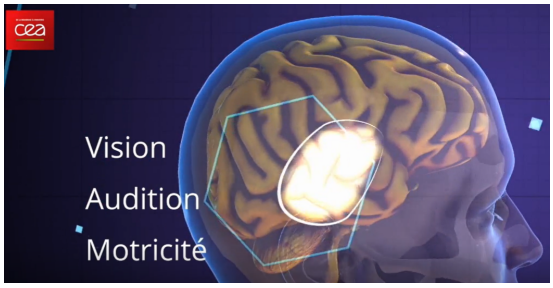
> *[Engineer] Remote-controlled car? Nothing new about that.*
>
> [Jessica Fletcher] *Yes, but what would a car like that be used for?*
>
> [Engineer] *Automotive testing, military maneuvers, motion picture stunts. And it would be a great boon for people like you, Jess. People who don't drive. You could get in your car, program your destination, and the computer would take you there.*
>
> [Jessica Fletcher] *Sounds wonderful. If a little far-fetched.*
>
> [Engineer's Wife] *No, no. I remember those designs. It was a wonderful concept. No more highway accidents. This car had a built-in self-protection system against collisions.*
>
> [text extracted from *"Murder, She Wrote"*, the first series, episode 8]

Figure 4.2 – KITT - Knight Industries Two Thousand. Excerpt from the *"Knight Rider"* original series. `https://www.youtube.com/watch?v=hfRiedxPQhs`

Half a century later, autonomous vehicles have a constantly growing trend (with all the problems related to liability and to technology in *stricto sensu*) as well as exoskeletons, whether controlled by a muscular impulse for the transport and manipulation of loads (Figure 4.3 (a)), from a chest pulse, in the case of paraplegic people (Figure 4.3 (b)), or from a neuronal pulse, in the case of quadriplegic people (Figure 4.3(d)), through embedded chips (Figure 4.3 (a-c)).

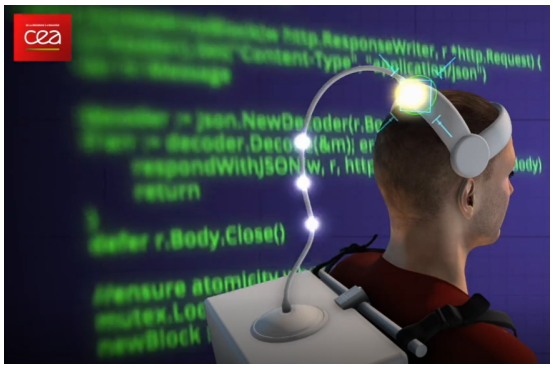
(a) *Hercules - RB3D*

(b) *Exoskeleton*

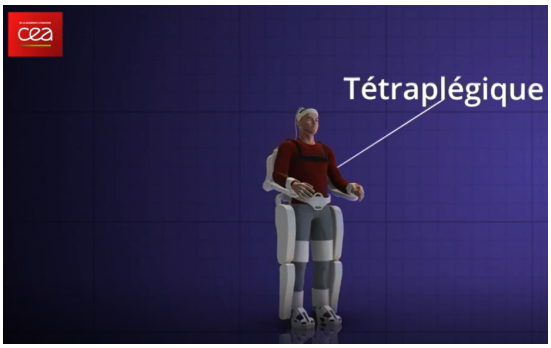


(c) *Brain's area dedicated to neuronal impulses for motor skills*

(d) *Embedded Chip able to process neural signals*



(e) *Embedded chip transforms neural signals into commands for the exoskeleton*



(f) *Exoskeleton activated by the command issued by the human brain*

Figure 4.3 – Figure (a) taken from [Her; RB3], Figures (b)-(f) figures excerpt from [CEAb].

The examples in Figure 4.3 well represent the current technological evolution, whose trend is expected to increase. More generally, we expect a growth of CPS civil applications, closer to the end-user, with a high autonomy and AI. This type of systems is referred to by the acronym ACPS, promoted by the European commission [San17], in which the properties of autonomy and adaptability are emphasized.

Autonomy in ACPS intrinsically involves automatic decision-making and, then, more extensively, embraces Artificial Intelligence, as clearly stated in [CyP14; Tör+16]. The introduction of AI in ACPS is revolutionizing safety techniques as traditionally used. Still a few years ago, some distinguished scientists wrote *Transcending Complacency on Superintelligent Machines*:

> "So, facing possible futures of incalculable AI benefits and AI risks, the experts are surely doing everything possible to ensure the best outcome, right?
>
> Wrong. [...] Some of us — not only scientists, industrialists and generals — should ask ourselves what can we do now to improve the chances of reaping the AI benefits and avoiding the risks." [Ste].

After that, in 2015, S. Russell *et al.* [RDT15], in a letter signed by many scientists, point out three challenges related to AI safety:

- *verification* (how to prove that a system satisfies safety-related properties);
- *validation* (how to ensure that a system meets its formal requirements and does not have unwanted behaviors); and
- *control* (how to enable meaningful human control over an AI system after it begins to operate).

In 2016, the problems raised by safety and AI have been further analyzed in [Amo+16], where the authors focus on accidents in the machine learning systems. At the same time, the Future of Life Institute is becoming a leading actor for AI Safety Research [Fut].

ACPS are expected to bring a technological revolution that will influence the lives of a large part of the world population and will have a deep impact on nearly all market sectors. Increasing levels of AI will allow ACPS to drive on our roads, fly over our heads, move alongside us in our daily lives and work in our factories, offices and shops, soon. In spite of this disruptive landscape, deployment and broader adoption of ACPS in safety-critical scenarios remains challenging.

This chapter discusses the scientific and technical results, investigating on critical ACPS applications. In those applications I chose to apply a contract-based approach to guarantee the system's coherence among different heterogeneous requirements and the different applied solutions. In doing so, my research activities contributed to a solid, rich and fruitful structural collaboration with KTH (Sweden), the University of Trento (Italy), the University of Berkeley (USA), and my workplace, CEA.

4.2 Contribution: Safety and Resilience for critical ACPS

Along my research trajectory so far, I firstly applied contract-based approaches [BCP07] to enforce and prove correctness by construction of an automatic code generation in high-integrity embedded systems (see Section 2.2 and [Can+10]). As briefly discussed in the introduction of Chapter 3, contract-based approaches are identified in [SDP12] as a prominent approach to seamlessly combine the various dimensions of the multi-layers CPS design.

However, it was still necessary to evaluate the expressiveness of contract-based approaches on concrete examples of CPS, from the textual specification down to the embedded code, and to

understand if the identified approach could also be extended to ACPS, where autonomy, AI and decision-making is even more relevant. Finally, it was still necessary to investigate solutions at tool level, both at the design and validation phases, to integrate contracts into the development of critical ACPS with new, highly interdisciplinary development methods.

The majority of my research work at CEA is situated within this space - thus contributing to the vision *"From Research To Industry"* brought by the *Technological Research Direction* (*Direction de la Recherche Technologique*) of CEA [CEAa] to which my department, (*Integrated Circuits and Digital System Division*), belongs. More precisely, in the last 10 years, I have studied the following research problems related to (mobile) critical ACPS:

- Are contract-based approaches able to provide a seamless guarantee of safety-related properties from ACPS design to execution platform?
- Can contract-based approaches be extended to guarantee seamlessly safety-related properties of a cross-layers architecture, from sensors to AI, and from AI to actuators, of an ACPS?
- Connected ACPS bring to the foreground the importance of ensuring safety and security properties together. In May 2015, for example, the Germanwings flight crash showed how a security property (the closed door of the cockpit cannot be opened by any staff outside the cockpit) affected a safety property (the pilot could have had a heart attack and no one could have intervened). Then, can safety and security properties effectively be specified and analyzed together at the design level via CBD approaches?
- Is a Two-Dimensional representation of the ACPS design enough representative of the environment in which an ACPS system operates? Do we need to evolve the design and validation tools?
- Mobile ACPS include an autonomous navigation system. Can AI improve such a system by guarantying at least the same level of safety as traditional autonomous navigation (without AI) and preserving performance of the mobile ACPS?
- Are AI-based methods capable of improving accuracy and uncertainty in critical CPS?

Previous research investigated the use of contract-based approaches in several civil applications of critical ACPS. They are the following:

- autonomous metro (2014-2016)
- a platoon of connected autonomous vehicles (2017-2019)
- autonomous drone (2017-2020)
- autonomous mobile robot (2016-2021)
- nuclear reactor (2017-2021)

Unlike the first two use cases, the last three include AI.

For each aforementioned use case, the next sections discuss the rationale, the specification of the use case, the scientific and technological problem studied, and how it was addressed. The jury can refer to publications for more details. Finally, my analysis and reflection considers ACPS applications as a whole (e.g. autonomous vehicles, drones, robots) and is grouped at the end of this chapter.

APPLICATION DOMAIN	SELECTED PUBLICATIONS and VIDEOS
Railway Application Domain	Publications [CSP14; CZP15]
Autonomous Mobile Robot	Publications [May+19], Video [CMS]
Automobile Application Domain	Publications [Pas+19]
Three-Dimensional animated scenarios	Publications [CSP14; CZP15; Pas+19; Laa+], Videos [ZC; LCa; LCb]
Autonomous Drones	Publications [Laa+]
Nuclear Reactor	Publications [CCC21]

Table 4.1 – List of the articles for the technical details

Railway Application Domain: Autonomous Metro As discussed in Section 3.3, the railway industry is increasing the automation of metros and trains. Only in Paris, for example, we have witnessed an automation of Metro 4 and RER B, as well as the extension of the automatic Metro 14. The automation of metros depends to a large extent on the CBTC system (Communication Based Train Control) and more precisely on a subset of the Automatic Train Control subsystem (ATC).

The associated **operational scenario** is the following. A train stops at a station that is equipped with a physical barrier and automatic doors, whose purpose is to protect passengers from the moving train (see Figure 4.4). In order to be able to operate train and platform doors, the doors of the train and the doors of the platform need to be aligned. At that point, both of them are automatically opened - thus allowing the passengers to get on and off the train. Finally, the train is authorized to move on if and only if both platform and train doors are closed. This operational scenario is often referred to by the technical term *passenger exchange*.

Figure 4.4 – On the left, automatic opened doors, on the right, the platform doors are automatically closing (images excerpt from YouTube)

The function *passenger exchange* is an important functionality of the CBTC, and this case study is obviously representative of ACPS. Indeed, it integrates not only computational and physical processes with feedback loops, but also the human factor. This function takes control of platform and train doors when the train is safely docked at a station; then it organizes the exchange of passengers (e.g. manage train and station doors opening/closing and doors blocking by passengers) while protecting them from any untimely train movement or non-aligned doors opening. It finally gives the authorization to depart when all safety conditions are met.

The operational phase linked to this case study is critical since doors are open and passengers can move freely between the train and the station.

I study the following research question:

> *Are contract-based approaches able to provide a seamless guarantee of safety-related properties from ACPS design to execution platform?*

Safety Regulation To demonstrate the safety-related *passenger exchange* properties, it is necessary first to identify hazards that cause accidents and/or near-miss accidents, then to establish contracts between the system components to define the necessary conditions that ensure safety, and finally to refine those contracts down to software.

Moreover, the approach I promote had to be compliant with CENELEC safety standards and integrated into the methodology and (semi-)formal languages adopted in the industry that gave the real use case and specified the safety requirements.

Autonomous Mobile Robot For some years now, we have seen an industrial trend to the development of mobile robots, with different levels of autonomy with/without AI, in different civil applications, such as household robots or toys. Compared to the past, where robots were mainly used in industry with little direct contact with the operator, the current type of applications target the masses and have a closer interaction to humans. This trend is expected to grow over the next decades.

To increase the presence of robots near humans, a few guarantees should be met. First, robots have to operate safely and not hurt users. Second, the user should not be required to program, reset or maintain the robot. Third, unlike a factory, the environments in which the robot operates are not controlled in any way by the constructor or the developer. Thus, the robot must adapt itself to its environment and manage its own resources such as available memory, computing power and battery.

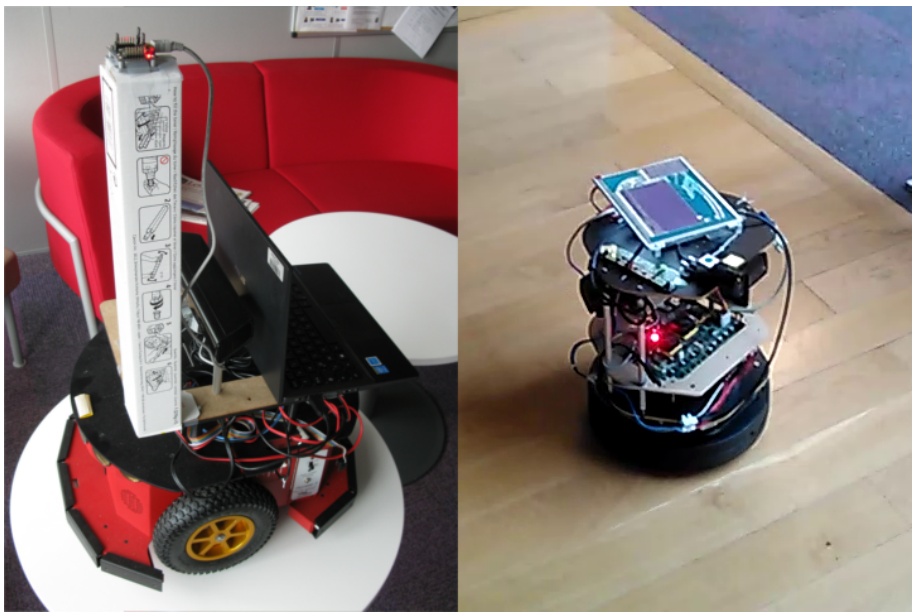


Figure 4.5 – The autonomous mobile robot having AI

Figure 4.5 shows an autonomous mobile robot, having AI, owned by the CEA department I work for. The LEN (Lifelong Exploratory Navigation) architecture has been deployed in the robot. LEN has been firstly developed by my colleagues, Laurent Soulier and Fabrice Mayran de Chamisso [CLM15; MSA16; May16]. LEN is based on a cross-layer architecture from the robot's sensors to the occupation grid map,

to the generalized Voronoi graph (GVG) until to the navigation and exploration algorithms [CLM15; MSA16] and rising up to the high-level AI, and from here down across the architecture until the actuators [May+19]. LEN allows a robot navigation and space exploration in dynamic and unknown environments (i.e. without the need to create a static digital cartography of the environment, which is very expensive in terms of cost and time).

One of the difficulties related to ensure safety properties for these applications is given by the level and type of acceptable uncertainty. These are the cases, for example, in detecting, and avoiding, obstacles, or recognizing a place already visited, by making safe decisions accordingly (e.g. brake, turn right).

I study safety-related properties of LEN and I investigate the following research question:

> *Can contract-based approaches be extended to guarantee seamlessly safety-related properties of a cross-layers architecture, from sensors to AI, and from AI to actuators, of an ACPS?*

Automobile Application Domain: A Platoon of Connected Autonomous Vehicles The trend towards the adoption of autonomous vehicles has been tremendously increasing in the last few years, with the expectation of a significant reduction of road accidents, increased fuel economy and an overall higher traffic throughput [TM16; MW19; LK19]. Today, road tests of autonomous vehicles are in place in several countries, such as those from Uber in Pittsburgh [BN16] and from Valeo in France [Bou19]. For example, in the Valeo test, an autonomous car drives on the Paris beltway, automatically adapting its speed to traffic conditions. During the road tests, the French authorities demand a human driver supervisor (in the car), whose duty is to act in emergency situations.

A careful examination of road tests reveals several open problems both at the scientific and technical level, as well as the economic and social level. Clearly, traffic control is best tackled through a cooperative approach, where vehicles exchange information to jointly build a detailed picture of the current situation [Nie+16; ZLL18]. This, in turn, requires an efficient and *secure* vehicle-to-vehicle (V2V) communication mechanism to exchange information (e.g., an obstacle on the road) [BTD06]. At the same time, the heterogeneous nature of the problem, and its fast evolution, demand a standardized infrastructure and a design methodology by which designers can unambiguously formulate the requirements and the properties of the system, to deliver functional as well as non-functional *safety* guarantees [Can+10]. This is essential to drive the adoption of the technology, and lower the risk perceived by the user. For this reason, there appears to be an increasing demand of attention to both safety and security properties together. Yet, this approach is impaired by the fact that those properties are specified, analyzed and developed by different teams having different cultural background and tools.

The use case is based on [SAR; SD90] and depicted in Figure 4.6. It is composed of a platoon of three connected autonomous vehicles. The first vehicle is the leader of the platoon and the other vehicles follow.

The addressed **operational scenarios** are the following:

- if a vehicle brakes, e.g. because it has detected an obstacle, it sends the *brake command* signal to the following vehicle, which must then brake accordingly, and so on until all the vehicles in the platoon received the signal.
- If the minimal safety distance between two vehicles is not satisfied anymore, e.g. during a descent, the following vehicle must slow down until it meets again the safe distance between the front vehicle and itself.

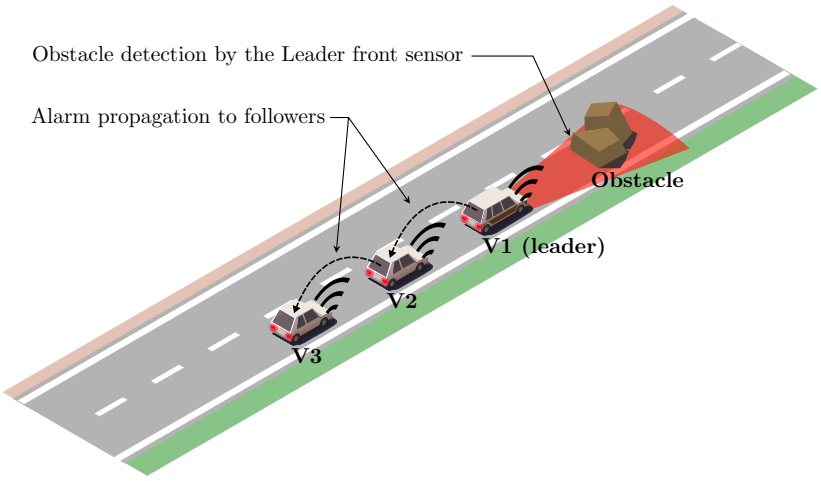


Figure 4.6 – A Platoon of Connected Autonomous Vehicles (Figure by Sebti Mouelhi and extracted from [Pas+19]).

- If the maximal distance between two vehicles is not satisfied anymore, e.g. because an obstacle between the first and second vehicle is suddenly occurred, then an *alert command* is sent to the front vehicle which must slow down (or brake depending on the messages of the communication). (if it is an obstacle, then the vehicle first brakes and sends the brake command to the following vehicles, then alerts the front vehicle).

In that application, I address the following research question:

Can safety and security properties effectively be specified and analyzed together at the design level via CBD approaches?

More specifically, the following two properties have been addressed.

- First, the V2V communication ought to be safe: if vehicle V1 detects an obstacle on its route, then V1 alerts V2 through a brake signal. This V2V communication shall be guaranteed within a given elapsed time.
- Second, the V2V communication ought to be secure. This involves that the authorization and authentication phases should be guaranteed.

The use case has been developed from the highest level requirements, provided in textual form, to the specification of the system in semi-formal languages, down to the software components and the Ada code, up to a toy prototype.

The development of the use case intentionally required the active collaboration of several research teams in Europe. E.g. TTTech (Austria) and VTT (Finland) have provided the requirements; the CEA team (France) and the university of Trento (Italy) specified accordingly the system in a semi-formal language, proved safety properties, in compliance to the automotive safety norm [ISO18], and coordinated the teams; the Budapest University (Hungary), the AITA International Inc. (a small industry in Hungary) and the Aalborg University (Denmark) analyzed the security properties, thanks to the Arrowhead Tools framework [Del+17]; the ECE team (France) developed the code and the prototype. This fruitful collaboration made it possible to simulate, with all due cautions, a rich and multidisciplinary work environment.

The international team promoted a contract-based methodology that, starting from natural language requirements, reaches the prototyping stage of a platooning autonomous vehicle system, with an additional focus on safety and security requirements.

Safety Regulation Moreover, the proposed methodology is compliant with the ISO26262 [ISO18] safety norm. A newer norm called ISO/PAS 21448:2019 Road vehicles - Safety Of The Intended Functionality (SOTIF), addressing the different levels of autonomy especially in emergency intervention systems, was released during the publication. SOTIF is not discussed but, being an extension of ISO26262, we are confident that the methodology is compliant with it (but it needs to be proven).

The Importance of Three-Dimensional animated scenarios As discussed in Chapter 3, CPS constitute a disruptive new discipline of systems engineering, whose development would require new, highly interdisciplinary development methods. This requirement also effects the tool level, i.e. whether the tools designed for the High-Integrity Embedded Systems can also be suitable for ACPS, and in particular for critical ACPS that require additional analysis and verification.

The research question I investigated is then:

> *Is a two-dimentional representation of the ACPS design representative enough of the environment in which an ACPS system operates? Do we need to evolve the design and validation tools?*

In the early years of the second decade of the 21th century, most of the design and analysis tools are essentially based on a static, two-dimensional representation of the system (with a few exceptions, such as the animated state machine diagram), as is the case for PlotemyII [Pto14], Simulink [Matb], MATLAB [Mata], UPPAAL [Upp], and the eclipse-based tools (such as Obeo, Papyrus) [The]. These type of design and analysis tools provide countless advantages as shown by the breadth of related literature.

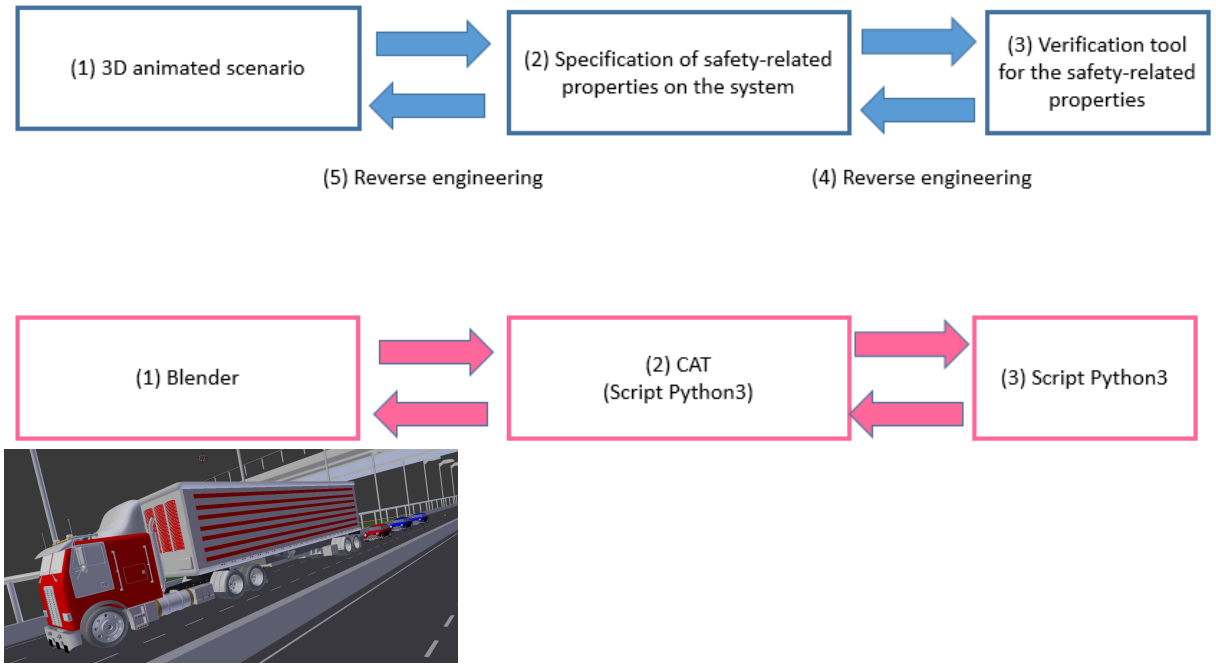


Figure 4.7 – CAT. In Blue color, the design of CAT. In Pink color, its implementation. The Image is taken from the 3D animated scenario for the Platoon of Connected Autonomous Vehicles Use Case and represents Figure 4.6

I wondered whether relying exclusively on a static and two-dimensional (2D) specification was sufficient to represent the environment and the operational scenarios where ACPS work. More specifically, it was my impression that a prospective change was necessary when we study ACPS with respect to how we were used to reason about High-Integrity Embedded Systems and to represent them in such reasoning. Certain concepts and ideas cannot easily be represented in illustrations or 2D representations, and if they were, time and effort are needed to understand the complex representation.

From the quest for understanding which tooling specification was best for reasoning on ACPS, I have the idea of creating an animated three-dimensional (3D) scenarios (like a film or video game for instance) that could have represented the operational scenarios and the environment in which the ACPS operate.

To the best of my knowledge, 3D tools that existed at the time did not allow the specification of complex operational scenarios in which the ACPS under study was immersed.

In this effort to investigate on which tools could provide a representation of animated 3D scenarios, suitable for ACPS, I discovered that the United States are also going in that direction, e.g. with the excellent works of Taha [Tah+13].

In this context, I proposed CAT (Contract Analysis Tools). CAT is intentionally based on the work of [Ham]. CAT has been successfully tested on several use cases [Pas+19; CZP15; ZC; LCa; LCb].

Autonomous Drones Figure 4.8 represents two real examples of the use of drones in civilian emergency situations. In 2018, persistent heavy rains in France resulted in the Seine river breaking its banks. For a couple of weeks, the situation caused an imminent alert in Paris. The Prefecture of Paris used drones to control the embankments of the river [Pre]. In 2019, during the Notre-Dame fire, drones were used to have a "real-time" map of the emergency's situation and to quickly understand how to best manage the situation (to put out the fire and save the artistic patrimony of the cathedral).

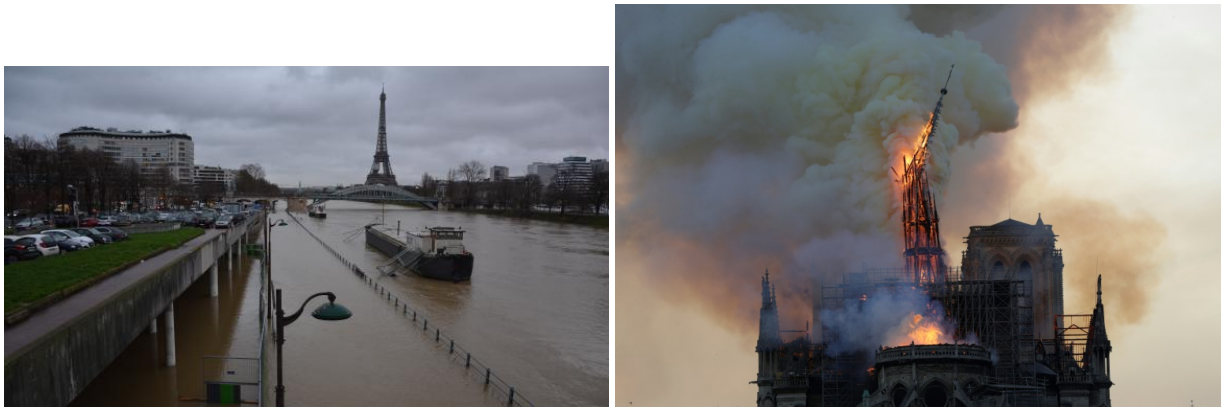


Figure 4.8 – Flooding of the Seine river in 2018 and Notre Dame on fire in 2019

In the above real examples, drones are directly piloted by police squads. I think that reactivity in the mitigation of the accident could further improve, if drones were capable to reach the place of the accident by themselves, in autonomous way.

I then studied the following **operational scenario** (see Figure 4.9): In the case of an accident, a fully-autonomous drone autonomously navigates until the accident site. Then the drone sends images and videos to the ground control station. The information is immediately treated by first aid personnel who can decide promptly and with full knowledge of the causes to send the right aid for the situation (e.g. 3 ambulances, 2 fire brigade units and 3 police units and staff to handle the circulation). The drone's navigation map is autonomously calculated by the ground control station

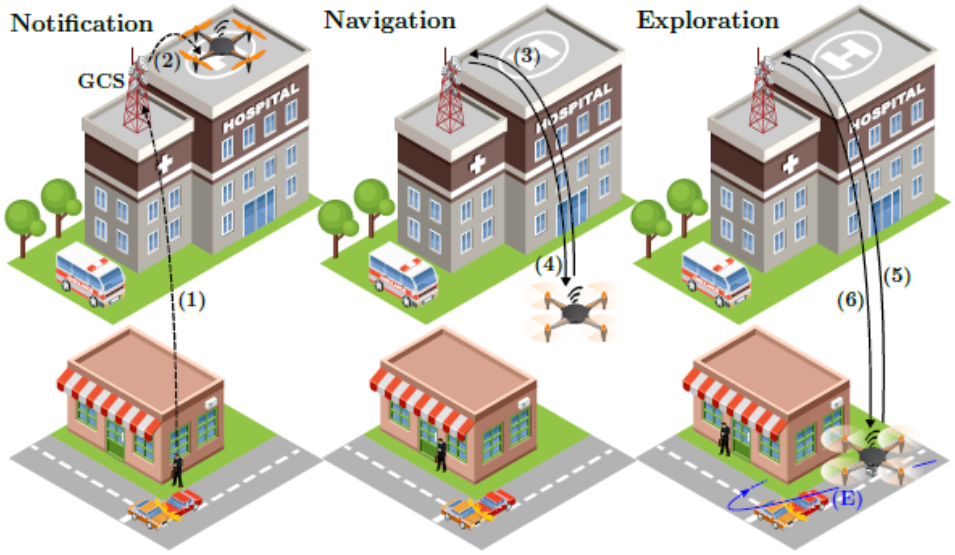


Figure 4.9 – (1) Notification phase: a person on-site sends an alert to the GCS by smartphone; (2) GCS activates the drone; (3) Navigation phase: the drone continuously sends navigation data to GCS; (4) GCS controls the drone when necessary; (E) Exploration phase: while hovering, (5) the drone broadcasts the accident scene to GCS, which in turn (6) may order it to provide medical supply for victims if needed. Steps (3)/(4) and (5)/(6) are repeated with different frequencies (depicted by $\longrightarrow$) [Laa20].

and then sent to the drone. This solution allows preserving the drone battery, which is necessary to get to the destination point and to be able to send the video. Finally, for the sake of privacy issues, the drone activates the camera only when it has arrived to the accident site.

Traditional mobile navigation systems for ACPS exploit algorithms based on A*. The latter has the advantage of finding a path in a deterministic way, which is a salient property for safety viewpoint.

The research question I addressed was:

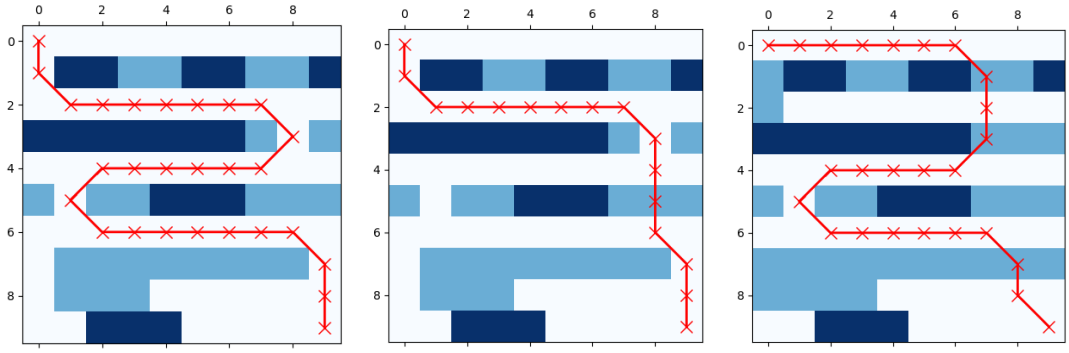
> *Can AI improve mobile ACPS by guarantying at least the same level of safety as traditional autonomous navigation (without AI) and preserving performance?*

Safety Regulation The European regulation of unmanned aircraft operations established by the European Aviation Safety Agency (EASA) [Eur15b] divides drone operations into three categories as follows: open (low risk), specific (medium risk) and certified (high risk). For the open category, the aircraft is not allowed to operate at a height exceeding 150 meters above ground [Eur15b]. Hence, the drone must respect a minimal distance with regard to physical obstacles such as building roofs. We arbitrarily define this distance to be equal to 15 meters. Consequently, each cell in the digital environment representation with a terrain elevation higher than 135 meters is considered as a physical obstacle that the drone must avoid in its journey (trajectory).

In addition to the physical obstacles, the European regulation of unmanned aircraft consider regulatory-related obstacles, which include airports, stadiums, embassies, factories, etc. The drone is not allowed to fly over such areas unless the regulation authority allows it to.

During the path planning process, both physical and regulatory-related obstacles must be taken into account.

Three operational scenarios were studied (see Figure 4.10):



(a) (b) (c)

Figure 4.10 – The dark blue color represents physical obstacles. The light blue color represents the regulatory obstacles. The red line is the path found by the hybrid genetic algorithms-based (GA) algorithm. In case (a), both algorithms (A* and GA) find a solution. In case (b), the path found by A * is not walkable due to the battery level of the drone. On the contrary, the proposed hybrid GA algorithm finds a path that fly over regulatory obstacles, because it takes into consideration the battery parameter. In case (c), A * finds no path. On the contrary, the proposed algorithm finds a path that can be flown over.

- Base case. There is an accessible path without any obstacles between the location of the drone and the accident site and the drone battery is sufficient to reach the target.
- Battery insufficient case. The level of the battery does not allow the drone to navigate the accessible path and reach the target.
- Case regulatory obstacles. There is no accessible path that does not overfly regulatory obstacles.

To compare the results quantitatively, for each operating scenario we introduced above, the navigation system was calculated with A*, which provides us the (safety) baseline, and with the promoted algorithm, which is based on hybrid genetics algorithms. Tests were carried out on i7 Intel Core processor with a 2.8 GHz frequency. The map of the environment is 10*10 because the goal was to compare the two solutions. However, when we repeated the test on a map with a bigger size (100*100 and 500*500), the computation time is not impacted for both algorithms (i.e. A* and our hybrid genetic algorithm). The computation time remained inferior to 1 second.

The result of the tests clearly shows that embedded AI into the autonomous navigation system not only achieves the same safety level as the baseline, but improves it, if it is necessary to find an alternative air travel that overflies regulatory-related obstacles (see Figure 4.10).

Nuclear Reactor In the second half of the last century, we have witnessed considerable deployment of several, different nuclear power plants in the world. They differ with respect to power, design, adopted technologies and digital systems (list not exhaustive). For example, in the seventies, 900 MegaWatt nuclear power plants had a hardwired *Instrumentation and Control* (I&C), while, in 1985, the French 1300 MegaWatt nuclear power plants deployed the first digital I&C system (ref. to DIPS - the Digital Integrated Protection System).

A nuclear power plant is an example of Cyber-Physical Systems (CPS) [Fra16], i.e. physical systems monitored and controlled by electronic systems, otherwise termed I&C systems. In a nuclear power plant, a nuclear reactor is, by namesake, the core of a power plant where nuclear fission happen (see Figure 4.11). A nuclear reactor is controlled by numerical I&C commands, sent either automatically by the software or directly by the operators in the main control room.

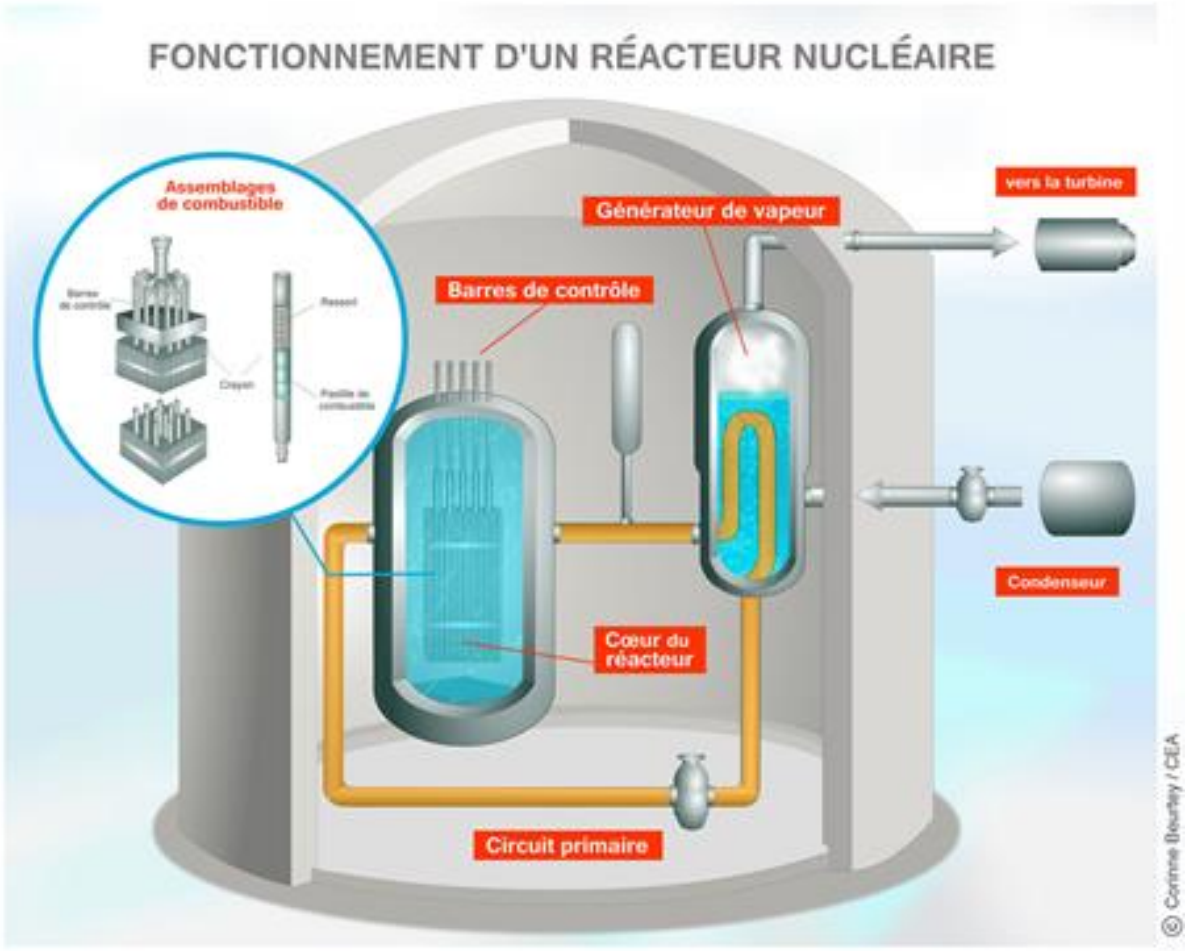


Figure 4.11 – Nuclear Reactor and Nuclear Power Plant. Image CEA [CEAa]

In the I&C systems, the protection system is one of the most critical systems. Its main mission is to control the nuclear reactor in the case of nominal as well as degraded behavior. Its functionality plays a leading role in setting "safety measures to prevent incidents and to mitigate their consequences if they were to occur" [Int18] - where "incidents include initiating events, accident precursors, near misses, accidents and unauthorized acts (including malicious acts and non-malicious acts)" [Int18].

To prevent incidents from happening, the nuclear safety authorities demand to assess a safe threshold (or safe margin) and, consequently, a second and depending threshold devoted to the margin of errors (list non-exhaustive). The latter ensures an additional protection measure, which ought to consider acceptable uncertainty. In both cases, the threshold depends on several parameters that involve different countermeasures, i.e. different I&C commands.

Nuclear safety standards IEC 60880 [IEC06] and IEC 61513 [IEC11] define the safety objectives for a safe system and critical software. They demand (A) measures for functional and performance requirements in response to specific signals or conditions. Among the parameters under examination, response time and accuracy have to be taken into account [IEC11]. Moreover, (B), all measures must be validated and proved (list non-exhaustive).

To achieve a greater level of accuracy, we need to control the acceptable uncertainty degree of each component and system integration involved in a nuclear reactor design. All these measures and controls are resource intensive. This constitutes a challenge, since they must be proved and validated in a reliable execution time (list non-exhaustive). However, a closer examination reveals that an

accurate computational model of a nuclear reactor behavior is too complex to be controlled and simulated in a predefined computing execution time. The heterogeneity of its data model is among the main factors generating such a complexity, which reflects different disciplines (e.g. neutronics, thermal-hydraulics, computer science). To complicate the scenario further, data influence each other. In the meantime, the increased hardware performance capability and advanced numerical methods are allowing innovation in multi-physics models to calculate and prove more accurate measures to assess safe margins.

In this context, I have studied the following problem:

Are AI-based methods capable of improving accuracy and uncertainty in critical CPS?

The underlying idea is that Machine Learning based algorithms have the competitive advantage of being able to analyze big volumes of information, composed by neutronics (neutron flux) and the thermal-hydraulics parameters (temperature, density, velocity fields), in high fidelity multi-physics modelling. The expectation is then to achieve a better model accuracy and finer control of the acceptable uncertainty degree, required by the nuclear safety regulations - thus avoiding or, *à défaut* reducing, human errors in the decision-making.

Safety Regulation I would highlight here that nuclear safety regulations are among the more conservative safety regulations and, hence, at the best of my knowledge, they do not allow AI today. One of the main blocking points is the (formal) proof that all provided measures, included accuracy, safe margins, uncertainty, response time and execution time (list non-exhaustive) must be validated and proved [IEC11]. However, advanced study in computer science (including results on mixed-criticality on multi- and many-core - of course with additional proofs and all possible cautions) could be better investigated for nuclear reactor analysis in the design phases and for the study of the parameters based on the fuel consumption.

4.3 Reflections on the Lessons Learned

If we consider ACPS applications as a whole (e.g. drones, robots, autonomous vehicles), we can make some observations that I want to share in this section.

Contract-Based Approaches As I have had the opportunity to test contract-based approaches on different application domains, I can affirm that contract-based approaches offer a seamless approach to demonstrate safety-related properties from the highest level requirements, often specified in textual form, up to the code. Moreover, contract-based approaches are easily capable of being integrated with the methodology used in the industry. This feature plays an important role, whenever a researcher aims to transfer the achieved scientific and technical results to industry (thus satisfying one of the CEA missions). One of the difficulties to overcome when industrial transfer is contemplated is that an industry, with the exception perhaps of start-ups or small industries, already has its own methodologies, its tools and has already spent money and time in training its workforce. Then, to avoid the industrial comment "very interesting, but it demands us too much effort in time and money", it is better to develop a methodology that is easy to (1) understand, (2) use and (3) which does not require an overturning of the industrial methodologies in place.

In my experience, contract-based approaches have been a means to facilitate the demonstration that the promoted methodology was compliant with the safety norms. More generally, as I have shown on the use cases analyzed in the previous section, contract-based approaches are a rigorous method of developing and measuring the quality of the development of a product and then of safety (Figure 4.12).

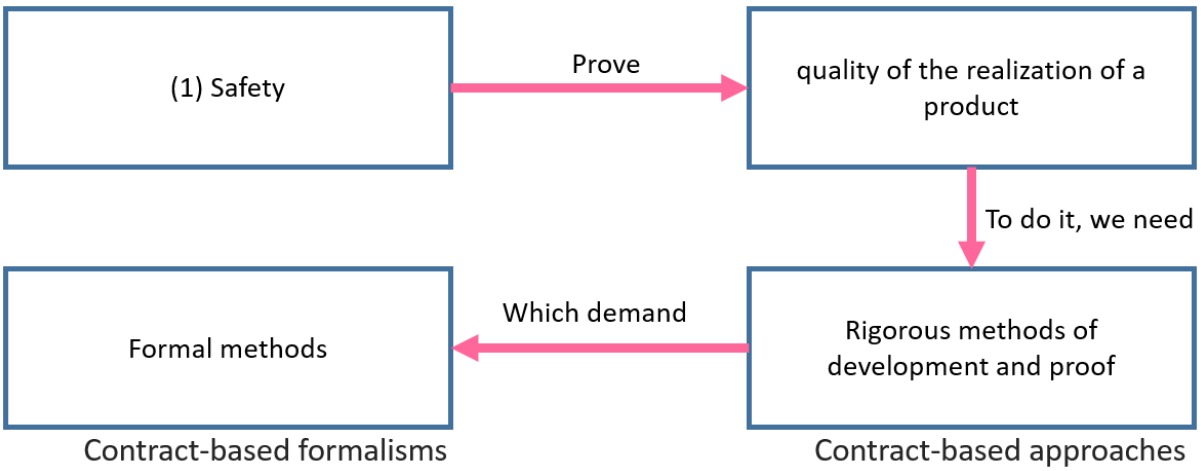


Figure 4.12 – Relationship between contract-based approaches and Safety

Finally, contract-based approaches are used as a rigorous method to ensuring both safety and security properties together at design phase. They have also improved communication between the different involved teams. The use of contracts showed that secure communication protocols between autonomous vehicles can raise conflicts with safety requirements and then decrease the overall safety level of the ACPS. To overcome this conflict, we have to find alternative solutions at design level in order to satisfy both requirements (i.e. by adopting another security protocol and including an embedded architecture capable of sending the *brake command* regardless of the received signal).

A Tool Perspective I tested CAT and the 3-dimensional animated scenarios on different use cases [ZC; LCa; LCb]. A considerable difficulty in terms of time-consuming effort is not based on the API script between the tool for the 3D animated scenarios and the selected analysis tool (e.g. CAT), but rises from specifying the environment where the ACPS work (e.g. the buildings of a city), a difficulty that increases with respect to the expected realistic, or even exact, reproduction of the environment. However, I strongly believe that in 2021 and with the new generation of engineers, a 3D animated scenarios should be integrated with traditional approaches. Based on my experiences, a 3D animated scenario representation of the environment, where the ACPS operate, can help to reason about the expected ACPS behaviour. To the best of my knowledge, in the last few years, there have been interesting studies and investigations (e.g. the CPSwarms project [CPSb]). Finally, if I were to redo the work today, I would look very carefully at the ROS2 tool suite [ROS] because it provides 3D animated scenarios by integrating real-time constraints, and the community behind ROS2 is closer to computer science. During the LEN works [May16; CMS], ROS has been used to simulate the environment in laboratory tests, before doing the tests in the real environment. The next development of LEN entails the adoption of ROS2.

From Safety to Resilience Along my research trajectory so far, I pursue the quest whether AI-based techniques could improve the safety level of a system. The way forward has been the researches on the safety navigation for autonomous drones, and on accuracy degree for the simulation model of a nuclear reactor.

Note that from these results and expectations it is not inferred that AI always improves the safety level, because they must be generalized. My studies give a first positive and encouraging feedback to investigate in, or at least not exclude, this direction.

However, one of the major challenges is proving the compliance of the proposed AI-driven methodology to the related safety regulations. As a matter of fact, if we consider ACPS applications

as a whole (e.g. drones, robots, autonomous vehicles), we discover that safety-related analysis, as applied in the traditional application domains such as railway or nuclear energy, suffers from given limits when we focus on ACPS.

Some of the most important assumptions in traditional safety engineering lie in models of how the world works. In this regard, a capital example is provided by accident causality models, currently bounded in scope, not extending to advanced levels of autonomy where system behavior depends on its memory and appraisal of received stimuli. Current approaches include probabilistic risk assessment such as Fault Tree Analysis (FTA) and Failure Modes and Effects Analysis (FMEA), working on the assumption that, once the system is deployed, it does not learn and evolve anymore. Another strong hypothesis of the traditional safety approach is based on the user-responsibility. In traditional domains, specialized teams use the final product and are responsible of their maintenance. In the railway, for example, we have specialized train-drivers, teams specialized in the train's maintenance, teams specialized in the physical infrastructure, team specialized in the electrical infrastructures, etc. The responsibility of an accident is totally supported by the (train) constructors and/or the (railway) company. Rarely, a passenger is fully responsible to mitigate a possible accident. This situation does not happen in some applications of ACPS. Compare the railway organization with the case of a drone. The drone driver is not a safety engineer and/or a specialized drone driver. Often, they have another work and use the product, for example, to control agriculture or to film video. In the case of an accident, the user's responsibility is primary engaged and analyzed. The role of a user of a (semi-autonomous) drone could be similar to the driver of a vehicle. Among the main differences, however, is the use of redundancy as a means to reduce the risk of an accident. The limitation of the physical space and cost constitute stronger constraints for ACPS applications than traditional ones. In many cases, heterogeneous architectures cannot be applied. Preventing accidents in ACPS requires using models that include the entire socio-technical aspects and treat safety as a dynamic control problem. Future intelligent autonomous systems need to be able to appraise safety issues in their environment, self-learn from experience and interactions with humans, and adapt and regulate their behavior appropriately.

Many safety communities, including ECS-SRIA and expert technical groups in the regulation, are studying how to change the safety standards to take AI into account. Of the extensive topic, I only highlight the need for paradigm changes, and promote a change from safety, as traditionally studied in high-integrity embedded systems, to resilience. The latter could allow us to release some hard constraints related to safety (e.g. redundancy, sensor quality) and provide a given level of no longer safety but system resilience.

In this regard, we should distinguish between Exogenous and Endogenous resilience.

Endogenous Resilience is the ability of the system to detect and manage internal faults and malicious attacks. For example, a drone should be resilient to internal software errors, manage battery properly, execute a command, detect physical damage of sensors and control it, eventually, with suitable safety measures.

Exogenous Resilience deals with the system's external environment in which it is operated. Avoiding an obstacle constitutes the leading example of Exogenous Resilience.

Exogenous and Endogenous Resilience are pretty new in ACPS and, in at the best of my knowledge, established methodologies and tools for ACPS, including AI, do not exist yet.

Certification Similar to what is happening for safety, the certification of ACPS shows its limits (Example 4 provides a short and simplistic description of the certification process).

Although the benefits of certified systems are marked and outstanding, the cost of the certification remains expensive and, in many case, prohibitive. In avionics, the cost to ensure the most critical

level (level A) is estimated to be more than 55% with respect to the minimum level (level E) [HB07]. In the drone application domain, only one category requires certification ('Certified' Category for operations with an associated higher risk) [Eur15a; Eur16].

As in the case of safety, certification is based on the hypothesis that the system to be certified will have the same behavior (it always implements only the specified functionality). In other words, no learning phase is allowed. For example, in the avionic application domain, a program does not change its behavior with the increasing number of performed flights. Similarly, in the railway application domain, the control command to automatically open/close the doors of an automatic metro will have the same behavior forever. However, the AI introduction overstretches the fundamental hypothesis on which the traditional certification process is based, by allowing learning phases and changes in the system's behaviors. In addition, if AI is coupled with a fully or a high level of system's autonomy (i.e. human has not the control of the system), then they unwind civilian responsibility: who is responsible of an accident of ACPS having AI? How can we protect the ACPS and humans? Is it possible to certify an ACPS having AI? And how?

The **certification** of a system happens after the design and development and the integration of the subsystems. A certified system can include subsystems having different level of criticality, i.e., different levels of certification [HB07]. Broadly speaking, once the system is integrated and analyzed, safety teams provide the arguments for the certification. A third certification entity firstly analyzes them and, then, discusses them via audit. We highlight that a system is always certified with respect to a particular use and related norms. For example, a system having a Technical Standard Order (TSO) authorization cannot be installed and used in an aircraft without passing the avionics certification [FAA].

Example 4 – Simplistic Overview of the Certification Process

Social Trust, Privacy and Sustenability The investigation on safety for critical ACPS involves potentially new solutions, which should meet constraints of three disciplines: privacy, social trust and sustainability, which have nothing to do with safety or computer science in *scricto sensu*.

Privacy for drone applications is an extremely sensitive issue, as analyzed by the works of Silvana Pedrozo and Francisco Klauser (sociologists). The authors are (also) known for a poll they made on the Swiss population on the acceptability or not to drones in 2015. The results have been discussed via several means (radio, scientific journals, etc.) [KP17; Sil18]. In [KP17], the authors state that: "whilst the majority of respondents are supportive of the use of unarmed military and police drones (65 and 72% respectively), relative numbers of approval decrease to 23 and 32% when it comes to commercial and hobby drones". The underlying reason of such acceptability is not based on the guarantee of safety level of that systems. For example, in 2015, a drone injured a woman at the Pride Parade in Seattle [Mil17] and the news had a very broad resonance [BCC15]. Instead, it is based on privacy concern and individual freedom [Sil18].

The recent accidents involving autonomous systems (e.g., Tesla fatal car accidents), show that sole engineering progress in the technology is not enough to guarantee a safe and productive partnership between a human and an autonomous system. Then, parameters based on (given levels of) social trustworthiness, should be taken into account. As a result, some of the open problems to be overcome are the following. How to verify (reason about) such specifications in the context of a given human-machine collaborative context? How to synthesize (design) an autonomous system such that, in a collaborative context with a human, these specifications are guaranteed?

Finally, the attention to the environment has increased significantly in the population. This implies that any safety solution is not enough, and we have the moral obligation to conceive and develop solutions that can also meet eco-sustainability constraints.

Chapter 5

The CPS Education Challenge

Contents

Over the last 8 years, I carried out a number of actions and provided consistent contributions around the challenge of designing a new education path for CPS engineering. I collaborated with the International Workshop on Embedded and Cyber-Physical Systems Education (WESE), part of the EsWEEK conference, as program committee. In WESE, I have had the opportunity to interact with the international community on education methods for CPS and directly contributed with two papers. I was the Principal Investigator and Organizer for the EIT Digital CPS Summer School in 2016 and delivered a lecture at PhD Doctoral School in 2015. Finally, in 2018, I was the Principal Investigator of the Arrowhead Tools project on behalf of CEA, where I set the main contribution of the French consortium towards the design and the development of a tool devoted to the education on CPS and AI-driven systems. The tool mainly targets the CPS education of high school students, putting the school in disadvantaged neighborhoods as a priority.

This chapter reflects these experiences. It should be noted that this activity is not my main priority activity at CEA. Therefore, the results and the analysis in this chapter are not "original" in the sense of the those discussed in the other chapters of this document. Most of them are firstly introduced by my colleagues in Europe at KTH, TU Dortmund, Université de Toulouse - ISAE, and in USA at Berkeley, Halmstad University, Rice University. However, as I am part of the WESE community, I actively contributed to the WESE intent by applying these methods in my working space. The discussion I present in this chapter is the reward of several years of my reflection on this topic and actions that I am performing on it.

5.1 Context

As discussed in Chapter 3, in the second decade of the 21^{th} century, we have witnessed an emerging discipline, widely known under the term Cyber-Physical Systems, i.e. systems which couple cyber (computing and networking) with a physical side (mechanical, electrical, and chemical processes) [SDP12]. Since the first appearance of CPS, the scientific community had it very clear that the CPS actually

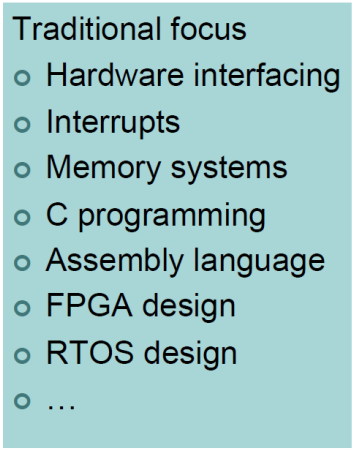


Figure 5.1 – Embedded Systems Traditional topics. Figure drawn from [Edw09]

constituted a disruptively new discipline of system engineering, whose development would require new, highly interdisciplinary, methods, design and tools.

Figure 5.1 is drawn from [Edw09]. Although we should always be wary of generalisation, we can state that the figure shows the main topics of a traditional university course on Embedded Systems, independently of the location and the year. The emphasis is given on programming languages, synchronization mechanisms, coordinating access to shared resources, be they software (such as data) or hardware (e.g. memory access). Particular attention is paid to Operating Systems, Kernel and FPGA control. An interesting and comprehensive presentation of Embedded Systems is given in the book [Mar03]. Its first edition was published in 2003.

The scientific community has been wondering whether a new education tailored to CPS is necessary or if the pre-existing one on embedded systems is sufficient. In this regards, the term *education* is not only restricted to the content of a CPS course, it embraces also the ways of teaching the new discipline. Then, the question *"is the pedagogy in use sufficient to address the interdisciplinary nature of and the challenges brought forward by CPS?"* is added to the traditional questions on the content and the relationship between the two disciplines, for example *"What is covered by one topic and what is not"? "Do we need to change the contents of an ES course to take into account CPS topics?" "Do we need to add a CPS course to a preexisting embedded systems course or should we substitute a CPS course to a preexisting embedded systems course?"*

Since 2005, researchers, educators, and industrial seniors gather at the Workshop on Embedded and Cyber-Physical Systems Education (WESE) joint with the EsWEEK conference. "The WESE community aims to assess needs and share design, research, and experiences in embedded and cyber-physical systems education. Demand for embedded and cyber-physical system engineers has motivated a growing interest in the question of educating specialists in this domain. For this reason, WESE community addresses questions such as *What skills and capabilities are required by the engineers of tomorrow*, and *How should the corresponding educational programs be formed, in order to provide experts ready to engineer the Cyber-Physical Systems that will greatly impact our future society*"? [WES]

The quest for contents and pedagogy fit for CPS education should reflect the drive of the European Commission towards an approach to education (and research) closer to industry, innovation and entrepreneurship. J. M. Barroso, president of the European commission between 2004 and 2014, promotes a development of an ecosystem based on three pillars: education, science and industry (See Figure 5.2). He states that:

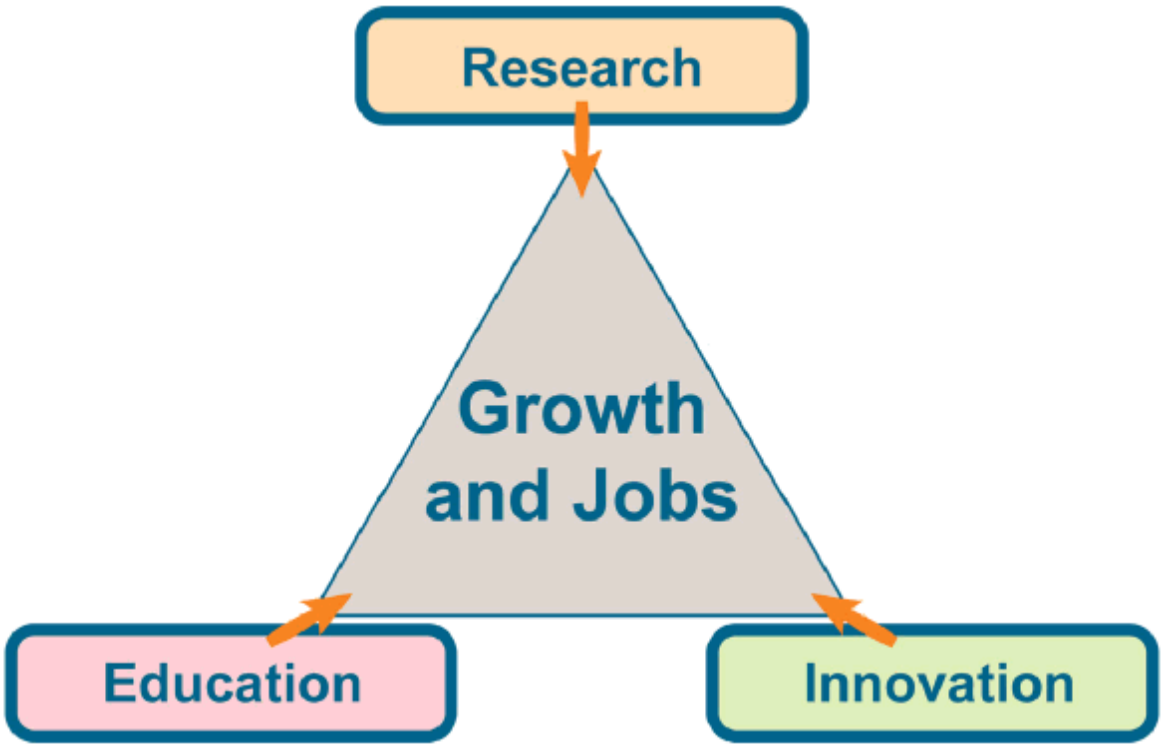


Figure 5.2 – The Three Pillars. Figure drawn from [LM13]

> "History shows that there is no sustainable path to growth and prosperity outside the research-innovation-education triangle" [Bar12]

This way forward is reinforced by the former European Commissioner for Research, Innovation and Science, Máire Geoghegan-Quinn,

> "Our future - she says in [Eur13] - beyond the crisis depends on having the capacity to transform the structure of the economy towards more knowledge-intensive and innovative industries and services"

European universities are then confronted with the EU political drive to carry education through a more societal and economical support than they have done in the past, as the capacity for innovation by new generations of CPS engineers is expected to play a predominant and leading role. Much has been written about the importance of innovation. Conversely, much less has been said about the evidence that innovation is necessary but not sufficient to have an impact on society and economy. In this regard, if several economists consider innovation as a major force in economic growth [Ros04], economical studies confirm that technological innovation could be unprofitable, when time is not considered in the economy performance equation [Ros04; Nor96]. This last concept is also highlighted by the European Artemis (Advanced Research & technology for Embedded Intelligence and Systems) initiative which points out the greater impact and time-to-market results as the target priorities for the next coming period (2017-2025) [ART]. As a result, the ability to keep innovation together with time-to-market is necessary to any achievement in this domain.

This requires several actions on several levels.

Edition 2003	**Editions 2011, 2018 and 2021**
Introduction	Introduction
Specifications	Specifications and Modeling
Embedded System Hardware	Embedded System Hardware
Standard Software: Embedded Operating Systems, Middleware, and Scheduling	System Software
Implementing Embedded Systems: Hardware/Software Codesign	–
Validation	Evaluation and Validation
–	Application mapping
–	Optimization
–	Test

Table 5.1 – Table of Contents of the four editions of "Embedded System Design". The second edition introduces the subtitle "Embedded Systems Foundations of Cyber-Physical Systems" [Mar11]. The third and fourth editions have the subtitle: "Embedded Systems Foundations of Cyber-Physical Systems, and the Internet of Things" [Mar18; Mar21]

5.2 What Content for an Education on CPS?

Sixteen years have now passed since 2005 and the WESE community has identified the main body of a CPS course, which becomes more open to modeling, design, evaluation and validation topics (see e.g. [LS10; ME11]). We have already mentioned the book *Embedded System Design* [Mar03]. It is interesting to study the evolution of the editions, because they show the educational shift between embedded systems and CPS and how the latter also is accompanied by the technological evolution over the last 15 years. The table of contents of the four editions is provided in Table 5.1. Already from the second edition, one can notice the subtitle "Embedded Systems Foundations of Cyber-Physical Systems" and, coherently, the appearance of concepts such as modeling and system software. Despite a sharp distinction between embedded systems and CPS was not entirely intelligible, the awareness that CPS is a new discipline which needs to interface with other matters and *savoir-faire* is clear. The third edition extends the subtitle to "The Internet of Things" and introduces concepts such as multi-core platforms and the use of a flipped classroom-based teaching. The fourth edition is characterized by closer links with AI, and a more in depth investigation on the assurance of safety and security properties.

In 2020, an interesting survey [Mar+20] on CPS education was published by Peter Marwedel, Martin E. Grimheden, Tulika Mitra and Hugo A. Andrade. It compares several CPS educational programs worldwide. We observe that teaching of CPS not only requires the introduction of a single course, it demands an entire organization over several years under the constraints of a limited total number of course. In this regards, Peter Marwedel states in [MRM13]

> "Teaching embedded system design is difficult: Embedded system education is typically implemented as a concentration (specialization) within some other well-known program such as computer science or electrical engineering. [...]. With the introduction of CPS, we are passing the limits of this approach. It is therefore necessary to consider the introduction of an integrated CPS program."

The quest of which content for a CPS program rises with the quest for how to "link knowledge for the involved disciplines" [Mar+20].

In December 2020, the interesting article "Cyber-physical systems research and education in 2030: Scenarios and strategies" [BBT21] was published. The paper emphasizes the intrinsic interdisciplinarity

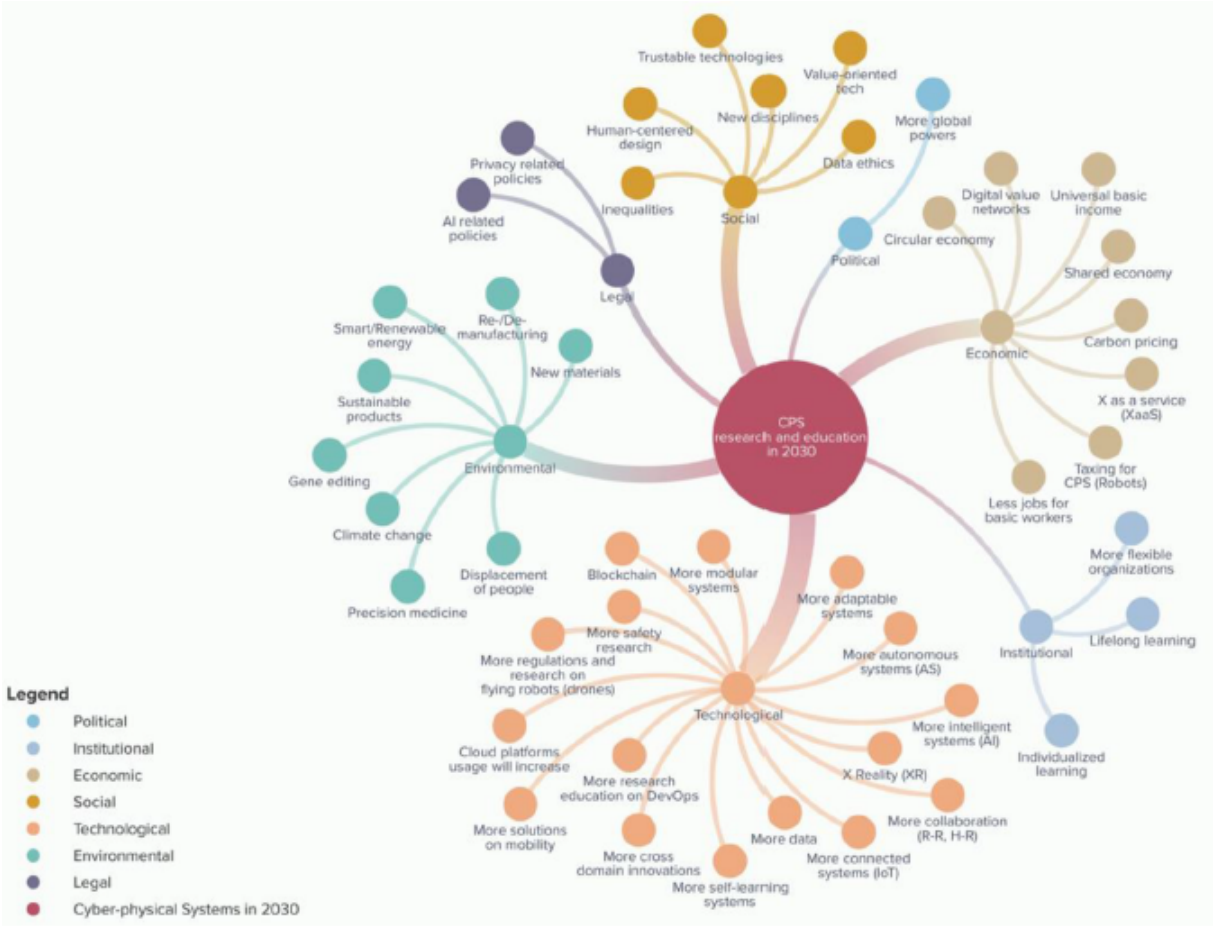


Figure 5.3 – Figure drawn from [BBT21]. The way forward to CPS interdisciplinarity over the 2030 horizon.

of CPS and the relationship between the different disciplines. The starting point of this study is the United Nations agenda for 2030, which highlights the political drive to achieve a list of objectives [Nat15]. This direction is directly related with education. It is precisely by analyzing this agenda that the authors identify some emerging challenges for CPS, such as the ecological sustainability of technological solutions. The authors state that "CPS is pushing the boundaries of traditional engineering education and we should start to consider the changes not only in technology but also in economics, politics, societies, and legal needs" [BBT21]. Figure 5.3 shows the relationship between the different disciplines, identified by the authors in [BBT21].

To the best of my knowledge, the relationship between the different disciplines is not yet structured, formalized or acquired and, therefore, there is not a firm feedback. I believe that interdisciplinary collaboration can play a decisive role for the development of CPS content, which is no longer based on the disciplinary segregation of the world as before, in which experts, be they professors, students, industrialists, researchers, were driven to a "siloed" work for a particular application domain and whose goal was to master that discipline alone. I believe that in order to face the societal challenges of the present and immediate future it is necessary to educate new generations in interdisciplinarity in the sense that the (integrated) union of the knowledge of the individual members of a team is greater than the mere sum of it. In other words, it is not that everyone has to know everything about everything, but it is the collaboration between the individuals of a team that is decisive in solving a problem.

From a pragmatic and factual point of view, I believe that this collaboration can be established through various actions.

From the **contents perspective**, a university-level CPS course should consider the design and the development of a realistic CPS, the use of fragments of videos of real CPS (mostly accessible on YouTube) and the study of CPS by adopting different viewpoints (e.g. design methodologies and processes, safety, included mixed criticality, modular pre-certification, mathematics frameworks). New CPS engineers should stay up to date with new emerging concepts joint with technical knowledge and scientific method. Finally, I suggest that new generation of engineers should be able to root theory in facts and to always get the proper feedback from the facts.

From the **curriculum management perspective**, it is necessary that the university or the head of masters give clear indications to professors and substitute teachers on what is the objective of their course and of the module, and give this overview. Students must be able to perceive this unity and overall vision which should also shine through the instructors. Since 2014 I have been teaching *System and Software Dependability* at the *Master of Nuclear Energy* (France) in two modules *Operation* and *Design*. During these years, we have had several meetings with the managers of the module, of the master and with the professors who were taking the course in the same module in the same period of the year, in order to identify the respective contents and to integrate our courses in a collaborative way to respond to the needs of the master and students. In 2019, my lecture series achieved the International Institute of Nuclear Energy (I2EN) accreditation of the module for a five-year period, from 2019 to 2024. This has been possible thanks to the cooperation with the professors of the module (3 courses) and the support of the master and option manager.

For the **supervision of master's or doctoral students**, a student can take advantage from a multicultural environment, thanks to more supervisors and/or industrial and scientific collaborations. In my experience, firstly as a PhD student and then as a PhD supervisor, it is possible to have multiple supervisors and I can affirm that the result, although restricted to a few students and my experience alone, is definitely positive. The tipping point is to clearly outline the role of the supervisors, who must themselves collaborate. For several years, I have collaborated with prof. Antonio Cammi (Politecnico di Milano, Department of Energy, CeSNEF - Enrico Fermi Center for Nuclear Studies) in the master course. We both provide some lectures to the same course. Our topics and knowledge-bases are different, and our approaches are slightly dissimilar. Over the years, we understood that our teaching can benefit from a closer collaboration among us. This is the reason why we took the joint supervision of the (former) PhD student, Christian Castagna. For the sake of ensuring a coherence into the PhD Thesis, we decided that the main supervision and research direction was provided by prof. Antonio Cammi, whereas my supervision was minor. The outcome of our approach does not concern only the student and his career, it overcomes — as expected — our way of teaching courses. Therefore, it is my intention to continue on this walk of fruitful collaborations between the university, research and industrial representatives.

5.3 Balance Between CPS and Innovation & Entrepreneurship?

In Section 5.1, we highlighted how innovation is one of the key elements of the European political agenda. To achieve innovation, education should open CPS topics not only to the aforementioned themes (e.g. modeling, D&D, system) but also to Innovation and Entrepreneurship (I&E), as a means to providing students the intellectual instruments in order to create 'disruptive innovative technologies'.

Coupling education on CPS and I&E has been provided by the EIT Digital Summer School in the years between 2014 and 2016. The 2014 edition was held in Trento (Italy), that of 2015 in KTH (Sweden) and the 2016 edition in Paris (France), in which I intervened in the role of

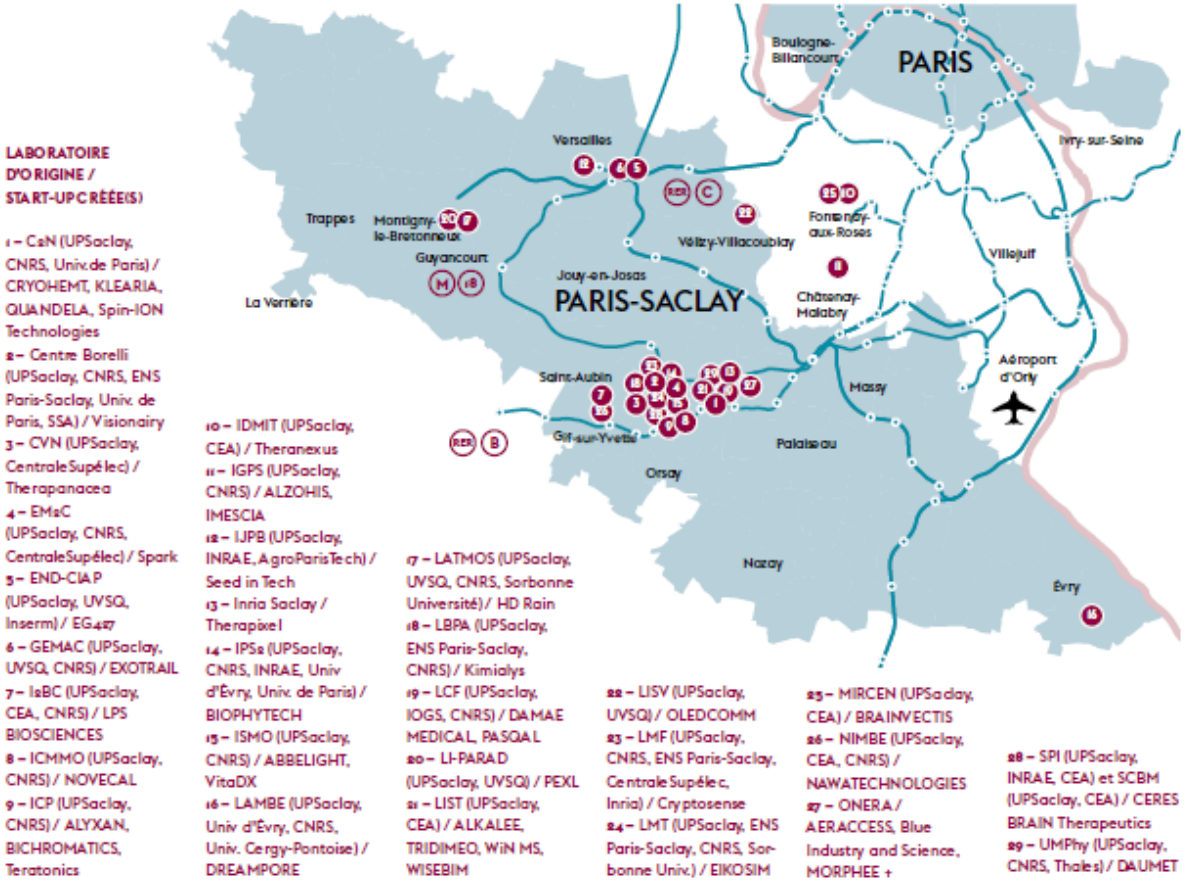


Figure 5.4 – Figure drawn from [Par21]. Start-Up rises from the Paris-Saclay Campus

Principal Investigator and Organizer. The feedback on the three years of experience was discussed in WESE [Can+16]. Our conclusion on the challenge of finding the best balance between I&E was that both of them need time to expect good results as well as a deep introduction and analysis. In my opinion, the balance between I&E and CPS contents is not fixed yet and it is still an open issue.

If we abstract from the CPS Summer School experience, and we move towards a CPS university course, we can suggest that education should provide training opportunity for full-immersion Entrepreneurship with valuable industries. The ability to integrate the internship into an industrial environment, capable of innovative projects, does not arise without an important management effort. The Paris-Saclay Campus is emblematic in creating a vibrant multi-disciplinary environment (industry, research laboratories, universities) and represents a success story. We have witnessed an innovation towards the creation of start-ups rise from research laboratories and universities, as shown in Figure 5.4 and an increasing number of students (about 40 %) that consider Entrepreneurship as foreseeable [Par21].

5.4 Which Pedagogical Education for CPS?

The WESE community also regarded the "Flipped Classroom" as a most appropriate means to teach CPS. The third edition of *Embedded System Design* [Mar18] adopted such a methodology. A flipped classroom consists of providing educational materials before the lecture. Students should read such materials before the CPS class.

Today, we have resources of great value at our disposal and it is unimaginable not to use them. For example, during my course, I use the videos, available on YouTube, provided by *Institut de Radioprotection et Sûreté Nucléaire* (IRSN) to introduce and analyze some nuclear accidents [IRS]. In the CPS topic, some high-potential videos are on Peter Marwedel's YouTube channel [Mar]. Moreover, I suggest the use of videos and texts, which show different viewpoints. For example, the book *A Nuclear Crisis in Historical Perspective: Three Mile Island*, written by Samuel Walker, introduces one of the most famous nuclear accidents, known with the acronym TMI-2. Although the accident did not cause deaths or contamination cases, it has been a societal shock, with the birth of the antinuclear movement, and modified the standardization process of nuclear safety standards. The book is important because it merges the accident in its historical context and provides technical details and images of the accident. In my experience, this perspective impresses students very much.

If the students access study materials at home before the class, what then becomes a classroom course? The classroom is the place where critical reasoning is deepened, open to interdisciplinary and creativity via team work and team building.

5.5 What Responsibility in Achieving Time-to-Market Innovation?

The idea of the European Commission to promote faster time-to-market innovation must be better reasoned.

Research has its own time: it can be made more efficient and productive, but it still requires time. What can and must be cut is what is around research.

Firstly, the administrative part toward entrepreneurship. The time between the idea or prototype of an innovation and the market must be faster. It plays a key role, for example, in the success of the American campus. This action requires a change in society and is the politics duty.

Secondly, research can integrate applications and look at innovative products. This last point has to do with education.

5.6 Conclusion

This chapter reflects my experiences on the CPS Educational Challenge over the last 8 years. Along this research trajectory, I want to strengthen my collaboration with universities and industries, which I consider a value added to the research formation of PhD students, as diversity is an essential ingredient for innovative and impacting researches to our ecosystem.

Chapter 6

Summary and Research Perspectives

Last summer I took a journey in the countryside, driving a hybrid cabriolet, *with all options including ADAS, that I had rented for the occasion.*

At some point, I found myself at a crossroad. The road sign indicated that my destination was on the right, unlike the cabriolet's navigator, which indicated the opposite direction.

This situation reminded me of a passage from the book D-Day, by Antony Beevor, 2009, in which the Allies had to decide the day of the landing in Normandy based on the weather forecast. They made the same inquiry to two different weather station teams, who provided two analyses that simply did not match up: "no two of the expert participants in the discussion could agree on the likely weather even for the next 24 hours". In reporting the result to the Eisenhower's chief of staff, unable to disclose this divergence, he said "the situation is complex and difficult".

To me, at that crossroad, it wasn't like that anymore. I acted as someone born in the new millennium probably would, placing trust in the technology in which they are immersed. I chose to follow the direction of the navigator because I believe that the engineers and designers of that system had all the tools to conceive and develop this technology at the best.

This *Mémoire* reflects my research trajectory after attaining the PhD in Computer Science in 2003 and over 15 years, collaborating with different teams in different cities and countries.

As discussed in Chapter 3 and 4, for the incoming period we expect a massive increase in mobile ACPS, with AI, in our personal, social and professional spheres. Such systems must respond to challenges arising outside of multiple disciplines and social concerns, such as the ecological impact of the safety-related solutions.

The way forward that I envision for my future research in this domain is to address methods and analysis techniques for the design and development of eco-sustainable and trustable architectures of AI-driven mobile and autonomous critical cyber-physical systems operating in highly dynamic environments.

In doing so, I focus my research activities on "Lifelong Exploration & Navigation" (LEN) for autonomous mobile robots, with artificial intelligence, in a semi-unknown environment. The decision-making process of LEN is done in a bounded time window, consuming a bounded amount of computational resources. LEN is developed at CEA and is patented by Laurent Soulier (CEA). LEN offers an excellent case study to experiment the achieved scientific results, because LEN is implemented in a mobile robot having AI. A first result on the architecture was published in IEEE Design and Test [May+19].

In the short term, I would like to understand what methods and techniques are needed to guarantee and ensure a sustainable and trustable low-level architecture, i.e from sensors to the occupancy grid

and the Generalized Voronoi Graph (GVG) and then back down to the actuators. Category and Graph Theory can help to reasoning about the system (position of the robot in the digital representation of the environment vs the position of the robot in the real environment). LEN is used as a means to reason on the problem and experiment with the scientific results.

In the short/medium term, I want to generalize the results of my PhD student, Emine Laarouchi, to the more complex case of LEN. In that PhD project, we have quantitatively demonstrated that the introduction of genetic algorithms improves or equals the safety levels of a system. The question is now if this result can be generalized to AI-based decision making and what is the impact on resilience. This investigation intends to contribute to Challenge "Decision making under uncertainty" introduced in the EU-Robotics road map 2030 - Technical Group Safety, where I actively contributed to: "Suitable (AI) decision-making, where the consequences are acceptable both for frequency and severity. We expect that results are proved with respect to measurable quantitative and heuristic baseline".

In the medium/long term, I want to study methods and techniques that can help make an architecture trustworthy and sustainable from the bottom up to the high-level architecture (i.e. to the high-level AI functionality of the robot). In this case, it is possible to exploit the whole architecture of LEN (from sensors to AI functionality and down-back to actuators), to understand if there exist some constraints and the associated cost to implement the measures to make the architecture trustworthy.

Throughout those efforts, I want to strengthen my collaboration with universities and industries, which I consider a value added to the research formation of PhD students, as diversity is an essential ingredient for innovative and impacting researches to our ecosystem.

Appendix A

Correctness by Construction in RCM

A complete analysis, discussion and proof can be found in my article [Can+10]. For the sake of completeness, I include here an extract of the aforementioned article tailored with the proof.

Attributes on the provided and required services are used to determine the structure of the run-time components [CPV08; CV09]. In this appendix we analyze the behavior of the threads of the RCM run-time components. We show that, assuming terminating services, the interaction between the components always results in the corresponding jobs to safely terminate. Because services are assumed to terminate, safe termination can be reduced to checking that no cycles are introduced in the call graph. In fact, certain cycles produce no harm, as they involve the activation of different jobs. We show next that we can use the graphs to check these conditions.

To do so, we must analyze all the possible sequences of invocations that are induced by the call graph, which in our case corresponds to the hypergraph HG_I of the Interface view. We start by looking at the possible ways in which each of the PIs in the system might be invoked. We can obtain this information starting from a PI and traversing the edges in the hypergraph backwards to follow the sequence of invocations.

We call this a *walk* in the hypergraph.

Definition 1 (Walk) *A* walk *is a sequence of the form*

$$PI_0\ x_0\ RI_1\ \delta_1\ PI_1\ x_1\ RI_2 \ldots PI_k,$$

such that $x_i = \langle RI_{i+1}, PI_i \rangle$ *and* $\delta_i = \{PI_i, \langle \ldots, RI_i, \ldots \rangle\}$.

In other words, a walk is a path in the hypergraph that represents the transitive closure of calls from service PI_k to service PI_0.

Walks are always finite (by definition), but their length is potentially unbounded, because of cycles, whose presence denounces a non terminating behavior. This situation in our system is prevented by the design tool, which analyzes the relevant walks in the specification to detect and proactively prevent recursion. We focus on the number of walks to be considered to only those that are potentially harmful. When a deferred PI is invoked, the calling thread never executes the sequence of calls originating from the deferred service, which are initiated by the called thread. We can account for this condition by considering only those walks that terminate as soon as they traverse a deferred PI (with the exception of, possibly, PI_0). We call this a *call sequence*.

Definition 2 (Call sequence) *A* call sequence *is a walk* $PI_0 \ldots PI_k$ *such that* PI_1 *to* PI_{k-1} *are all immediate,* PI_k *is nominal, and* PI_0 *is either deferred, or otherwise an immediate PI with no related RIs, i.e., there is no* $\delta_0 \in \Delta$ *such that* $\delta_0 = \{PI_0, \langle \ldots, RI, \ldots \rangle\}$ *for some* $RI \in I_R$.

Thus, a call sequence may include deferred PIs (which are associated to threads) only at the beginning and at the end of the sequence. In addition, a call sequence is maximal and can not be extended, since if PI_0 is immediate, then it has an empty set of RIs.

In some cases, cycles in a call sequence can be erroneously introduced by designers. However, they are always detected as soon as they occur and reported to the designer that should modify the design until it is cycle-free. Code is always automatically generated from a cycle-free design.

Lemma 1 *If a cycle is present in a call sequence, then it is possible to detect it.*

Proof 1 *Let $PI_0 PI_1 \dots PI_k$ be a call sequence. We have two cases.*

In the first case, a cycle is present in the call sequence if a PI is traversed twice. This occurs whenever there exist PIs PI_i and PI_j in the sequence, such that $i \neq j$ and $PI_i = PI_j$. This condition can be detected by traversing the call sequences originating from all deferred PIs in a depth-first manner, and verifying if an interface service has already been traversed. This process terminates since RCM has a statically defined number of PIs and RIs.

A cycle may still be present if PI_0 is deferred and $PI_0 = PI_k$. This case, however, does not pose any problem since the flow of control is broken by the presence of the OBCS access control structure in the RCM run-time component.

A "harmful" cycle is therefore one that involves only immediate services. When a cycle is detected, an error is automatically reported in the Interface view and the system is not validated until the cycle is removed by designers.

Given a deferred PI, we identify a subgraph of the hypergraph that includes all the services reachable from that PI.

Definition 3 (Thread graph) *Let $HG_I = (V_I, E_I)$ be a hypergraph and let PI_T.nominal $\in V_I$ be a nominal PI. A* Thread graph *$HG_T = (V_T, E_T)$ for PI_T.nominal is the subgraph of HG_I such that V_T and E_T include all and only the vertices and the edges that either belong to a call sequence that terminates in PI_T.nominal or belong to at least one call sequence that terminates in PI.deferred and there is a hyperedge $\psi \in E_I$ such that $\psi = \langle PI_T.\text{nominal}, PI.\text{deferred} \rangle$.*

Like call sequences, harmful cycles in a thread graph are always detected.

Proposition 2 *If a cycle is present in a thread graph, then it is possible to detect it.*

Proof 2 *The proof follows the same "structure" as that of Lemma 1. Each call sequence in a thread graph is analyzed. Harmful cycles, if any, are then detected by Lemma 1. Note that hyperedge Ψ does not introduce harmful cycles in a thread graph since it only links two deferred methods.*

Once harmful cycles are detected (if any), errors are automatically reported in the Interface view. The system is not validated until the designer removes all harmful cycles in the thread graph. Therefore, every thread activation always terminates without cycles, and threads are always able to return to their initial condition and be ready to service a new invocation.

We can extend our discussion to a composition of subsystems, in which required interfaces from one subsystem are linked to provided interfaces in the other. Cycles that begin and terminate at the same deferred PI are again harmless in that they represent the call flow of a sporadic thread whose execution produces the release event for its own next activation. The remaining possible cycles are ruled out by the tool, which immediately invalidates the composition if it detects a cycle in a call sequence within the same thread of control. In this case, the check can be done incrementally, by considering only the call sequences that span across the subsystems. There are, in particular, two cases of interest. In the first, a designer composes two open systems S_1 and S_2 so that each RI of one subsystem invokes a deferred PI of the other subsystem. While cycles may be generated, they are

again harmless, because the thread that belongs, for instance, to S_1 never executes directly any of the PIs of S_2, but always proceeds through the OBCS control structure (and vice-versa). In this case, the correctness of the system composition can be inferred compositionally from the correctness of the components, and no further verification is required. The second case is more complex, and involves the invocation of an immediate PI of one subsystem from one RI of the other. The compositionality principle cannot be applied in this case, and the tool is employed to detect and report the presence of potential cycles before the system can be validated.

The attributes set on the Functional and Interface view are sufficient for our tool to conduct the analysis and validate a system. The Implementation view is then automatically generated according to our RCM Interface Grammar [CV09], which guarantees the implementation of the services with the appropriate run-time components. The generated code is *by construction* guaranteed to be free of harmful cycles. (See Proposition 2).

Appendix B

Contribution to Publications

In this Appendix, I identify the specific type of contribution (whether in vision, technical content, writing, or project management) to the publications that I have co-authored, which I list here broken down by topic and by importance. To improve readability, Keynotes are always introduced after my publications.

B.1 Integrity Embedded System: Ravenscar Computational Model

[Can+10] Technical contribution: application contracts-based design to RCM; RCM grammar; Hypergraph Theory.
Concept [50%]; writing [33%]; project management [80%].

[CPV08] Technical contribution: composability in model transformations; RCM Grammar.
Concept [50%]; writing [33%]; project management [80%]

[CP08] Technical contribution: RCM grammar; Hypergraph Theory.
Concept [50%]; writing [50%]; project management [80%]

[CV09] Technical contribution: RCM Grammar.
Concept [50%]; writing [50%]; project management [20%]

[Can08] Keynote. Technical contribution and slides [100%]

B.2 Integrity Embedded System: Guaranteeing Correctness Under Heterogeneity

[Pas+09] Technical contribution: survey on the use of metamodels and tools in EU.
Implication of my colleagues at CEA in the paper [100%]; First draft [100%] of sections "A profile for safety analysis", "Defining formal execution descriptions", "Heavyweight vs. lightweight design" "Language design strategies" and "Introduction". writing [30%]

[Esp+09] Technical contribution: analysis of the syntax and semantics overlapping in SysML and Marte; state of the art.
Concept [50%]; writing [33%]; project management [80%]

[Can+09] Technical contribution: SOPHIA (conception and first implementation).
Concept [90%]; writing [80%]; project management [100%]

B.3 CPS: Roadmap and Dissemination

[ECS21] Project Management: Organization of "safety and resilience" experts team and "security and privacy" experts team [100%] - Chapter "Quality, Reliability, Safety and Cybersecurity".
Writing [100%] of Sections "Technology-Enable Societal Benefits"; "Major Challenge 3: ensuring cyber-security and privacy" and "Major Challenge 4: ensuring of safety and resilience" - Chapter "Quality, Reliability, Safety and Cybersecurity".
Regular meetings within the chapter and in the ECS-SRIA (with all chapter leaders) [100%]
Technical contribution [20%] and writing [20%] to the following chapters: "Embedded Software and Beyond"; "Architecture and Design: Methods and Tools"; "Long-Term Vision" (for Chapter "Quality, Reliability, Safety and Cybersecurity").

[Can+19] Technical contribution [100%] to Sections Introduction, Scientific Challenges of Safety-Critical ACPS, Ability to Appraise Safety Issues and to Learn from Experience, Certification, Privacy and Conclusions.
Concept [80%]; writing [33%]; project management [80%]

[Can+14] and [CG14] Papers originated from the organization of "Challenges and New Approaches for Dependable and Cyber-Physical System Engineering" (DeCPS),in the Ada-Europe International Conference on Reliable Software Technologies (AEiC 2014).
Concept [90%]; writing [50%]; project management [80%]

[Can15b] Keynote. Technical contribution and slides [100%]

[Can15a] Keynote. Technical contribution and slides [100%]

[Can12] Keynote. Technical contribution and slides [100%]

[Ant21] Technical contribution and slides [100%] on ECS-SRIA vision.

B.4 ACPS: Autonomous Metros

[Sou+16] Technical contribution: integration of safety and engineering collaborative process [50%]; methodology [50%]; Comparison with Existing Approaches [90%]; seamless and early verification based on the standard GSN (Global Structuring Notation) and a Contract- Based Design [90%].
Concept [90%]; writing [75%]; project management [60%]
[CSP14] Technical contribution: identification and application of contracts; First design and development of CAT (contract Analysis Tool) as Eclipse plugin, which is integrated with the industrial methodology and tools.
Concept [100%]; writing [50%]; project management [90%]

[Dal+10] Technical contribution: Sections Introduction [50%] Related work [70%] Methodology [50%] SysML model [100%]
Concept [80%]; writing [50%]; project management [20%]

B.5 ACPS: Autonomous Drones

[Mou+19] Project management: supervisor [100%] of the PhD student work for the Section Introduction; Related works; Drone Rescue System; Predictive Methodology ; Verification process.

[LCC17]. Technical contribution: design Methodology via 3D Iterative scenarios.
Concept [90%]; writing [30%]; project management: Teaching how to write a paper to PhD student [90%].

[CLB17]. Concept [100%]; writing [50%]; project management[100%] and supervisor of PhD student [100%].

[CS21]. Webinar at HiPEAC. Technical contribution: Introduction, Conclusion and on "Path planning embedded in autonomous drone"
Concept [95%]; writing [50%]; project management[50%].

B.6 ACPS: Autonomous Mobile Robots

[May+19] Technical contribution: contract-based approach and cross-layer CPS integration methodology.
Concept [80%]; writing [50%]; project management [90%]

[CCS21] Concept [100%]; script and video editing [100%]; project management [90%]

B.7 ACPS: Autonomous Connected Vehicles

[Pas+19] Technical contribution : safety analysis; compliance of the promoted methodology to safety norms.
Concept [95%]; writing [50%]; project management [90%]
[MCR17]. Technical contribution: discussion, perspectives and related works [50%]

B.8 ACPS: Nuclear Reactor

[CCC21] Technical contribution: introduction and compliance to the Nuclear Safety Regulations.
Concept [50%]; writing [50%]; project management [90%]

B.9 CPS Education Challenge

[Can+16] Technical contribution: organization of the CPS Summer School in 2016.
Concept [90%]; writing [50%]; project management [90%]

[CZP15] Technical contribution: master internship supervisor
Concept [90%]; writing [50%]; project management [90%]

Bibliography

[AD03] Christophe Aussaguès and Vincent David. *Introduction à la sûreté de fonctionnement dans OASIS (An introduction to safety related aspects of OASIS)*. Rapport de recherche RT/LIST/DTSI/SLA/03-448. CEA Commissariat à l'énergie atomique et aux énergies alternatives, 2003.

[AH01] Luca de Alfaro and Thomas A. Henzinger. "Interface Automata". In: *Proceedings of the Ninth Annual Symposium on Foundations of Software Engineering*. ACM Press, 2001, pp. 109–120.

[Amo+16] Dario Amodei et al. "Concrete Problems in AI Safety". In: *CoRR* (2016).

[Ant21] Marina Settembre Antonio Imbruglia Daniela Cancila. *Cybersecurity for interconnected systems - 5G" (Provided in Italian "Aspetti di cibersicurezza per sistemi interconnessi - 5G")*. Keynote at the Italian Association of Electrical, Electronics, Automation, Information and Communication, Technology - Society for Microelectronics Electronics Semiconductors (AEIT-AMES). 2021.

[ART] ARTEMIS. *Research Agenda 2016*. http://www.ecsel-austria.net/newsfull/items/artemis-strategic-research-agenda-2016-sra-31.html.

[Bal+98] Felice Balarin et al. "Scheduling for embedded real-time systems". In: *IEEE Design Test of Computers* 15.1 (1998), pp. 71–82.

[Bar12] José Manuel Barroso. "Science and innovation, an essential factor for competitiveness and growth in Europe". In: *Science and Innovation in Europe: the keys to tomorrow's growth*. 2012.

[BB04] Enrico Bini and Giorgio Buttazzo. "Schedulability Analysis of Periodic Fixed Priority Systems". In: *IEEE Transactions on Computers* 53.11 (2004), pp. 1462–1473.

[BBT21] Didem Gürdür Broo, Ulf Boman, and Martin Törngren. "Cyber-physical systems research and education in 2030: Scenarios and strategies". In: *Journal of Industrial Information Integration* 21 (2021).

[BCC15] BCC. *Seattle's Ferris wheel hit by drone*. `https://www.bbc.co.uk/news/technology-34797182`. 2015.

[BCP07] Albert Benveniste, Benoît Caillaud, and Roberto Passerone. *A Generic Model of Contracts for Embedded Systems*. Rapport de recherche 6214. Institut National de Recherche en Informatique et en Automatique, 2007.

[BN16] Greg Bensinger and Jack Nicas. *Uber to Put 100 Autonomous Volvo SUVs on Road in Pittsburgh*. The Wall Street Journal. 2016.

[Bon17] William B. Bonvillian. "The rise of advanced manufacturing institutes in the United States". In: *The Next Production Revolution: implication for gouvernments and business*. OECD, 2017, pp. 361–395.

[Bou19] Christhope Bounoux. *La voiture Autonome Valeo sur le périphérique parisien*. `https://www.youtube.com/watch?v=dd6pqIuhaHk`. 2019.

[BTD06] Subir Biswas, Raymond Tatchikou, and Francois Dion. "Vehicle-to-vehicle wireless communication protocols for enhancing highway traffic safety". In: *IEEE communications magazine* 44.1 (2006), pp. 74–82.

[Can+09] Daniela Cancila et al. "SOPHIA: A Modeling Language for Model-Based Safety Engineering". In: *Proceedings of the Workshop on Model based architecting and construction of embedded systems (ACES-MB), in conjunction with the MODELS conference*. 2009.

[Can+10] Daniela Cancila et al. "Toward Correctness in the Specification and Handling of Non-Functional Attributes of High-Integrity Real-Time Embedded Systems". In: *IEEE Transactions on Industrial Informatics* 6.2 (2010), pp. 181–194.

[Can+14] Daniela Cancila et al. "RoundTable on Challenges and New Approaches for Dependable and Cyber-Physical System Engineering (De-CPS)". In: *Ada-Europe International Conference on Reliable Software Technologies (AEiC 2014)*. Ada user Journal, 2014.

[Can+16] Daniela Cancila et al. "Experience and reflections on three years of CPS Summer School within EIT Digital". In: *Proceedings of the Workshop on Embedded Systems Education (WESE)*. ACM Digital Library, 2016.

[Can+19] Daniela Cancila et al. "Sharpening the Scythe of Technological Change: Socio-Technical Challenges of Autonomous and Adaptive Cyber-Physical Systems". In: *MDPI journal. Special issue on CPS Challenges* (2019).

[Can08] Daniela Cancila. *Correctness by Construction in the Model-Driven Engineering*. Keynote at Verimag. Invitation by J. Sifakis. 2008.

[Can12] Daniela Cancila. *Critical Systems*. Keynote at ITSLE international workshop. 2012.

[Can15a] Daniela Cancila. *CEA LIST 12:0. Systèmes Cyber-Physiques critiques: approach par contrat*. Keynote at CEA. 2015.

[Can15b] Daniela Cancila. *Contract-Based Design Tailored to Safety Issues for Cyber-physical Systems*. Keynote at the international conference on safety and security in IoT, conference EAI IoT. 2015.

[CCC21] Christian Castagna, Daniela Cancila, and Antonio Cammi. "Adoption of ACPS in Nuclear Reactor Analysis". In: *The 25th Ada-Europe International Conference on Reliable Software Technologies (AEiC 2021)*. Ada user Journal, 2021.

[CCS21] Daniela Cancila, Fabrice Mayran de Chamisso, and Laurent Soulier. *LEN: Lifelong Exploratory Navigation*. video and abstract for Workshop on Perception and Action in Dynamic Environments - International Conference on Robotics and Automation (ICRA 2021). 2021.

[CEAa] CEA. *Commissariat à l'Energie Atomique et aux Energies Alternatives*. http://www.cea.fr/.

[CEAb] CEA. *Exosquelettes*. `https://www.cea.fr/presse/Pages/actualites-communiques/sante-sciences-du-vivant/the-lancet-bci-clinatec-2019.aspx`.

[CEN03] CENELEC. *50129. Railway Applications - Communications, signaling and processing systems – Safety related electronic systems for signaling (CENELEC 50129)*. European Standards, 2003.

[CEN11] CENELEC. *50128. Railway applications - Communications, signalling and processing systems - Software for railway control and protection systems (CENELEC 50128)*. European Standards, 2011.

[CEN99] CENELEC. *EN-50126: Application ferroviaires -Spécification et démonstration de Fiabilité, Disponibilité, Maintenabilité et Sécurité (FMDS) (CENELEC 50126)*. 1999.

[CG14] Daniela Cancila and Jean-Louis Gerstenmayer. "Challenges and New Approaches for Dependable and Cyber-Physical System Engineering (De-CPS)". In: *Ada-Europe International Conference on Reliable Software Technologies (AEiC 2014),* 2014.

[Cha05] R. Chapman. "Correctness by construction: a manifesto for high integrity software". In: *SCS '05: Proceedings of the 10th Australian workshop on Safety critical systems and software*. Australian Computer Society, Inc., 2005.

[CLB17] Daniela Cancila, Emine Laarouchi, and Alessandra Bagnato. "Dependability of the Transport of the Future". In: *Ada-Europe International Conference on Reliable Software Technologies (AEiC)*. Ada user Journal, 2017.

[CLM15] F. Mayran de Chamisso, L. Soulier, and M. Aupetit. "Exploratory digraph navigation using A*". In: *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI)*. 2015.

[CMS] Daniela Cancila, Fabrice Mayran de Chamisso, and Laurent Soulier. *LEN: Lifelong Exploratory Navigation*. ICRA-PADE 2021. Video name "ICRA-PADE-2021-DaniConf7bis". Restricted Access to the HDR commission.

[CON] CONTREX FP7 Project. *Design of embedded mixed-criticality CONTRol systems under consideration of EXtra-functional properties*. `https://cordis.europa.eu/project/id/611146/fr`. Online; accessed: 10 September 2021.

[CP08] Daniela Cancila and Roberto Passerone. "Functional and Structural Properties in the Model-Driven Engineering Approach". In: *Proceedings of the 13^{th} IEEE International Conference on Emerging Technologies and Factory Automation (ETFA08)*. Hamburg, Germany, 2008.

[CPSa] CPSetis H2020 Project. *Towards Cyber-Physical Systems Engineering Tools, Interoperability Standardization (CP-Setis)*. `https://cp-setis.eu/`. Online; accessed: 10 September 2021.

[CPSb] CPSwarm Project. `http://www.cpswarm.eu/`.

[CPS17] CPSetis Project. *Strategic Agenda on Standardization for Cyber-Physical Systems. A proposal for an Update of the ARTEMIS Strategic Agenda for Standardization*. 2017.

[CPV08] Daniela Cancila, Roberto Passerone, and Tullio Vardanega. "Composability for High-Integrity Real-Time Embedded Systems". In: *Proceedings of the First Workshop on Compositional Theory and Technology for Real-Time Embedded Systems (CRTS 08)*. Barcelona, Spain, 2008.

[CS21] Daniela Cancila and Laurent Soulier. *LEN: Lifelong Exploratory Navigation*. Keynote at webminar HiPEAC. 2021.

[CSP14] Daniela Cancila, Elie Soubiran, and Roberto Passerone. "Feasibility Study in the use of contract-based approaches to deal with safety-related properties". In: *Ada-Europe International Conference on Reliable Software Technologies (AEiC 2014)*. Ada user Journal, 2014.

[CV09] Daniela Cancila and Tullio Vardanega. *RCM Interface Grammar*. Tech. rep. University of Padova, 2009.

[CyP13] CyPhERS FP7 Project. *CPS: State of the Art. Deliberable 5.1*. http://www.cyphers.eu/sites/default/files/D5.1.pdf. 2013.

[CyP14] CyPhERS FP7 Project. *Cyber-Physical European Roadmap and Strategy. Deliverable 5.1: State of the Art*. http://cyphers.eu/. Online; accessed: 18 Mai 2021. 2014.

[CZP15] Daniela Cancila, Hadi Zaatiti, and Roberto Passerone. "Cyber-Physical Systems and Contract-Based Approach". In: *Proceedings of the Workshop on Embedded Systems Education (WESE)*. ACM Digital Library, 2015.

[Dal+10] Stefano Dalpez et al. "An Industrial Case Study Using an MBE Approach: From Architecture to Safety Analysis". In: *13th IEEE International Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing Workshops*. 2010, pp. 116–122.

[DB05] Robert Davis and Alan Burns. "Hierarchical fixed priority pre-emptive scheduling". In: *26th IEEE International Real-Time Systems Symposium (RTSS'05)*. 2005, 10 pp.–398.

[DD00] Vincent David and J. Delcoigne. *Security Method Making Deterministic Real-Time Execution Of Multitasking Applications Of Control And Command Type With Error Confinement*. rance. WO 02/39277 A1. 2000.

[Del+17] Jerker Delsing et al. "The Arrowhead Framework architecture: Arrowhead Framework". In: Feb. 2017, pp. 43–88.

[Der+13] Patricia Derler et al. "Cyber-Physical System Design Contracts". In: *Proceedings of the ACM/IEEE 4th International Conference on Cyber-Physical Systems (ICCCPS)*. New York, NY, USA: Association for Computing Machinery, 2013, pp. 109–118.

[DMT00] Robert Davis, Nick Merriam, and Nigel Tracey. "How embedded applications using an RTOS can stay within on-chip memory limits". In: (Jan. 2000).

[Dou78] Douglas Adams. *The Hitchhiker's Guide To The Galaxy*. BBC (British Broadcasting Corporation), 1978.

[DRE] DREAMS FP7 Project. *Distributed REal-time Architecture for Mixed Criticality Systems*. https://www.uni-siegen.de/dreams/home/. Online; accessed: 10 September 2021.

[ECS] ECSEL (Electronic Components and Systems for European Leadership) industrial European research framework program. https://www.ecsel.eu/. Online; accessed: 10 September 2021.

[ECS21] ECS-SRIA. *Strategic Research and Innovation Agenda - Electronic Components and Systems*. 2021.

[Edw09] Edward Lee. *Introducing Embedded Systems:A Cyber- Physical Approach*. Invited Keynote Talk Workshop on Embedded Systems Education (WESE) https://ptolemy.berkeley.edu/projects/chess/pubs/619/EmbeddedSystemsEducation_WESE.pdf. Online; accessed: 08 September 2021. 2009.

[EMC] EMC2 Ecsel Project. *Embedded multi-core systems for mixed criticality applications in dynamic and changeable real-time environments*. https://www.artemis-emc2.eu/. Online; accessed: 10 September 2021.

[Esp+09] Huascar Espinoza et al. "Challenges in Combing SysML and MARTE for Model-Based Design of Embedded Systems". In: *Proceedings of the European Conference on Model Driven Architecture Fondations and Applications (ECMDA)*. Springer LNCS 5562, 2009, pp. 98–113.

[Eur13] European Commission. *Research and Innovation performance in EU Member States and Associated countries: Innovation Union progress at country level 2013*. Research and Innovation series, 2013.

[Eur15a] European Aviation Safety Agency. *Advance Notice of Proposed Amendment (A-NPA) 2015-10. Introduction of a regulatory framework for the operation of drones*. 2015.

[Eur15b] European Aviation Safety Agency. *Introduction of a regulatory framework for the operation of unmanned aircraft*. 2015.

[Eur16] European Aviation Safety Agency. *'Prototype' Commission Regulation on Unmanned Aircraft Operations*. 2016.

[FAA] FAA. *aircraft certification design approvals Technical Standard Order (TSO)*. https://www.faa.gov/aircraft/air_cert/design_approvals/tso/.

[Fra16] Michael D. Franusich. *Security Hardened Cyber Components for Nuclear Power Plants*. Tech. rep. Grant No. DE-SC0013808. US Department of Energy, Office of Science, Chicago Office, 2016.

[Fut] Future of Life. *AI Safety Research*. https://futureoflife.org/ai-safety-research/.

[Gei+12] Eva Geisberger et al. *Integrierte Forschungsagenda Cyber-Physical Systems*. Springer, 2012.

[GLN01] Paolo Gai, Giuseppe Lipari, and Marco Di Natale. "Minimizing memory utilization of real-time task sets in single and multi-processor systems-on-a-chip". In: *Proceedings 22nd IEEE Real-Time Systems Symposium (RTSS)*. 2001, pp. 73–83.

[Gon+01] M. Gonzáles Harbour et al. "MAST: Modeling and Analysis Suite for Real-Time Applications". In: *Proceedings of the Euromicro Conference on Real-Time Systems*. http://mast.unican.es. Delft, The Netherlands, 2001.

[Hal07] A. Hall. "Realising the Benefits of Formal Methods". In: *Journal of Universal Computer Science* 13.5 (2007), pp. 669–678.

[Ham] Hamed Zaghaghi. *Animation Temporal Verification*. http://www.foro3d.com/f230/animation-temporal-verification-77370.html.

[HB07] Vance Hilderman and Tony Baghai. *Avionics Certification: A Complete Guide to DO-178 (Software), DO-254 (Hardware): The Science of Microfabrication*. Avionics Communications Inc, 2007.

[HC02] A. Hall and R. Chapman. "Correctness by construction: developing a commercial Secure System". In: *Software, IEEE* 19 (2002), pp. 18–25.

[Her] Hercules. *Exosquelette mécanique Hercule V3*. https://www.directindustry.fr/prod/rb3d/product-180292-2299769.html.

[IEC00] IEC. *61508:1998 and 2000, part 1 to 7. Functional Safety of Electrical, Electronic and Programmable Electronic Systems*. 2000.

[IEC06] IEC. *IEC 60880 Nuclear power plants – Instrumentation and control systems important to safety – Software aspects for computer-based systems performing category A functions*. 2006.

[IEC11] IEC. *IEC 61513. Nuclear power plants – Instrumentation and control important to safety – General requirements for systems*. 2011.

[Int18] International Atomic Energy Agency. *IAEA Safety Glossary*. 2018.

[IRS] IRSN. *Institut de Radioprotection et Sûreté Nucléaire. YouTube Channel*. https://www.youtube.com/channel/UCBm7Wzuu7uEPDqE5sD775Sg.

[ISO18] ISO/TC 22/SC 32. *ISO-26262-1:2018 : Road vehicles – Functional safety*. Safety International Standard. International Organization for Standardization, 2018.

[KP17] Francisco Klauser and Silvana Pedrozo. "Big data from the sky: popular perceptions of private drones in Switzerland". In: 72 (2017), pp. 231–239.

[KRO] KRONO-SAFE. *Krono-Safe: Safe Design in real-Time*. https://www.krono-safe.com/.

[Laa+] Emine Laarouchi et al. *Genetic Algorithms based ACPS Safety*. Restricted Access to the HDR commission.

[Laa20] Emine Laarouchi. "A safety approach for CPS-IoT". PhD thesis. Institut Polytechnnique de Paris, 2020.

[LCa] Emine Laarouchi and Daniela Cancila. *CAT: Drone scenario*. Video name "DroneScenarioPathSuccess". Restricted Access to the HDR commission.

[LCb] Emine Laarouchi and Daniela Cancila. *Vehicle Platooning System with obstacle detection*. `https://www.youtube.com/watch?v=jSvLUz4hURM`.

[LCC17] Emine Laarouchi, Daniela Cancila, and Hakima Chouchi. "Safety and degraded mode in civilian applications of unmanned aerial systems". In: *IEEE/AIAA Digital Avionics Systems Conference (DASC)*. IEEE Interactive Electronic Library (IEL), IEEE Xplore, 2017.

[Lee07] Edward A. Lee. *Computing foundations and practice for cyber-physical systems: A preliminary report.* Technical Report UCB/EECS-2007-72. EECS Department, University of California, Berkeley, 2007.

[Lin] Linked Data community. *Linked Data - Connect Distributed Data across the Web*. http://linkeddata.org/.

[LK19] Kibeom Lee and Dongsuk Kum. "Collision Avoidance/Mitigation System: Motion Planning of Autonomous Vehicle via Predictive Occupancy Map". In: *IEEE Access* 7 (2019).

[LM13] Audrey Linkenheld and Jacques Myard. *Rapport d'Information: déposé par la commission des affaires européennes sur le huitième programme-cadre pour la recherche et l'innovation « Horizon 2020 »*. Tech. rep. 1009. ASSEMBLÉE NATIONALE, 2013.

[LS10] Edward Lee and Sanjit Seshia. "An Introductory Textbook on Cyber-Physical Systems". In: *Proceedings of the Workshop on Embedded Systems Education (WESE)*. WESE '10. Association for Computing Machinery (ACM), 2010.

[Mar] Peter Marwedel. *cyphysystems. YouTube Channel*. `https://www.youtube.com/channel/UCAk3i_XTihlAF-UzOAAatbA`.

[Mar+20] Peter Marwedel et al. "Survey on Education for Cyber-Physical Systems". In: *IEEE Design & Test* 37.6 (2020), pp. 56–70.

[Mar03] Peter Marwedel. *Embedded System Design*. Kluwer Academic Publishers, Dordrecht, 2003.

[Mar11] Peter Marwedel. *Embedded System Design:Embedded Systems Foundations of Cyber-Physical Systems*. Kluwer Academic Publishers, Dordrecht, 2011.

[Mar18] Peter Marwedel. *Embedded System Design: Embedded Systems Foundations of Cyber-Physical Systems, and the Internet of Things*. Embedded Systems. Springer International Publishingt, 2018.

[Mar21] Peter Marwedel. *Embedded System Design: Embedded Systems Foundations of Cyber-Physical Systems, and the Internet of Things*. Embedded Systems. Springer International Publishingt, 2021.

[Mata] MathWorks. *MATLAB: Maths. Graphiques. Programmation*. `https://fr.mathworks.com/products/matlab.html`.

[Matb] MathWorks. *Simulink: Simulation et Model-Based Design*. https://fr.mathworks.com/products/simulink.html.

[May+19] Fabrice Mayran de Chamisso et al. "Lifelong Exploratory Navigation: an Architecture for Safer Mobile Robots". In: *IEEE Design and Test* (2019).

[May16] Fabrice Mayran de Chamisso. "Lifelong Exploratory Navigation". PhD thesis. University of Paris-Saclay, 2016. URL: http://www.theses.fr/en/2016SACLS413.

[MCR17] Sebti Mouelhi, Daniela Cancila, and Amar Ramdane-Cherif. "Distributed Object-Oriented Design of Autonomous Control Systems for Connected Vehicle Platoons". In: *IEEE International Conference on Engineering of Complex Computer Systems (ICECCS)*. 2017.

[ME11] Peter Marwedel and Michael Engel. "Embedded System Design 2.0: Rationale Behind a Textbook Revision". In: *Proceedings of the Workshop on Embedded Systems Education (WESE)*. WESE '11. Association for Computing Machinery (ACM), 2011.

[Mil17] Steve Miletich. *Pilot of drone that struck woman at Pride Parade gets 30 days in jail*. https://www.seattletimes.com/seattle-news/crime/pilot-of-drone-that-struck-woman-at-pride-parade-sentenced-to-30-days-in-jail/. 2017.

[Mou+19] Sebti Mouelhi et al. "Predictive Formal Analysis of Resilience in Cyber-Physical Systems". In: *IEEE Access* (2019).

[MRM13] Peter Marwedel, Wolfgang Rhode, and Katharina Morik. "The TU Dortmund Cyber-Physical Systems Program A Step Towards Multi-Disciplinary Education". In: *Proceedings of the Workshop on Embedded Systems Education (WESE)*. WESE '13. 2013.

[MSA16] Fabrice Mayran de Chamisso, Laurent Soulier, and Michael Aupetit. "Robust topological skeleton extraction from occupancy grids for mobile robot navigation". In: *Proceedings of the twentieth national congress on Shape Recognition and Artificial Intelligence (RFIA'16)*. 2016.

[MW19] Ke Ma and Hao Wang. "Influence of Exclusive Lanes for Connected and Autonomous Vehicles on Freeway Traffic Flow". In: *IEEE Access* 7 (2019).

[Nat] National Science Foundation. *Cyber-Physical Systems (CPS), Program Solicitation, NSF 08-611*. https://www.nsf.gov/pubs/2008/nsf08611/nsf08611.htm.

[Nat12] National Institute of Standards and Technology (NIST). *Cyber-Physical Systems*. the NIST CPS Workshop. https://ptolemy.berkeley.edu/projects/cps/. Online; accessed: 10 September 2021. 2012.

[Nat15] United Nations. *Transforming our world: the 2030 Agenda for Sustainable Development. In The General Assembly*. Tech. rep. A/RES/70/1. 2015.

[Nat21] National Institute of Standards and Technology (NIST). *Cyber-Physical Systems*. https://www.nist.gov/el/cyber-physical-systems. Online; accessed: 10 September 2021. 2021.

[Nie+16] J. Nie et al. "Decentralized Cooperative Lane-Changing Decision-Making for Connected Autonomous Vehicles". In: *IEEE Access* 4 (2016), pp. 9413–9420.

[Nor96] Douglass C. North. "Economic Performance Through Time: The Limits to Knowledge". In: *Economic History* (1996).

[NS14] Marco Di Natale and Alberto Sangiovanni-Vincentelli. "Are we Losing Focus on the Cyber-Physical Aspects in the CPS Research Agenda?" In: *CPS20: CPS 20 years from now*. Online; accessed: 10 September 2021. CyPhERS, Cyber-Physical European Roadmap and Strategy, 2014.

[Obs13] Observatory Of Automated Metros. *Annual World Report*. Tech. rep. UITP, 2013. URL: http://metroautomation.org/about-the-report/.

[OMGa] OMG. http://www.omg.org/.

[OMGb] OMG. *UML Profile for MARTE: Modeling and Analysis of Real-Time Embedded systems*. www.omgmarte.org.

[OMGc] OMG. *Unified Modeling Language*. https://www.uml.org/.

[Par21] Paris-Saclay. *L'édition de l'université paris-saclay: été 2021*. Tech. rep. 16. University of Paris-Saclay, 2021.

[Pas+09] Roberto Passerone et al. "Metamodels in Europe: Languages, Tools, and Applications". In: *IEEE Design & Test of Computers* 26.3 (2009), pp. 38–53.

[Pas+19] Roberto Passerone et al. "Methodology for the Design of Safety-Compliant and Secure Communication of Autonomous Vehicles". In: *IEEE Access* (2019).

[Pen] Penta industrial European research framework program. https://penta-eureka.eu/. Online; accessed: 10 September 2021.

[Pin+06] Alessandro Pinto et al. "System Level Design Paradigms: Platform-Based Design and Communication Synthesis". In: *ACM Transactions on Design Automation of Electronic Systems* 11.3 (2006), pp. 537–563.

[Pla] Platform4CPS European project. https://www.platforms4cps.eu/.

[Pre] Prefecture of Paris. https://twitter.com/prefpolice/status/956931690978594817. [Online; accessed 12-September-2021].

[PRO] PROXIMA FP7 Project. *Probabilistic real-time control of mixed-criticality multicore and manycore systems*. http://proxima-project.eu/. Online; accessed: 10 September 2021.

[Pto14] Claudius Ptolemaeus, ed. *System Design, Modeling, and Simulation using Ptolemy II*. Ptolemy.org, 2014. URL: http://ptolemy.org/books/Systems.

[RB3] RB3D. *Exosquelette*. http://www.rb3d.com/fr/exosquelettes.

[RDT15] Stuart J. Russell, Daniel Dewey, and Max Tegmark. "Research Priorities for Robust and Beneficial Artificial Intelligence". In: *AI Magazine* 36.4 (2015).

[ROS] ROS. *The Robot Operating System - ROS2*. https://docs.ros.org/.

[Ros04] Nathan Rosemberg. *Innovation and Economic Growth*. OECD. 2004.

[San07] Alberto Sangiovanni-Vincentelli. "Quo Vadis, SLD? Reasoning About the Trends and Challenges of System Level Design". In: *Proceedings of the IEEE* 95 (2007), pp. 467–506.

[San17] Sandro D'Elia. *CPS in EU programmes*. European Commission DG CONNECT. 2017.

[SAR] SARTRE EU Project. *Safe Road Trains for the Environment*. https://www.roadtraffic-technology.com/projects/the-sartre-project/.

[Sch06] D. Schmidt. "Model-Driven Engineering". In: *IEEE Computer* (2006), pp. 25–31.

[SD90] S. Sheikholeslam and C. A. Desoer. "Longitudinal Control of a Platoon of Vehicles". In: *1990 American Control Conference*. 1990, pp. 291–296.

[SDP12] Alberto Sangiovanni-Vincentelli, Werner Damm, and Roberto Passerone. "Taming Dr. Frankenstein: Contract-Based Design for Cyber-Physical Systems". In: *European Journal of Control* 18.3 (2012), pp. 217–238.

[Sel07] B. Selic. *From Model-Driven Development to Model-Driven Engineering*. Keynote talk at ECRTS'07. http://feanor.sssup.it/ecrts07/keynotes/k1-selic.pdf. 2007.

[Sif05a] Joseph Sifakis. "A Framework for Component-Based Construction Extended Abstract". In: *Proceedings of the Third IEEE International Conference on Software Engineering and Formal Methods (SEFM'05)*. IEEE Computer Society, 2005, pp. 293–300.

[Sif05b] Joseph Sifakis. "Embedded Systems - Challenges and Work Directions". In: *Principles of Distributed Systems*. Ed. by LNCS. Vol. 3544. 2005.

[Sil18] Silvana Pedrozo and Francisco Klauser. *Drones policiers : Une acceptabilité controversée*. https://www.espacestemps.net/articles/drones-policiers-une-acceptabilite-controversee/. 2018.

[Sou+16] Elie Soubiran et al. "Ensuring Dependability and Performance for CPS Design: Application to a Signaling System". In: *Cyber-Physical Systems: Foundations, Principles and Applications*. Elsevier Inc., 2016, pp. 363–375. URL: https://hal-cea.archives-ouvertes.fr/cea-01818394.

[Ste] Stephen Hawking and Max Tegmark and Stuart Russell and Frank Wilczek. *Transcending Complacency on Superintelligent Machines*. Huffington Post, April 19, 2014.

[Tah+13] Walid Taha et al. "Experiences with A First Course on Cyber-Physical Systems". In: *Proceedings of WESE workshop*. 2013.

[The] The Eclipe Foundation. https://www.eclipse.org/org/.

[Tho+18] Haydn Thompson et al. *Platforms4CPS: Key Outcomes and Recommendations*. 2018.

[Tho16] Haydn Thompson. *Mixed-Criticality Cluster Portfolio Analysis. Report from the Final Mixed-Criticality Workshop*. 2016.

[TM16] Alireza Talebpour and Hani S. Mahmassani. "Influence of connected and autonomous vehicles on traffic flow stability and throughput". In: *Transportation Research Part C: Emerging Technologies* 71 (2016), pp. 143–163.

[Tör+16] M. Törngren et al. "Characterization, Analysis, and Recommendations for Exploiting the Opportunities of Cyber-Physical Systems". In: *Cyber-Physical Systems. Foundations, Principles and Applications*. Ed. by Houbing Song et al. Intelligent Data Centric Systems. Academic Press, Elsevier, Sept. 2016. Chap. 1, pp. 3–14.

[Upp] Uppsala University and Aalborg University. *UPPAAL*. https://uppaal.org/.

[Var06] Tullio Vardanega. "A Property-Preserving Reuse-Geared Approach to Model-Driven Development". In: 12^{th} *IEEE Int. Conf. on Embedded and Real-Time Computing Systems and Applications*. IEEE, 2006, pp. 223–230.

[W3C] W3C. *Semantic Web*. http://www.w3.org/2001/sw/.

[WES] WESE. *Workshop on Embedded and Cyber-Physical Systems Education*. https://www.kth.se/mmk/mechatronics/2.75564/wese-2019/call-for-papers-1.893525.

[Xec] Xecs industrial European research framework program. https://eureka-xecs.com/. Online; accessed: 10 September 2021.

[ZC] Hadi Zaatiti and Daniela Cancila. *CAT: Passenger exchange scenario*. AFIS 2015. Video name "AFIS-2015". Restricted Access to the HDR commission.

[ZLL18] C. Zhai, F. Luo, and Y. Liu. "Cooperative Look-Ahead Control of Vehicle Platoon for Maximizing Fuel Efficiency Under System Constraints". In: *IEEE Access* 6 (2018), pp. 37700–37714.