Project EC2.20

ONNX - a safety related profile

contact@confiance-ai.fr | www.confiance.ai

Specific information to be provided in ANR reports only

Progress Report		Final Report	
From:	*AAAA/MM/JJ*	To:	*AAAA/MM/JJ*
Brief description of the project	********		
Contractual description of the project:		********	
Reference budget:		********	

Document reference: XXX

Contributors

	Name	Organisation	Role
Responsible for the deliverable	TEULIERES Marie-Charlotte	Airbus Protect	
Co-author	GABREAU Christophe	Airbus	

Document Control

Revision	Date	Commentary	Author
V0.0	30/08/24	Delivery	
V0.1			
V0.2			
V1.0			

Contents

Chapter A

Introduction

Embedding machine learning technique in safety related systems implies to demonstrate its conformity to standards and regulation authorities. Among others, it is necessary to ensure and demonstrate that the Machine Learning model implementation process preserves the safety / functional / operational properties of the model developed during the design process. The work presented has been based on the ongoing standard discussed within the joint standardization group EUROCAE WG-114 / SAE G-34, which introduced the notion of MLMD, Machine Learning Model Description, to propose an accurate and precise description of the ML Model. After evaluating several format to support the MLMD, ONNX has been selected. As it does not fulfill all the requirements for regulations, a profile should be specified for safety related systems.

Chapter B

Summary of discussions

B.1. Confiance.AI working group on MLMD

The working group on the MLMD[1] and Safety related profile of ONNX is composed of industries, such as Airbus, Embraer, Thales, Safran, and researchers such as IRT SystemX, IRT Saint Exupéry, Inria, ONERA, CEA. We meet every 2 weeks, with the following objectives :

- Based on the work previously done to identify gaps between the format and the standard requirements, refine and develop the constraints imposed by the standards and the industry.
- Bring together different domains, industries, and research institutes to support our work.
- Propose and discuss potential solutions, identifying their advantages and disadvantages.
- Construct the baseline discussions for specification of the profile.
- Create a working group within the ONNX community to be fully integrated to the environment.

B.2. Requirements and needs

In the development process of Machine Learning, as described in the ED-324 / ARP6983, the MLMD is the interface between the design and the implementation process. "It should contain sufficient details on the ML Model semantic to fully preserve this semantic in the implemented ML Model" [arp6983,]

The MLMD should support the validation and verification, but primilarly it has to support the exact, or approximative, replication of the model for the implementation process.

It is necessary to identify :

- what is needed in order to express an explicit and precise description of the semantic
- what is useful in the development process

[1] In the document, MLMD is the acronym for Machine Learning Model Description

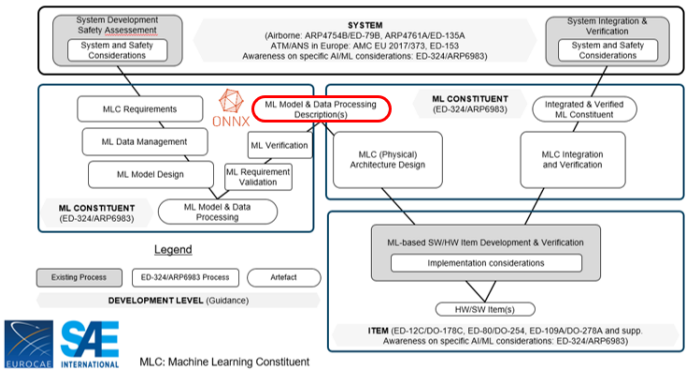

Figure B.1: W cycles described in the ED-324/ARP6983

The MLMD should include :

- The operator semantics,
- the graph semantic,
- the datatypes,
- the ONNX abstract and concrete syntax.

B.3. Specification of the semantic

The specification of the semantic is the core element of the Machine learning Model Description. In order to identify needs and illustrate the requirements, the work of specification has been done on the convolution layer. It should contain :

- A definition of the operator role and behaviour (textual),
- A description of Inputs, and Attributes,
  - Definition,
  - If it implies a modification of the semantic, specification of the impact,
  - Constraints, and dependencies with other attributes,
- Textual, or abstract specification, could be mathematical, textual etc.
- Formal specification,
- Reference implementation,

Concerning the formal specification, two solutions are being discussed : ACSL and C, and Why3, see figure B.2. ACSL and C, has the advantage to be well-known by the community and so easy to use, whereas Why3 combine a global solution, both the reference implementation and

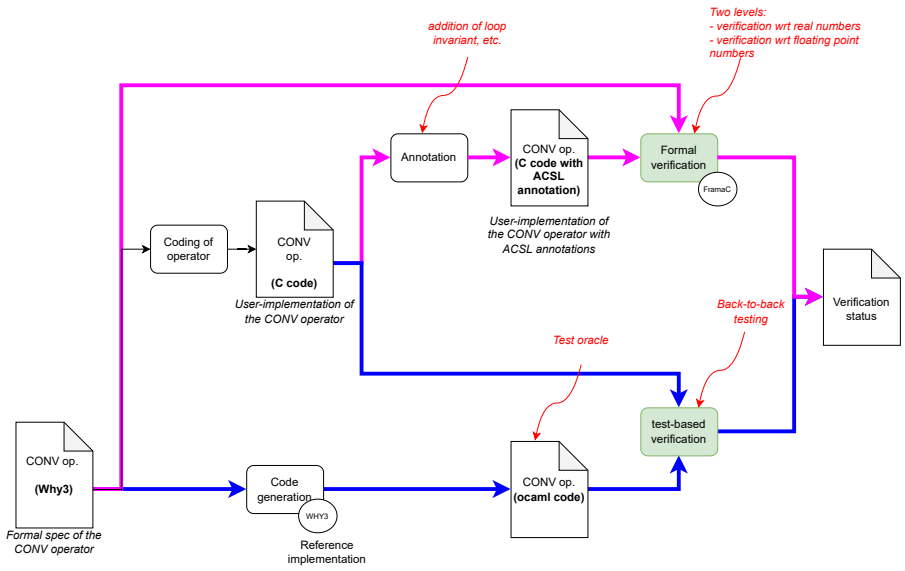

Figure B.2: Reference implementation workflow on proposed solutions

a formal specification. However, as a first step, manual specification ("pencil and paper") could be helpful to understand the behaviour and fix the needs.

The reference implementation could be used as a test oracle, in order to perform test-based verification.

B.4. Syntax and readability

The human readable objective depends on the definition of "human readable" and what is expected of the concept.
The definition and requirements, user needs depend on the task to be performed. Depending on the objective, it should be more or less precise, and containing specific information. Needs on the precision and content depend on the task that can be : design, certification argumentation, validation and verification, implementation etc. Moreover it should be auditable.
It is difficult to have a human readable form that matches all the requirements. The conclusion of our discussions is that the MLMD community should ensure that the model record all the request information in order for all users to develop their own visualization tool. The MLMD community has to ensure that the model will be interpreted the same way between activities, even though the visualization is not the same.
The main material for synthax description available on ONNX GitHub is the Intermediate Representation description.

B.5. Scheduling

The scheduling notion represents the sequencing of operations, and data flow semantic expression. It takes place as a transition between the MLMD and MLMIDs[2] (Machine Learning Model Item Description). MLMD can be decomposed into several items, hardware or software that are connected to each other. These connection, and this decomposition should be specified. The work on [Gauffriau, 2024] develops the notion of expressing scheduling using Petri Nets over NNEF. An exemple is given in figure B.3

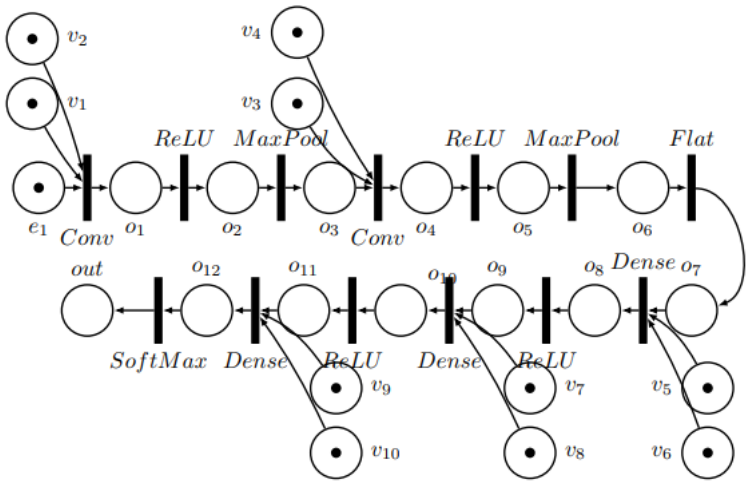

Figure B.3: NNEF semantics of the LeNet-5 express with Petri net. Initial marking. From [Gauffriau, 2024]

It describes two main objectives :

- Express all possible implementation, specify which implementations are valids. Define a static scheduling. In the example : Paths of the petri nets specify all the possible implementation. And fulfill the need of DataFlow semantic expression which is not yet proposed by NNEF (nor ONNX).
- There are several ways to implement Neural Networks. For example on how to split Neural Networks according to Hardware items availables. The objective is to demonstrate split is coherent to description and semantic. In the example, over NNEF, proposition to include subgraphs and extension of the syntax to express interaction between subgraphs. It introduces a colored Petri Nets to be compared to initial petri nets demonstrating all possible implementations in order to prove the preservation of the semantic.

The objective of the MLMD is to define the semantic of the model without ambiguity. The need to express the sequencing is a derived requirement from the implementation. It is an additional information not included in the semantic. Moreover, the notion of MLMID should be

[2] defined in the [arp6983,]

precisely defined with its scope. The challenge is to determine at what is an item, and at which level the MLMID operates, ie a unit, a core etc.

B.6. Some challenges

Floating Point numbers Floating point numbers introduce approximation in the model that can lead to errors. It is not possible to not specify any accuracy, that would mean that there is no error boundaries, but the accuracy of computation may be complicated.
Proposed solution : Specify the generic specification in the real domain, and some specific specifications for different data types.

Chapter C

1er Chapitre

C.1. ONNX Working Group : Safety-related profile

To continue the work beyond Confiance.AI, a working group has been proposed and accepted within ONNX. The group will be led by Eric Jenn and Jean Souyris. All information on the working group can be found on github : here.

The objective of the working group is to provide a definition of the formalism used to represent a trained ML model:

- *that has an understandable and non-ambiguous syntax and semantics.*
  The description of the ML model expressed using this formalism must be a Low-level Requirement for the implementation phase in the sense that its interpretation and implementation shall not require any further information that the one given by the description of the ML model.

- *that allows multiple levels of accuracy and precision for the description of a given model.*
  The language used to describe the model (i.e., its syntax and semantics) must be non-ambiguous, but a model may be ambiguous if this ambiguity is acceptable or even necessary to leave some freedom to the implementer (e.g., for optimization). The objective is to identify, control, and possibly remove this ambiguity by an appropriate refinement of the ML model description.

The kick-off meeting is scheduled on the 26th of September, 2024, to participate please contact **Eric Jenn**.

Three main tasks has been identified :

- Capture the needs and elicit the requirements for the safety-related profile (not limited to avionics)
- Analyse the ONNX standard with respect to the requirements
- Develop the "safety-related profile" for ONNX

Bibliography

[arp6983,] arp6983. *ED-324 | ARP6983 : Process Standard for Development and Certification/Approval of aeronautical Safety-Related Products implementing Artificial intelligence*.

[Gauffriau, 2024] Gauffriau, Adrien De Albuquerque Silva, I. P. C. (2024). Formal description of ml models for unambiguous implementation. *12th European Congress on Embedded Real Time Software and Systems (ERTS 2024)*.

Titre: ONNX : un profile pour les systèmes critiques

Mots Clés: Confiance, MLMD, implémentation, certification, systèmes critiques

Abstract: La description du modèle de Machine Learning est un élément clef dans le cycle de développement d'un composant. En effet, il fait l'interface entre la phase de design, et l'implémentation pour assurer la réplication des propriété. ONNX ne couvre actuellement pas tous les besoins concernant la description pour des systèmes critiques, qui doit comprendre une sémantique explicite et claire, et une synthaxe définie. La création du groupe de travail a pu mettre en avant les éléments haut-niveau qui permettent de décrire un modèle de machine learning. Ce qui a aboutit à l'intégration du groupe à la communauté ONNX.

Title: ONNX : a safety related profile

Keywords: Trust, Safety critical, MLMD, implementation

Abstract: The description of the Machine Learning model is a key element in the development cycle of a component. It serves as the interface between the design phase and implementation to ensure the replication of properties. Currently, ONNX does not fully address the needs for description in critical systems, which must include explicit and clear semantics, as well as a defined syntax. The creation of the working group highlighted the high-level elements necessary to describe a machine learning model, leading to the group's integration into the ONNX community.

Our partners: