FOCETA – Foundations for Continuous Engineering of Trustworthy Autonomy

Project Overview and Early Results

Chih-Hong Cheng

DENSO AUTOMOTIVE Deutschland GmbH

ISO 26262 Digital Conference

DENSO
Crafting the Core

Foundations for
Continuous Engineering
Trustworthy Autonomy

FOCETA

Goals and Partners

The breakthrough targeted by FOCETA is a **practical** and **demonstrably effective** methodological framework for the **formal modelling and verification of dependable learning-enabled components**, and for the **rigorous design of learning-enabled autonomous systems**. The new method will combine the advantages of data-based and model-based techniques towards ensuring Safety, Security and improving Performance, at lower costs and in shorter time.

H2020 ICT-50-2020 - Software Technologies

Agenda

- Some of the FOCETA addressed challenges
  - Introduce essential dependability and performance **specifications** for learning-enabled autonomous systems
  - Scalable **verification** techniques for learning-enabled components
  - Unified approach that combines **learning** from data and synthesis from specification
  - **Monitoring** the faithfulness of decision-making at runtime
- Concluding remarks

Agenda

- Some of the FOCETA addressed challenges
  - Introduce essential dependability and performance specifications for learning-enabled autonomous systems
  - Scalable verification techniques for learning-enabled components
  - Unified approach that combines learning from data and synthesis from specification
  - Monitoring the faithfulness of decision-making at runtime
- Concluding remarks

Requirement for learning-enabled components

- Element out-of Context (EooC)
  - Assumptions
  - Guarantees
  - List of triggering conditions that may lead to functional insufficiencies
  - …
- Research questions:
  - How far can we go with formal specification?
    - Background: to integrate the GSN-subtree of EooC to the main GSN
  - Can formalized specification sharpen our understanding?
    - E.g., decide the proper validation target

Perception (Object Detection)

S.ObjDetect	**UC-AVP 01**	**Functional**	**Within X meters of range from the Ego vehicle, the object detection component shall identify pedestrians in their correct position.**

Source: L. Gauerhof, R. Hawkins, C. Picardi, C. Paterson, Y. Hagiwara, I. Habli: Assuring the Safety of Machine Learning for Pedestrian Detection at Crossings. SafeComp'20, pp 197-212

Let's try to formulate it more precisely

Perception (Object Detection)

S.ObjDetect	**UC-AVP 01**	**Functional**	**Within X meters of range from the Ego vehicle, the object detection component shall identify pedestrians in their correct position.**

Source: L. Gauerhof, R. Hawkins, C. Picardi, C. Paterson, Y. Hagiwara, I. Habli: Assuring the Safety of Machine Learning for Pedestrian Detection at Crossings. SafeComp'20, pp 197-212

In one brainstorming session:

```
def (sensor) If sensor <=5
    if (detection() = 1) return 1
```

Always 'Detected correctly' **Until** 'distance>X'

Object detection is **Always** True when object distance in [0, X]

the requirement is under-specified: depends on how realistic we want to be, need to know how frequent we want to execute.

Always(y \in detected() iff \exists x. Pedestrian(x) and distance(ego, x) <= X and and distance(x,y) <= epsilon)

```
main := forall x in Objects,
 ALWAYS ( range(x) -> pedestrian(x) -> detect(x) )
range(x) := distance(x,ego) <= X
detect(x) := ODC_pedestrian(x) &&
| ODC_position(x) - position(x) | < epsilon
```

\forall Pd, Ps, X : ped_pos(Pd, Ps) /\ ego_range(X) /\ in_range(Ps, X) -> identified(Pd)

Lead to

- Proper **hiding** of formal specifications, if one wants to use it
- Need for a central **ontology** for specification, together with tool support

Perception (Object Detection)

S.ObjDetect	**UC-AVP 01**	**Functional**	**Within X meters of range from the Ego vehicle, the object detection component shall identify pedestrians in their correct position.**

∀ pedestrian:
always (is_in_X_meters(pedestrian)
→ (od_identified(pedestrian)
&&
| od_pred_location(pedestrian) – act_location(pedestrian) | <= K)

Lead to

- [Unrealizability issues] Some pedestrians can be occluded by all sensors.
- [Monitorability issues] There exists certain predicates that are associated with the ground-truth

Agenda

- Some of the FOCETA addressed challenges
  - Introduce essential dependability and performance specification for learning-enabled autonomous systems
  - Scalable verification techniques for learning-enabled components
  - Unified approach that combines learning from data and synthesis from specification
  - Monitoring the faithfulness of decision-making at runtime
- Concluding remarks

Verification techniques covered by FOCETA

- Testing
  - Model-based testing (this talk)
  - Additional coverage criterion for relative completeness
  - AI-assisted testing of CPS

- Formal verification
  - Understand the applicability and limitation of formal verification for learning-enabled systems

Combinatorial Testing for Safe ML

Munich Schwabing A9 highway

Source:

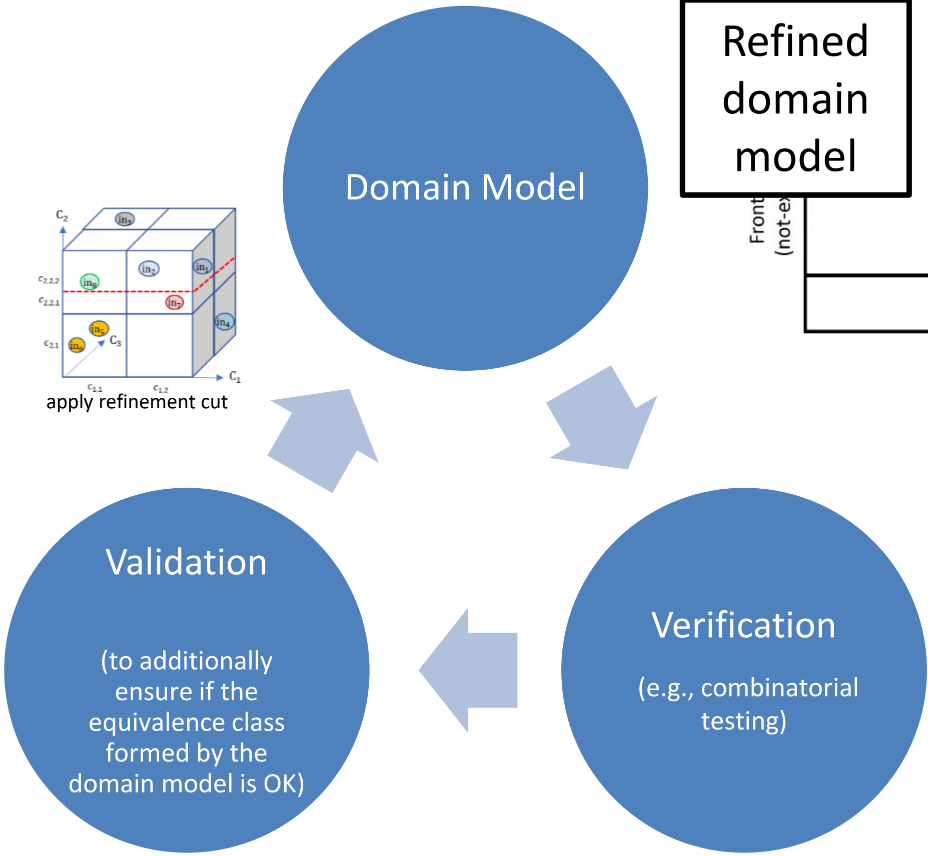
- C.-H. Cheng, C.-H. Huang, G. Nührenberg. *"nn-dependability-kit: Engineering Neural Networks for Safety-Critical Autonomous Driving Systems"*. In: ICCAD'19
- C.-H. Cheng, C.-H. Huang, H. Yasuoka. "Quantitative Projection coverage for testing ML-enabled autonomous systems". In: ATVA'18

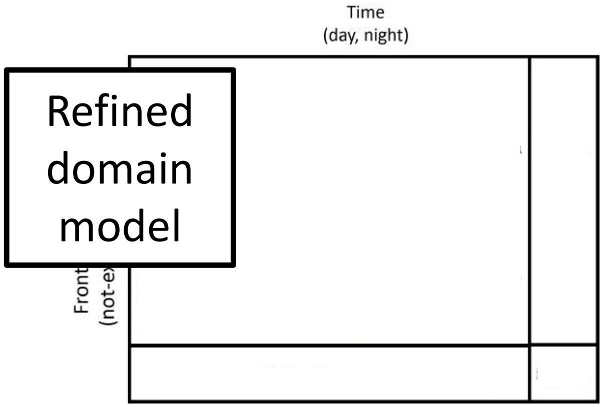
DAYTIME	*morning*	*day*	*evening*	*night*	
HAZE/FOG	*no*		*yes*		
STREET CONDITION	*dry*	*wet*	*icy*	*snow*	*broken*
SKY	*cloudy*	*no*	*clear*		
RAIN	*no*		*yes*		
REFLECTION ON ROAD	*no*		*yes*		
SHADOW ON ROAD	*no*		*yes*		
VRU TYPE	*adult*		*child*		
VRU POSE	*pedestrian*	*jogger*	*cyclist*		
VRU CONTRAST TO BG	*low*		*high*		

Source: C. Gladisch, C. Heinzemann, M. Herrmann, M. Woehrle. "*Leveraging combinatorial testing for safety-critical computer vision*." In: SAIAD'20 (BMWi KI Absicherung)

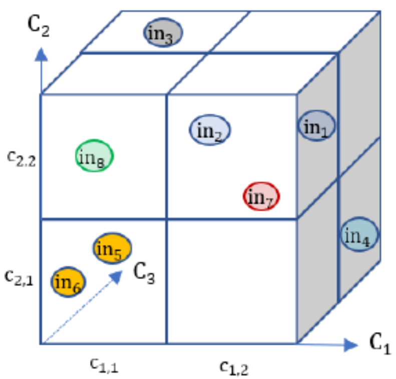
(Concept in https://arxiv.org/abs/2103.04578)

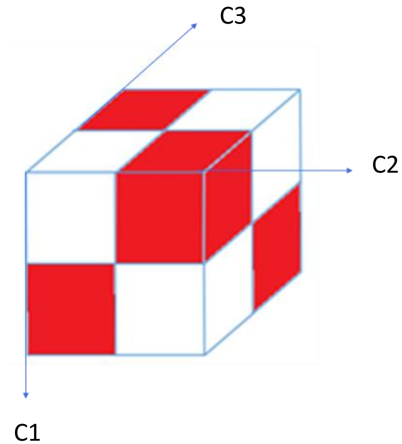
Continuous Verification & Validation





apply refinement cut





Agenda

- Some of the FOCETA addressed challenges
  - Introduce essential dependability and performance requirements for learning-enabled autonomous systems
  - Scalable verification techniques for learning-enabled components
  - Unified approach that combines learning from data and synthesis from specification
  - Monitoring the faithfulness of decision-making at runtime
- Concluding remarks

Combining learning from data and synthesis from specification

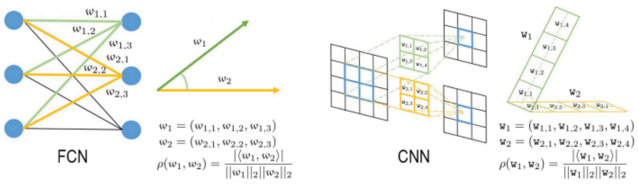
- Deep learning with domain knowledge integrated
  - Potential integration point: layer, loss, data

- Simple output range constraints → enforce the output range in the design of DNN
  - Example: Imitate MPC controller using DNN for Lane-Keeping Assist (*)
  - Output "steering angle" is constrained to be [-60,60] degrees
    - The output value should only be between [-1.04, 1.04] →
    - In architecture design, tanh [-1,1] followed by constant 1.04 scaling

(*) https://www.mathworks.com/help/reinforcement-learning/ug/imitate-mpc-controller-for-lane-keeping-assist.html

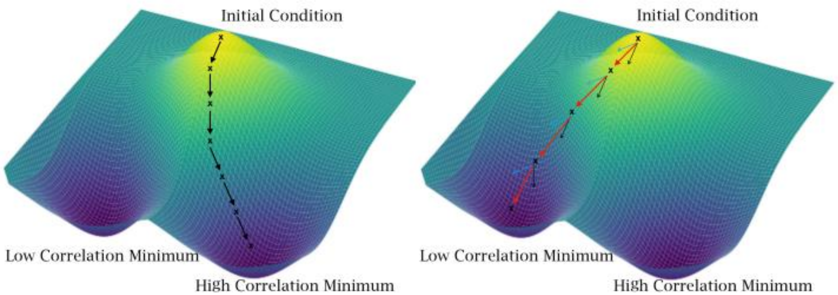
How does Weight Correlation Affect the Generalisation Ability of Deep Neural Networks?

Definition of Weight Correlation



(FCN) The WC of any two neurons is the cosine similarity of the associated weight vectors. (CNN) The WC of any two filters is the cosine similarity of the reshaped filter matrices. This is an open question and there are other correlation metrics for matrices, such as chordal distance and subspace colinearity, we adopt the cosine similarity metric because of its low computational complexity.

Regularisation Based on Weight Correlation



Intuition: find a low correlation minimum

$$\nabla_\theta \tilde{J}(\theta; X, y) = \nabla_\theta J(\theta; X, y) + \alpha \nabla_w \mathbf{g}(w)$$

Standard loss function
hyper-parameter
Regularisation term

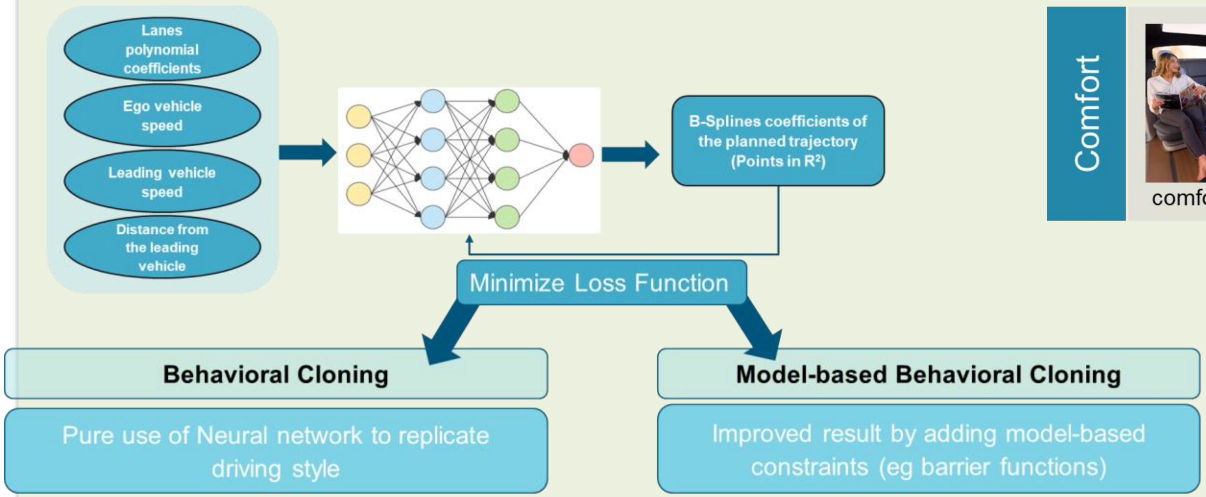
PAC-Bayesian Framework

- Training considering weight-correlation-based regularization improves the generalization ability of CNNs (convolutional) and FCNs (fully-connected).

SIEMENS
Ingenuity for life

Imitation learning with safety-aware loss function

A combined of model-based and AI based control algorithm to learn human-like autonomous driving, that satisfy both safety and comfort objectives



Agenda

- Some of the FOCETA addressed challenges
  - Introduce essential dependability and performance requirements for learning-enabled autonomous systems
  - Scalable verification techniques for learning-enabled components
  - Unified approach that combines learning from data and synthesis from specification
  - Monitoring the faithfulness of decision-making at runtime
- Concluding remarks

Decision faithfulness

- Focus not on using additional modalities for cross-check → monitoring within the single modality is the target
- Methods exists to estimate uncertainty
  - E.g., MC-dropout, entropy, …
  - However, the uncertainty should be **calibrated** in order to be used
    - Calibration requires a **<u>defined process</u>**
    - Recent results in NeurIPS’20 (*) shows that, even under the binary classification setup, the sharpness of calibrated confidence is hard to be guaranteed without prior knowledge of the distribution

(*) C. Gupta, A. Podkopaev, and A. Ramdas, “Distribution-free binary classification: prediction sets, confidence intervals and calibration,” NeurIPS’20.

Agenda

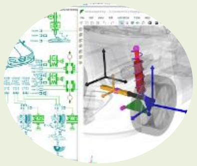
- Some of the FOCETA addressed challenges
  - Introduce essential dependability and performance requirements for learning-enabled autonomous systems
  - Scalable verification techniques for learning-enabled components
  - Unified approach that combines learning from data and synthesis from specification
  - Monitor the faithfulness of decision-making at runtime
- Concluding remarks

FOCETA

- Model-based engineering & data-driven engineering **meet in the middle**, towards something better

- Challenges
  - Specification
  - Verification
  - Learning
  - Monitoring

Use Case Ongoing – Automated Valet Parking



The currently constructed use case utilizes a virtual engineering framework, with the baseline from the Siemens tool chain, e.g.,

- Simcenter Amesim
- Simcenter Prescan

Integrate all partners contribution in design, verification, and validation of safe AI functionalities.

Thank you

FOCETA

Foundations for
Continuous Engineering Trustworthy Autonomy

Chih-Hong Cheng
DENSO AUTOMOTIVE Deutschland GmbH

c.cheng@eu.denso.com

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 956123.