A Context-Aware, Confidence-Disclosing and Fail-Operational Dynamic Risk Assessment Architecture

Patrik Feth, Rasmus Adler, Daniel Schneider

Fraunhofer Institute for Experimental Software Engineering
name.surname@iese.fraunhofer.de

Abstract. Future automotive systems will be highly automated and they will cooperate to optimize important system qualities and performance. Established safety assurance approaches and standards have been designed with manually controlled stand-alone systems in mind and are thus not fit to ensure safety of this next generation of systems. We argue that, given frequent dynamic changes and unknown contexts, systems need to be enabled to dynamically assess and manage their risks. In doing so, systems become resilient from a safety perspective, i.e. they are able to maintain a state of acceptable risk even when facing changes. This work presents a Dynamic Risk Assessment architecture that implements the concepts of context-awareness, confidence-disclosure and fail-operational. In particular, we demonstrate the utilization of these concepts for the calculation of automotive collision risk metrics, which are at the heart of our architecture.

1 Introduction and Related Work

Coping with uncertainties is a major challenge of next generation automated and cooperative systems. Uncertainties are induced by the system complexity (e.g. size, complex behaviors or even utilizing behaviors based on artificial intelligence, genetic algorithms and the likes) and by the complexity, dynamism and unknowns of the environment (e.g. properties of cooperating systems but also behaviors of non-cooperative systems, humans, etc.). It is the goal of a proper engineering process to reduce the amount of such uncertainties during development time. However, the complexity of systems of higher automation levels acting in open environments is too high to remove all uncertainties during development time. For example the behavior of an autonomous system cannot be completely analyzed as it might not follow an explicit behavior specification [1] or for a connected system it is not completely known which information is available at which point in time. Traditionally, dependability engineering addresses those uncertainties with worst-case assumptions. As this does not lead to systems that perform sufficiently well, Laprie required in 2008 to transition from dependability to resilience, defined as *the persistence of dependability when facing changes* [9]. In parallel work to this paper, we introduce Dynamic Risk Management

(DRM) as the safety aspect of resilience. A technical system performing DRM is maintaining autonomously a state of acceptable risk during operation. A crucial function for achieving this is the self-localization of the system within the risk state space. We refer to this activity as Dynamic Risk Assessment (DRA).

In the automotive domain, the safe state is the state in which a collision is sufficiently unlikely to be caused by the considered vehicle. DRM can thus be considered as Collision Avoidance (CA) for the automotive case. State-of-the-art CA systems work with risk metrics as *Time-To-Collision* (TTC) to decide when to become active and which risk reduction strategy to perform [6]. The used risk metrics have evolved over the last years to very mature concepts for Dynamic Risk Assessment. The original version of the TTC metric (distance divided by relative speed) can only be used for the scenario in which two vehicles with the same speed drive behind each other [16]. This is due to the simple constant turn and constant velocity assumption and due to the fact that only the vehicle in front is considered for the calculation. Dijkstra and Drolenga [3] have then generalized the metric to consider more than two vehicles, still using very simple prediction models. How more complex probabilistic prediction models can be integrated in the TTC calculation was presented in [2]. Schreier et al. then performed this step and used a Bayesian, maneuver-based, long-term trajectory prediction for the calculation of TTC, which they call *Time-To-Critical-Collision-Probability* [12]. This risk assessment method was used in the PRORETA 3 system [17].

As impressive as this development is, there is still an important link missing for the usage of risk metrics for a genuine Dynamic Risk Assessment: For the calculation if a state is of acceptable risk it is required to know the probability of an accident and the severity of an accident. The relationship between the value of a risk metric and these two fundamental risk parameters is unclear [13]. The most popular metrics address only the likelihood of an accident by estimating the remaining time till a collision using simple assumptions about the evolution of the current situation. These assumptions do not take into account vehicle-specific properties like the maximum deceleration capabilities, passive safety measures and many other aspects that affect the risk of a collision in a certain driving situation. Consequently, the actual relationship between the value of a collision risk metric and the probability and potential severity of a collision is unclear. However, improving risk metrics is a topic on its own and the existing metrics are at least good enough to build effective CA systems by formalizing the notion of how close the current driving situation is to an accident. We will therefore use existing collision risk metrics for the conduction of DRA in the automotive domain.

Already the very simple version of Time-To-Collision depends on information from the environment (distance and relative speed). Environmental perception is in general a big challenge for vehicles of higher automation levels. If the information from the perception component is used in the context of Advanced Driver Assistance Systems (ADAS), the using ADAS system needs to have great confidence in the integrity of that information. This confidence is obviously also required by our DRA, because a risk has to be detected with a confidence level

that fits to the level of risk. The higher the risk, the higher is the required confidence in detection. To separate the generation of environmental perception information and its usage and thus divide the responsibility, Johansson and Nilsson demand to instantiate each perception service multiple times with different confidence levels [7]. In the architecture that we present in Section 2 we assume that the environment perception subsystem follows this paradigm. This confidence-disclosure is the main novelty of our Dynamic Risk Assessment architecture.

2 Dynamic Risk Assessment Architecture

This section introduces our novel architecture for Dynamic Risk Assessment and highlights how this architecture utilizes the concepts of context-awareness, confidence-disclosure and fail-operational. In order to attain all relevant information (e.g. vehicle speed) in our architecture, we assume a perception component that was realized to address multiple confidence levels simultaneously as described in [7]. Following this concept, most of the input is now considered to be confidence-disclosing by being explicit which confidence level each signal and port provides. For example, there are three levels of confidence for the information of ego speed: ego speed high confidence, ego speed mid confidence and ego speed low confidence. For each of these levels of confidence, there is a dedicated signal and corresponding ports. All such perception input signals are consumed by the actual Dynamic Risk Assessment components, which again reflect different levels of confidence. I.e. we stipulate a *High Confidence DRA*, a *Mid Confidence DRA* and a *Low Confidence DRA* component.

Figure 1 zooms into the *Low Confidence DRA* component which uses the low confidence input signals. Unlike the higher-confidence DRA components, this component contains a *CNN-Based Risk Estimator*. Convolutional Neural Networks (CNNs) have been used widely recently for image classification and scene understanding. In some applications CNNs have been able to perform better than humans on these tasks [8]. However, as CNNs are a lot different from source code in higher programming languages, there is not yet a sufficient body of knowledge on the verification of such Neural Networks. This hinders currently the application of CNNs in the context of applications with high confidence requirements. The confidence-disclosure of our architecture allows the integration of such advanced techniques, into the architecture of safety-critical systems nevertheless. Certainly, the verification of Neural Networks to increase the achievable confidence level requires future work and for now, they should only be used with a disclosure of their low confidence.

The other two sub-components of the *Low Confidence DRA* component are the *TTC (Lane Keeping) Risk Estimator* and the *TTC (Arbitrary Driving Situation) Risk Estimator*. Those two components add the concept of context-awareness to the architecture and are also part of the *Mid Confidence DRA* and *High Con-*

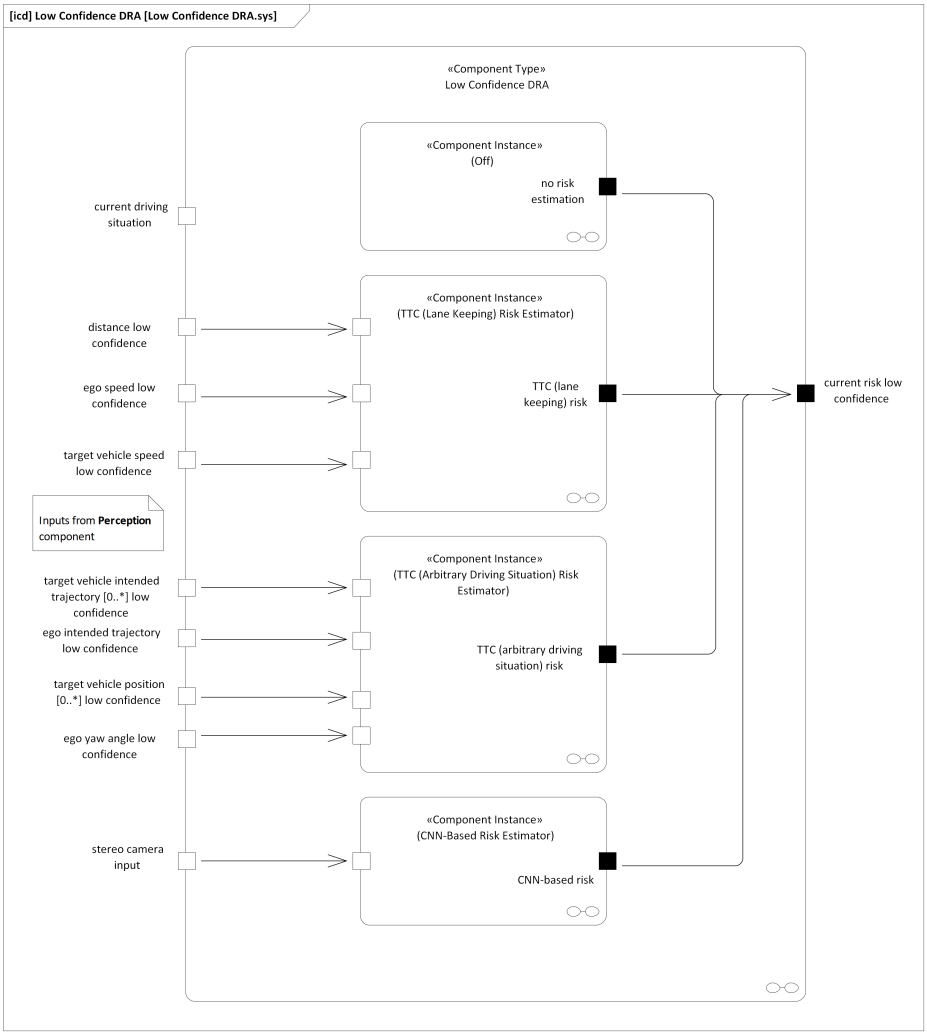

Fig. 1. Low Confidence Dynamic Risk Assessment

fidence DRA components. In the current driving situation of lane keeping, the simple version of the TTC metric described in [16] is sufficient for the dynamic risk assessment. In more general situations a more complex calculation of TTC is required as described in [3]. For robustness, and thus also fail-operational, reasons it makes sense to require, at any point in time, as least information as necessary to make a sufficiently valid risk assessment. As most of the time lane keeping is the current driving situation, it does not make sense to always require the larger set of information required for assessing arbitrary driving situations. For additional fail-operational purposes inside the *TTC (Arbitrary Driving Situation) Risk Estimator* two components for TTC calculation are available: *TTC Intended Trajectory* and *TTC Reachable Area*. The first one requires information about the intended trajectory of all vehicles in the vicinity while the second one only requires information about the current position of other vehicles and then performs a sort of worst-case TTC calculation as described in [14]. This *TTC Reachable Area* calculation obviously leads to an overestimation of the risk but it requires less information and can thus serve as a fallback layer for the more accurate *TTC Intended Trajectory* calculation. In Figure 1 the output of all calculation components are connected to a single output port. In fact, which calculation is used can be formulated by Boolean logic and depends on the current driving situation and a static situation-dependent prioritization.

The outputs of the three *Dynamic Risk Assessment* components (High Confidence, Mid Confidence, Low Confidence) are forwarded to a *Dynamic Risk Control* component. This component contains three different risk reduction strategies in accordance to the ISO 22839 standard for forward vehicle collision mitigation systems [6]: *Driver Warning Risk Reduction*, *Speed Reduction Braking Risk Reduction* and *Mitigation Braking Risk Reduction*. Each of these actions has a different level of inherent risk. This level of inherent risk in turn dictates the requirements with respect to the confidence level of the Dynamic Risk Assessment: *Mitigation Braking Risk Reduction* only works with a high confidence risk assessment, *Speed Reduction Braking Risk Reduction* can work with high and mid confidence risk assessment while *Driver Warning Risk Reduction* can also be employed based on low confidence risk assessment results. At this level, the different confidence level do also add redundancy to the risk assessment and thus allow a fail-operational behavior of the risk control. The *Driver Warning Risk Reduction* has three redundant sources of information it can use for the risk assessment. In case of multiple available sources, a prioritization or aggregation needs to be defined. In case of a missing source, the risk reduction function can compensate it with the remaining sources.

3 Summary and Future Work

In this paper we presented a novel architecture for Dynamic Risk Assessment that utilizes the concepts of context-awareness, confidence-disclosure and fail-operational. The introduced DRA architecture is specific to the automotive do-

main and uses automotive collision risk metrics for assessing the current risk. As input for the DRA we assumed a perception architecture addressing multiple confidence levels simultaneously as introduced in [7]. We transferred that concept of confidence-disclosure also to the calculation of the risk metric giving it multiple outputs for different confidence levels. We demonstrated the possibility to integrate low-confidence but high-performance DRA techniques, e.g. realized by means of Machine Learning techniques, by following the concept of confidence-disclosure. The output of the Dynamic Risk Assessment is used to trigger risk reduction strategies of varying inherent risk. Strategies with a high inherent risk require high confidence DRA results, while for such with a low inherent risk also low confidence DRA results are sufficient. The aspect of context-awareness is represented in the architecture by different ways of performing the risk metric calculation specific to the current driving situation. The DRA architecture enables fail-operational behavior in such a sense that varying availability of information and varying levels of confidence could trigger different configurations for the calculation of metrics.

In future work we plan to implement a demonstrator utilizing this architecture in CARLA, an open-source simulator for autonomous driving research [4]. Further, we plan to add the concepts of a mission-level scope and connectivity to the Dynamic Risk Assessment architecture. Instead of performing the risk assessment for a single vehicle, it will then be performed for an intersection scenario (i.e. traffic light assistant) as a mission that involves multiple vehicles. For such a DRA with mission-level scope, information needs to be shared among the mission participants, thus adding the aspect of cooperation into the mix. For the functionally safe sharing of corresponding information we will instantiate in a first step the Conditional Safety Certificates (ConSerts) concept [10] for the use case of mission-level Dynamic Risk Assessment. With ConSerts, a provided signal may have several guarantees with different levels of confidence, which in turn depend on the dynamic fulfillment of assumptions regarding the environment (e.g. other cooperating systems). In this way, ConSerts are a means to enhance the confidence consideration that we described in the architecture above. The level of confidence will still be represented by single values so that the sender and the receiver need to have a standardized agreed basis about what creates how much confidence. In order to avoid misunderstandings in this respect and to allow greater flexibility, we are currently working on a concept called Digital Dependability Identities, which augments ConSerts and the way levels of confidence are specified in the sense of a dynamic assurance case [11].

Acknowledgments. This project has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 783119. The JU receives support from the European Unions Horizon 2020 research and innovation programme and Netherlands, Austria, Belgium, Czech Republic, Germany, Spain, Finland, France, Hungary, Italy, Poland, Portugal, Romania, Sweden, United Kingdom, Tunisia.

References

1. Adler, R., Feth, P.: Safety assurance for highly and fully automated driving. 26th Aachen Colloquium Automobile and Engine Technology (2017)
2. Berthelot, A., Tamke, A., Breuel, G.: A novel approach for the probabilistic computation of time-to-collision. IEEE Intelligent Vehicles Symposium (2012)
3. Dijkstra, A., Drolenga, H.: Safety effects of route choice in a road network: simulation of changing route choice
4. Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., Koltun, V.: CARLA: An open urban driving simulator. In: Proceedings of the 1st Annual Conference on Robot Learning. pp. 1–16 (2017)
5. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. Human Factors: The Journal of the Human Factors and Ergonomics Society 37(1) (1995)
6. ISO: Intelligent transport systems - forward vehicle collision mitigation systems - operation, performance, and verification requirements (2013)
7. Johansson, R., Nilsson, J.: The need for an environment perception block to address all asil levels simultaneously. IEEE Intelligent Vehicles Symposium (2016)
8. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems (NIPS) (2012)
9. Laprie, J.C.: From dependability to resilience. IEEE/IFIP International Conference on Dependable Systems and Networks (2008)
10. Schneider, D., Trapp, M.: Conditional safety certification of open adaptive systems. ACM Transactions on Autonomous and Adaptive Systems 8(2) (2013)
11. Schneider, D., Trapp, M., Papadopoulos, Y., Armengaud, E., Zeller, M., Hofig, K.: Wap: Digital dependability identities. IEEE International Symposium on Software Reliability Engineering (2015)
12. Schreier, M., Willert, V., Adamy, J.: Bayesian, maneuver-based, long-term trajectory prediction and criticality assessment for driver assistance systems. IEEE Intelligent Vehicles Symposium (2014)
13. Wachenfeld, W., Winner, H.: Do autonomous vehicles learn? In: Maurer, M., Gerdes, C.J., Lenz, B., Winner, H. (eds.) Autonomous Driving. Springer Open (2015)
14. Wachenfeld, W., Junietz, P., Wenzel, R., Winner, H.: The worst-time-to-collision metric for situation identification. IEEE Intelligent Vehicles Symposium (2016)
15. Wardzinski, A.: Dynamic risk assessment in autonomous vehicles motion planning. International Conference on Information Technology (2008)
16. Winner, H., Geyer, S., Sefati, M.: Maße für den sicherheitsgewinn von fahrerassistenzsystemen. Darmstädter Kolloquium Mensch + Fahrzeug (2013)
17. Winner, H., Lotz, F., Bauer, E., Konigorski, U., Schreier, M., Adamy, J., Pfromm, M., Bruder, R., Lueke, S., Cieler, S.: Proreta 3: comprehensive driver assistance by safety corridor and cooperative automation. In: Winner, H., Hakuli, S., Lotz, F., Singer, C. (eds.) Handbook of Driver Assistance Systems. Springer International Publishing (2016)