Virtual Assessment of Automation in Field Operation
A New Runtime Validation Method

Walther Wachenfeld* and Hermann Winner†

Summary: Highly automated vehicles being a new technology in public traffic have to fulfill the demanding safety requirements resulting from human driving. To assess automated systems in means of safety a new runtime validation method the “Virtual Assessment of Automation in Field Operation” is introduced. Potential benefits like reliable test case generation, minimal additional risk and enlarged test case coverage are motivated.

Keywords: automated vehicles, virtual assessment, field operation, test method.

1 Introduction

A highly or fully automated vehicle (SAE J3016 Taxonomy [1]) will be a new technology when it enters public traffic. Nevertheless the safety of this new technology will be compared to that of human drivers assisted by advanced driver assistant systems (ADAS [2]). This combination of high safety requirements and new, higher levels of complexity leads to the so called “approval trap” that could befall the introduction of automated vehicles. The “approval trap” describes the situation in which there is no known way to prove that an automated system could match a human driver in terms of safety, even if such a system should exist. It is claimed that the state of the art in vehicle testing needs to be enhanced to bring these vehicles to everyday use [3].

Regarding these challenges this paper motivates a new runtime validation method called Virtual Assessment of Automation in Field Operation (VAAFO). Based on that, the VAAFO concept is discussed with the aid of a developed architecture and use cases that illustrate the principles. The benefits and drawbacks are highlighted.

2 Motivation

The required enhancement of test methods could be addressed by improving real world testing or virtual testing independently. However, structural drawbacks of both test methods will still exist:

- Real world tests reliably assess the object under test (OUT) because real situation

* Walther Wachenfeld, research associate at the Institute of Automotive Engineering at TU Darmstadt, Otto-Berndt-Straße 2, 64287 Darmstadt (wachenfeld@fzd.tu-darmstadt.de).

† Hermann Winner, head of the Institute of Automotive Engineering at TU Darmstadt, Otto-Berndt-Straße 2, 64287 Darmstadt (winner@fzd.tu-darmstadt.de).

setups and real systems are tested. But due to the costs of prototype vehicles and test drivers, real world tests can only cover a limited number of kilometers. In addition, real world tests are always connected to additional risks introduced by prototype systems that need to be controlled by test drivers.

- Virtual tests can cover an increased number of kilometers due to acceleration and parallelization of software simulations. Additionally, virtual tests are executed without taking any additional risk. However, these assessments are questioned due to the limited validity of software models and artificial test case generation.

To reduce the drawbacks of both systems, approaches exist that try to combine these two test methods. On the one side, virtual obstacles are injected into the field of view by augmented reality for human drivers [4] or as additional objects in the object list for automated driving [5]. Both approaches are used on test fields and therefore reduce additional risks and increase validity by replacing ego-vehicle software models. However, both approaches still suffer the problems faced by artificial test case generation.

This issue could be addressed by scenario or test case extraction during real world driving. Publications [6] and [7] discuss a (semi-)automated scenario generation for Software-in-the-Loop tests based on recorded real world driving. However, these publications do not address how to proceed with these generated scenarios and how to generate the necessary amount of *relevant* scenarios. According to statistics, the average necessary number of kilometers to experience at least one accident with casualties on a German highway is 210 million [8]. When trying to evaluate safety just by randomly collecting situations in public traffic, even billions of kilometers are necessary to cover a relevant part of the existing situations and to be able to scientifically prove safety [3]. Collecting this amount of data with only a few vehicles will not be meaningful.

The VAAFO approach described in the following is motivated by the described challenges above and will further be encouraged by:

- The handling of abnormal situations of other safety relevant transportation technologies. For example, in avionics, situations that indicate an unusual behavior of technical systems are reported also after start of production (SoP) during field operation [9].
- The requirements of new complex systems and the derived set of tests will be incomplete, especially during the first introduction phase.
- Automated driving will probably enter public traffic in an evolutionary manner. The use cases will expand and the level of automation will increase [10][1]. In the areas where the automation is not capable of driving safely the human together with ADAS still has to conduct the vehicle.

3 VAAFO Concept

The basic idea of the VAAFO concept is derived from the so-called Trojan Horse approach [11][12]. This approach addresses the testing of emergency intervening systems like emergency brake assist (EBA), which try to mitigate accidents. For this EBA, the results of assessment by means of false positive and false negative rates are clear. When assessing systems that control vehicle dynamics constantly, this unambiguity isn’t granted anymore. For that reason, the Trojan Horse has to be developed further, resulting in the VAAFO concept. Similar, but less concrete ideas are written down in a patent from Hoye et al. [13] and at a press interview from an employee at Bosch [14]. Both mention ideas without giving further insight into their development. For the VAAFO concept the next sections will give further insight answering three questions:

- Which additional components are necessary for the VAAFO concept and how do they interact with the human controlled vehicle? → Section 3.1 VAAFO Concept Architecture
- How does the VAAFO concept assess the automation and why is it a new runtime validation method? → 3.2 Section Assessment of Automation
- Can the new validation method address all relevant cases in means of safety? → 3.3 Coverage of the concept

3.1 VAAFO Concept Architecture

Which additional components are necessary for the VAAFO concept and how do they interact with the human controlled vehicle?

The VAAFO concept pursues two major goals:

- The assessment of the automation in terms of safety.
- The identification of test cases relevant for the safety evaluation of the automation (OUT).

These goals seem achievable when combining Virtual Assessment of Automation with Field Operation (VAAFO) in the way depicted in Figure **1**. A human (light blue/first row) drives the vehicle. Therefore he perceives the real dynamic world, processes the information and executes by steering, accelerating and braking the vehicle. Additionally, the driver could be assisted by an ADAS or partial automation. This is the *field operation* of regular today's driving.

In addition, Figure **1** shows more components that are implemented in the vehicle. First of all the automation (light green/second row), that perceives the real dynamic world and processes the information, *but does not act on the real actuators.* The automation cannot change the vehicle's real behavior by steering, accelerating, or braking. This missing link to the real actuators leads to an open loop control. To evaluate the closed loop behavior of automation the VAAFO tool is added (yellow/third row). The tool initializes a virtual dynamic world based on the world model of the automation. In this virtual dynamic world, the automation (Obejct Under Test - OUT) changes the behavior of the vehicle. Consequently, this makes the automation assessable for a time of some seconds.

The situation assessment is based on the two world models that reflect the driver's behavior and/or the behavior of automation. Based on the retrospective situation assessment the virtual behavior of the OUT is assessed and relevant cases are identified and logged. This new way of assessing the automation will be described in the following part using an example.

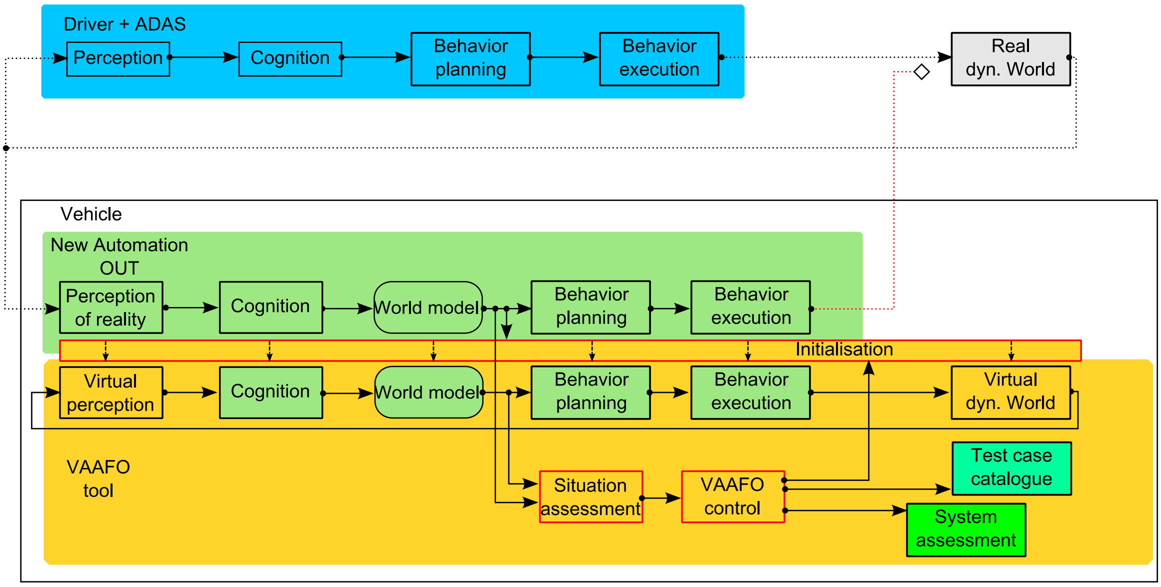

Figure 1: VAAFO concept architecture.

3.2 Assessment of Automation

How does the VAAFO concept assess the automation and why is it a new runtime validation method?

When automatically assessing vehicle automation, the assessment needs additional information the vehicle automation doesn't have during behavior planning. For example, a test driver who nowadays assesses a vehicle automation uses his own perception and cognition to compare his behavior planning with the execution of the automation to intervene if necessary. The test driver has additional information available due to his more advanced perception and cognition of the world (at least in most cases at present).

As the VAAFO concept is neither controlling the vehicle nor being a safety system, it accesses two information sources based on the same sensor setup that are not accessible for the vehicle automation. The one source is the real trajectory (defined by the human driver), accessed by interpreting the inertial sensors like speeds, accelerations, and localization techniques like differential GPS (Global Positioning System) or SLAM (Simultaneous Localization and Mapping). The real trajectory and the trajectory of the automation lead to two parallel worlds (section 3.2.1). The second source of information gives the retrospective view on the sensor data (section 3.2.2). The assessment doesn't need to be predictive or time synchronous. For that reason, the world model that is used for assessment can be enriched by information gained from a longer time span. Both sources of additional information will be explained now in more detail:

3.2.1 The Parallel World

One virtual world representation is built up based on the real sensor perception. In this virtual world, two trajectories or vehicle behaviors can be compared. One is the real trajectory the other the trajectory of the automation. Figure 2 illustrates this with an example:

In reality (row #1) a human-driven vehicle drives in the right lane. It approaches an obstacle and goes around it by moving one lane to the left.

The perceived worlds (rows #2 and #3) look similar to the first, but with the difference that the obstacle is not perceived before time step 2 s. A reason for this kind of false negative detection that is corrected over time could be the different characteristics of the mounted sensors. For example, the long range sensors like radar don't detect the bush but the sideways mounted short range sensors like radar, ultrasonic, or 360° camera do.

In this perceived world, one trajectory that is measured as well is the human driven one (row #2). The vehicle decelerates a bit at t_1 = 1 s and goes around the obstacle at t_2 = 2 s like in reality.

Based on the perceived world, a parallel world is started where the automation drives the vehicle (row #3). As the automation is not aware of the obstacle, it doesn't decelerate the vehicle and goes straight. In this example, the obstacle appears after the vehicle has passed it.

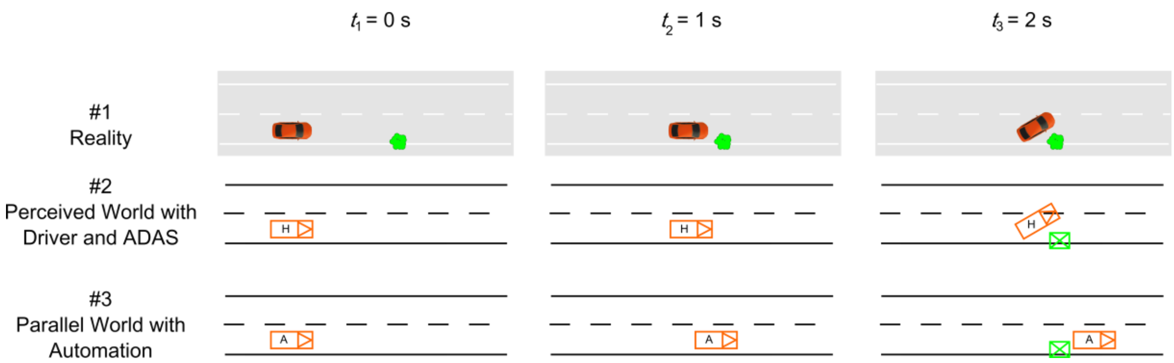

Figure 2: Comparison of parallel worlds with human and automation as vehicle drivers

Although the automation didn't collide in the parallel world, the different trajectories give a first indicator (trigger) that the behavior of the automation was possibly unsuitable. This uncertainty motivates further evaluation of this situation. This leads to the retrospective approach for assessment.

3.2.2 *Retrospective Approach*

The different behavior shown in the example above as well as the vague statement about the trajectory of the automation result from the uncertainty of perception. Uncertainty is one difference between the ground truth and the vehicle's understanding of the real world. Dietmayer [15] distinguishes three uncertainties:

- uncertainty about state (Zustandsunsicherheit)
- uncertainty about existence (Existenzunsicherheit)
- uncertainty about classification (Klassenunsicherheit)

When trying to use the same sensors for assessing, one can never get rid of these uncertainties. This is valid also for the VAAFO concept. But, as the VAAFO concept is not controlling the vehicle behavior, a new perspective can be taken onto the data of one situation. In particular, perceiving the environment gets more accurate when having more time for sensing, getting closer to the object (without falling below the lowest range), and getting access to more perspectives of one object. This motivates the retrospective post-processing of the world model. Information that is collected over a certain time span is summarized in a new world model called the retrospective world (row #4). This enriched world is used for assessing the trajectory of the automated vehicle (in general, the human can be assessed as well).

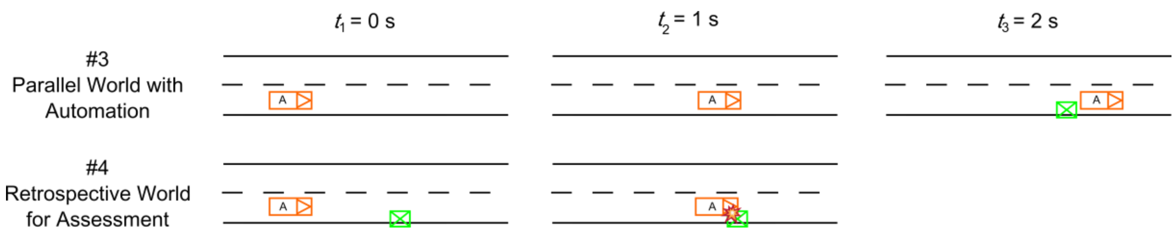

Figure 3: Comparison between world model for control and world model for assessment

Figure 3 shows this for the example stated above. The collision-free world (row #3) is post-processed and the obstacle detected in time step 2 s is already placed in the world model (row #4) from the beginning, as the static object doesn't change during this time span. When repeating the trajectory of the automated vehicle, the vehicle collides with the obstacle.

A second indicator that the automated vehicle behavior is not adequate is generated. Due to the uncertainties mentioned above, the clarity of this case is not always given, as discussed in the following.

3.2.3 *Credibility of Assessment*

Above, an ideal situation for using the VAAFO concept was motivated. A difference in trajectories as well as the detection for the reason of this difference is given. Other cases exist where this clarity isn't given. Let us assume that the obstacle isn't detected in time step 2 s (Figure 4). In this case the retrospective evaluation does not provide an indicator for any wrong behavior of the automation (row #4). Only the comparison between the real trajectory and that from automation indicates differences.

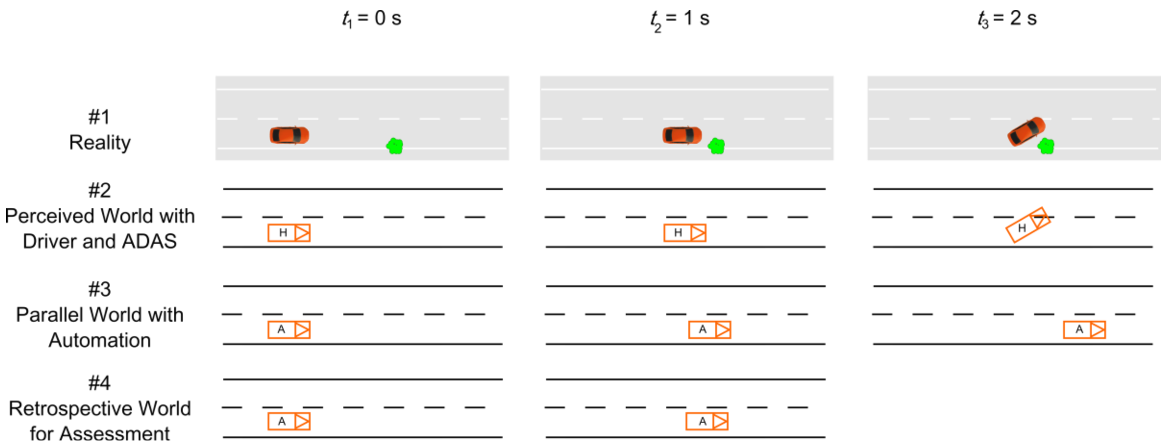

Figure 4: VAAFO case with uncertain assessment

A similar challenge occurs when the trajectory is the same but the retrospective assessment identifies a collision. This could result from two causes: either the perception suffers false positive detections or an accident isn't reported (or isn't severe).

One could argue that these cases challenge the credibility of the assessment of the VAAFO concept. However, these situations are identified as difficult to assess even with additional information. Consequently these situations are relevant to record and evaluate later because the automation is not able to understand what happened.

3.3 Coverage of the concept

Can the new validation method address all relevant cases in means of safety?

In examples it was shown how the concept works and which challenges exist. The level of coverage that is reachable is relevant for a test concept. The test coverage of the VAAFO concept is limited by two factors.
First, the coverage is limited by the situations that are accessible. The VAAFO concept will never uncover challenges in handover situations or in situations that originate in the misunderstandings of other road users when not being able to communicate with a human driver. To assess the safety for these situations, another validation method (for example the Wizard of Oz experiment) needs to be performed.
Second, the coverage is limited by the sensitivity of the comparison between human and automation trajectory. Similar to the ROC (Receiver Operating Characteristic) for object detection, the VAAFO concept will suffer the compromise between false positive detection and missing (false negative) detection of the faulty behavior which should be detected (true events). The right sensitivity for trajectory comparison needs to be derived in the future.

3.4 VAAFO Concept Benefits and Drawbacks

The potential benefits of the VAAFO concept become obvious when looking back to the motivation. The coincidental/random nature of the real world serves as a reliable test case generator. When covering a huge amount of kilometers, the safety evaluation is valid. As the VAAFO concept is not changing the vehicle behavior in real world, no additional risk to public traffic is added, which means that every driver is able to collect these kilometers.

When trying to collect these huge numbers of kilometers using some small amount of vehicles isn't enough. But, when placed on 1000 vehicles in German public traffic, it would be possible to obtain more than 10.000.000 kilometers' worth of data each year [16]. This would require vehicles equipped with both sensors and processing power. These vehicles that would be over equipped compared to the vehicles with a customers noticeable benefit would lead to higher costs for the carmaker and contradict the conventional cost-minimizing approach of the car industry. However, since the VDA and other roadmaps to highly automated vehicles foresee an evolutionary approach, a large number of these vehicles must be seen on German roads before highly or fully automated vehicles be released for production [17]. Besides an evolutionary approach from one vehicle generation to the other, a different introduction strategy of automated driving supports the VAAFO concept even more strongly. When introducing vehicles with a deactivated automated driving function, an assessment of the functions can be done with the VAAFO concept before activation. In this case, the necessary hardware is already installed and was financed by the outlook to updated and activated functionality. A step by step increase of the range of application in terms of, for example, speed, weather conditions, and maneuvers is imaginable. In fact, the over air update is already in place at some vehicle manufacturers (homepage article [18]).

Even with all these benefits in mind, the concept still faces the challenge of valid simulation models. Simulation models that replace vehicle dynamics do exist, as along with inertial sensors and actuators like those seen in [19]. But how do the humans in the automated vehicle's surroundings behave, or how does a valid environment sensor model look? Both questions always challenge virtual results. The VAAFO concept, however, comes with the following two additional advantages addressing these challenges. First, the simulations will only be as short as a few seconds. Similar to those used for accident reconstruction, the simulations for automated vehicles are more accurate the fewer simulation steps away from initialization are performed. The fidelity of the behavior of other road participants will be a decision-making argument for the length of the virtual simulation. The second benefit comes with the fact that sensor characteristics are by concept introduced into the world model as real sensors are used to initiate and update the virtual world. For example, the timing of a formerly covered obstacle becoming visible for the perception sensors comes from the real used sensor.

A bunch of new research topics follow from the VAAFO concept, as it introduces additional "intelligence" into testing and widens the focus of testing beyond the start of production. The more complex the system, the more the question arises whether such methods are necessary to support classical safety assessment. This counts not just for automotive but also for the general automation of human tasks as it occurs in, for example, avionics or automated surgeries.

Literature

[1] SAE International Standard J3016: Taxonomy and Definitions for Terms related to On-Road Motor Vehicle Automated Driving Systems, SAE International (2014)

[2] Winner, H., Hakuli, S., Lotz, F., Singer, C. (eds.): Handbuch Fahrerassistenzsysteme, 3rd edn. Vieweg-Teubner-Verlag (2015)

[3] Wachenfeld, W., Winner, H.: Die Freigabe des autonomen Fahrens. In: Maurer, M., Gerdes, C., Lenz, B., Winner, H. (Hrsg.) Autonomes Fahren. Technische, rechtliche und gesellschaftliche Aspekte. Springer Vieweg (2015)

[4] Bock, T.: Bewertung von Fahrerassistenzsystemen mittels der Vehicle in the Loop-Simulation. In: Winner, H., Hakuli, S., Wolf, G. (Hrsg.) Handbuch Fahrerassistenzsysteme, pp. 76-83. Vieweg+Teubner Verlag (2012)

[5] Sefati, M., Stoff, A., Winner, H.: Testing Method for Autonomous Safety Functions Based on Combined Steering/Braking Maneuvers for Collision Avoidance and Mitigation. 6. Tagung Fahrerassistenz, München, (2013)

[6] Lages, U., Spencer, M., Katz, R. (eds.): Automatic scenario generation based on laserscanner reference data and advanced offline processing. Intelligent Vehicles Symposium (IV), 2013 IEEE (2013)

[7] Folie, M., Krug, M., Pfeffer, R., Greger, M.: Chain of effects for testing camera-based ADAS along the development process. Session 2, IPG Open House (2015)

[8] Statistisches Bundesamt (Destatis): Verkehrsunfälle - Fachserie 8 Reihe 7 (2012)

[9] Weitzel, A., Winner, H., Peng, C., Geyer, S., Lotz, F., Sefati, M.: Absicherungsstrategien für Fahrerassistenzsysteme mit Umfeldwahrnehmung. Berichte der Bundesanstalt für Strassenwesen - Fahrzeugtechnik (F), vol. 98. Wirtschaftsverl. NW Verl. für neue Wissenschaft, Bremerhaven (2014)

[10]Winner, H.: Quo vadis, FAS? In: Winner, H., Hakuli, S., Lotz, F., Singer, C. (Hrsg.) Handbuch Fahrerassistenzsysteme, 3rd edn. Vieweg-Teubner-Verlag (2015)

[11]Winner, H.: Einrichtung zum Bereitstellen von Signalen in einem Kraftfahrzeug. Patent DE 101(02), 771 (2001)

[12]Reschka, A., Rieken, J., Maurer, M.: Entwicklungsprozess von Kollisionsschutzsystemen für Frontkollisionen: Systeme zur Warnung, zur Unfallschwereminderung und zur Verhinderung. In: Winner, H., Hakuli, S., Lotz, F., Singer, C. (Hrsg.) Handbuch Fahrerassistenzsysteme, 3rd edn., pp. 913–935. Vieweg-Teubner-Verlag (2015)

[13]Hoye, B., Lambert, D., Sutton, G.: AUTONOMOUS DRIVING COMPARISON AND EVALUATION. US Patent 20,150,175,168 (2015)

[14]http://www.heise.de/newsticker/meldung/Autonome-Autos-Danke-dass-Sie-das-Auto-von-morgen-testen-2760591.html accessed 23/07/2015

[15]Dietmayer, K.: Prädiktion von maschineller Wahrnehmungsleistung beim automatisierten Fahren. In: Autonomes Fahren, pp. 419–438. Springer (2015)

[16]Brachat, H.: DAT report 2013. Published in Autohaus Extra 5 (2013)

[17]Bartels Arne, Ruchatz Thomas: Einführungsstrategie des Automatischen Fahrens. at - Automatisierungstechnik 63 (2015). doi: 10.1515/auto-2014-1145
[18]http://www.teslamotors.com/blog/model-s-has-you-covered accessed 31/07/2015
[19]Baake, U., Wüst, K., Maurer, M., Lutz, A.: Testing and simulation-based validation of ESP systems for vans. ATZ Worldw 116(2), 30-35 (2014). doi: 10.1007/s38311-014-0021-6