Response to DfT, BEIS and Centre for Connected and Autonomous Vehicles

Open consultation

Automated vehicle trialling code of practice: invitation to comment

On behalf of the UK Computing Research Committee, UKCRC.

Prepared by: Professor Chris Johnson,
School of Computing Science, University of Glasgow, Glasgow, G12 8RZ.
http://www.dcs.gla.ac.uk/~johnson

The UK CRC is an Expert Panel of all three UK Professional Bodies in Computing: the British Computer Society (BCS), the Institution of Engineering and Technology (IET), and the Council of Professors and Heads of Computing (CPHC). It was formed in November 2000 as a policy committee for computing research in the UK. Members of UKCRC are leading researchers who each have an established international reputation in computing. Our response thus covers UK research in computing, which is internationally strong and vigorous, and a major national asset. This response has been prepared after a widespread consultation amongst the membership of UKCRC and, as such, is an independent response on behalf of UKCRC and does not necessarily reflect the official opinion or position of the BCS or the IET.

Response

[1] (Addressing Section 1) – the revised Code of Practice provides a sound basis for work in this area. However, we note that no special provisions are made for the increasing levels of assurance that are appropriate with higher levels of automation; following the generally accepted distinctions introduced by the US National Highway and Transport Safety Agency [1]. A proportionate, risk-based approach should encourage trialists to recognise the need for greater assurance in the safety cases associated with vehicles exploiting higher levels of automation.

[2] (Section 1) – the revised Code of Practice has a particular view of automation that focuses on single operator, single vehicle control. It does not cover research/development of Automated Highways or Platooning of multiple vehicles. If these are assumed to meet the same criteria as single vehicles then it might be useful to add a clause that explicitly mentions this, see for example the EC SARTRE project [2] or the US truck demonstrators associated with Auburn University [3]. These have the potential to reduce costs, improve capacity and lower emissions hence should not be discounted from the revision.

[3] (Section 2.7) defines safety as *the absence of unreasonable risk*. The Health and Safety at Work Act (1974) is more specific, requiring that risks to health and safety should be reduced *so far as is reasonably practicable* (SFAIRP). In particular, this required trialists to conduct an analysis to demonstrate that the cost of reducing the risks further would be *grossly disproportionate* to the benefit achieved. UK industry has considerable experience of working under the HSAW Act and understands the broad concept of risk reduction SFAIRP. The Code

[1] https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety

[2] https://www.webcitation.org/5vpeyHyE3?url=http://www.sartre-project.eu/en/about/Sidor/default.aspx

[3] http://www.eng.auburn.edu/news/2018/09/truck-platooning-study.html

of Practice does not provide sufficient guidance on how to interpret “unreasonable risk” in terms of the consequences and likelihood of any residual hazards.

[4] (Section 2.7) – the second bullet requires ‘Evidence that the trial activity can be performed safely, whether with a safety driver in the vehicle or with a remote safety operator’. It is unclear what evidence can be used to demonstrate ‘the absence of unreasonable risk’ in this context. Exhaustive testing can only provide partial evidence for complex software-based systems because we cannot examine the many billions of potential states that the system might enter. Test coverage will only be able to consider a tiny fraction of the possible responses to complex environmental conditions. Hence, it is very likely that such tests will not consider the many possible states that could lead to accidents. Other forms of static analysis can help to address these limitations and more technical guidance should be provided to trialists on acceptable means of compliance with such high-level requirements.

[5] (Section 2.7) –existing trials mitigate risk by placing constraints on the operating environment of autonomous and connected vehicles. For example, certain demonstrators may not be operated at night. Others are restricted to roads that meet lane markings conventions etc. Any safety case must clearly identify those environmental constraints and they must be recognised by the operator of an autonomous vehicle.

[6] (Section 2.7) – safety cases typically record the evidence that supports particular assurance arguments. A primary aim of trials is to gather further evidence through experience ‘in use’. As described in paragraph 4, this cannot hope to cover all possible internal states of the vehicle software or environmental conditions. Equally, it is essential that the safety case be updated and maintained as new data becomes available from the road trials. For instance, if experience suggests new situations in which external communications links are interrupted then any risk assessment in the safety case should also be updated with the safety information gained through public trials.

[7] (Section 2.10) – One means of assessing whether or not the draft Code of Practice is ‘fit for purpose’ is to measure its provisions against recent cases in other jurisdictions. For example, the NHTSA recently intervened to halt a trial in which children were being transported to School in an autonomous vehicle [4]. The proposed clause might be strengthened beyond the use of the word ‘encouraged’ to engage with DfT and CCAV in such circumstances.

[8] (Section 2.11) - requires *plans for police investigators and relevant organisations to readily and immediately access data relating to an incident in a way that maintains the forensic integrity, security, and the preservation of the data* but police and emergency services are unlikely to have the resources to determine whether or not such plans are adequate, practical or complete. The police will typically require external support to determine whether or not a trialist has met their obligations in terms of the forensic logging of software related events. In particular, any such provision must enable analysts to identify the complex interactions between human and autonomous control conditioned by input from heterogeneous sensors over a range of network technologies.

[9] (Section 2.12) – experience in the United States has identified the need to retain forensic data, including that derived from near-miss incidents. If a more serious failure should occur then it should be possible to identify similar situations that might provide additional insights, for instance about the potential precursors to an accident. The revised Code should include provisions for the retention of safety-relevant data.

[10] (Section 2.12) – these requirements raise serious questions about the feasibility of any independent investigation given the complexity and volume of the data needed to determine the causes of any software malfunction. There may be petabytes of data and interpreting that data may require detailed knowledge of the system architecture and the system states where the data is generated. The vehicle manufacturer might have an insurmountable conflict of interest in providing that interpretation to a court and to an independent expert, and the time an cost of analysing it would be very high, as the expert evidence in *Bookout v Toyota* showed [5].

[11] Section 2.13 - places necessary but very extensive assurance requirements on every new version of the software. This will limit the frequency with which software can be patched or updated, which will deter manufacturers from implementing timely corrections for safety or security vulnerabilities that are discovered, because every software

[4] https://www.nhtsa.gov/press-releases/nhtsa-directs-driverless-shuttle-stop-transporting-school-children-florida
[5] https://en.wikipedia.org/wiki/2009–11_Toyota_vehicle_recalls

change must create a new version of the software. The Code should specify under what conditions it will be lawful for the vehicles to continue to be used if a vulnerability has not yet been corrected.

[12] Section 2.13 – as mentioned in paragraph [4], it is important to consider what would represent acceptable means of compliance within the context of this clause. For example, many autonomous vehicles have incorporated Commercial Off The Shelf (COTS) and third-party components for which they cannot access the source code. In such cases, they will not know the particular changes that are embedded within a particular patch or revision and so have no idea what aspects will require further bench testing prior to release. How should such situations be handled, given that many of these components were not intended for safety-related applications in autonomous vehicles?

[13] Section 2.14 – in practice, autonomous vehicles would seem to pose less of a threat to data protection than they might to public safety, especially in terms of the potential cyber risk. Sections 2.14-2.16 seem at odds to the earlier focus on safety and certainly could be swapped with 2.17 and 2.18 on Cyber Security.

[14] Section 2.17 – this passage lacks the level of detail necessary to protect future generation of vehicles, or even the present demonstrators, from the potential threats to cyber security. As mentioned in paragraph [12], many components will be repurposed from Commercial Off-The-Shelf products. The Tier 1 manufacturer is unlikely to know what software in included, let alone to have access to detailed assurance data for it. Current international and national safety standards do not adequately address the implications of security vulnerabilities for safety, so even full conformance with these standards would provide little evidence that risks had been reduced SFAIRP.

[15] Section 2.18 – one particular area of concern is the need to consider cyber security threats within the overall safety case. If this is not done then some other means is needed to document the measures taken to mitigate the putative risks. It is important that any security extensions to a safety case are updated – for example as new threats are demonstrated or, for instance, if a COTS supplier releases a patch that is not yet installed on an autonomous vehicle.

[16] Section 3.7 – there should also be a point of contact for any member of the public wishing to report an incident that they have witnessed or to determine how any previous notification of such an incident has been handled – this requirement stems from public concerns over some of the US demonstrators.

[17] Section 3.9 – the Code of Practice does not provide any mechanisms or processes by which a member of the public may lodge legitimate concerns about the conduct of any trial. This caveat should be seen in the context of paragraph [16].

[18] Section 3.10 - should require that detailed safety cases are available for independent scrutiny. An abridged version is likely to be a superficial and bland marketing document that lacks the evidence necessary to justify confidence. DfT might consider the development of a generic safety case or template to be instantiated by future trials. This approach has been used in aviation [6], reducing the costs for companies and improving consistency across the market.

[19] Section 3.10 – we welcome the requirement that safety cases should be regularly updated but suggest the provision should be strengthened in line with our comments in paragraph 6, above. At a minimum a safety case should be reviewed every 3-6 months and in the light of evidence from operational and bench testing. This requirement is particularly important to reflect any changing threat analysis for cyber security requirements.

[20] Section 3.11 - should require that **all** trial data captured by the company (along with any technical analyses carried out is maintained for a lengthy period so that it can be made available to public authorities when required, reiterating comments made initially in paragraph [9].

[21] Section 4.1 – This clause requires that drivers and operators “have a full view of the road ahead”. This might be strengthened to include an awareness of surrounding road users.

[6] https://www.eurocontrol.int/sites/default/files/field_tabs/content/documents/nm/safety/appendix-b-3-outline-safety-case-for-stca-system.pdf

[22] Section 4.2 – in situations where there are mixed modes of control, for example a remote driver and an on-board driver, it should always be possible to identify which has immediate control of the vehicle, with clear guidance on protocols for handing off and assuming control.

[23] Section 4.12 –we recommend a slight strengthening of this requirement to include a period of familiarisation on closed roads before any driver assumes control for the first time **or** transfers between different types of autonomous vehicle on a public road.

[24] Section 4.12 – the familiarisation recommended in paragraph [23] should cover a range of environmental conditions likely to be encountered on public roads.

[25] Section 4.14 – any subsequent incidents should be logged and the training envisaged in the revised Code of Practice should be updated both to alert all operators of the autonomous vehicle of potential concerns but also to trigger any necessary changes to the training of new drivers.

[26] Section 5.5 – even in closed roads, it is important that any testing is covered by the risk assessment processes that are advocated throughout the Code of Practice.

[27] Section 5.11- it is important that scenarios involving the unexpected loss of connectivity be tested prior to trails on public roads and that driver/operators demonstrate capability to respond in an appropriate manner. Similarly, a systemic approach to risk assessment should include a range of failure modes that might be anticipated during any field trials.

[28] Throughout – there is a general concern that the Code of Conduct does not consider the traceability issues that must be addressed if any concerns during trials are to trigger appropriate interventions in the underlying software and hardware. It should be possible for an external agency to view the results from a trial and then to trace back subsequent changes to the underlying systems that were triggered as a result of that test.