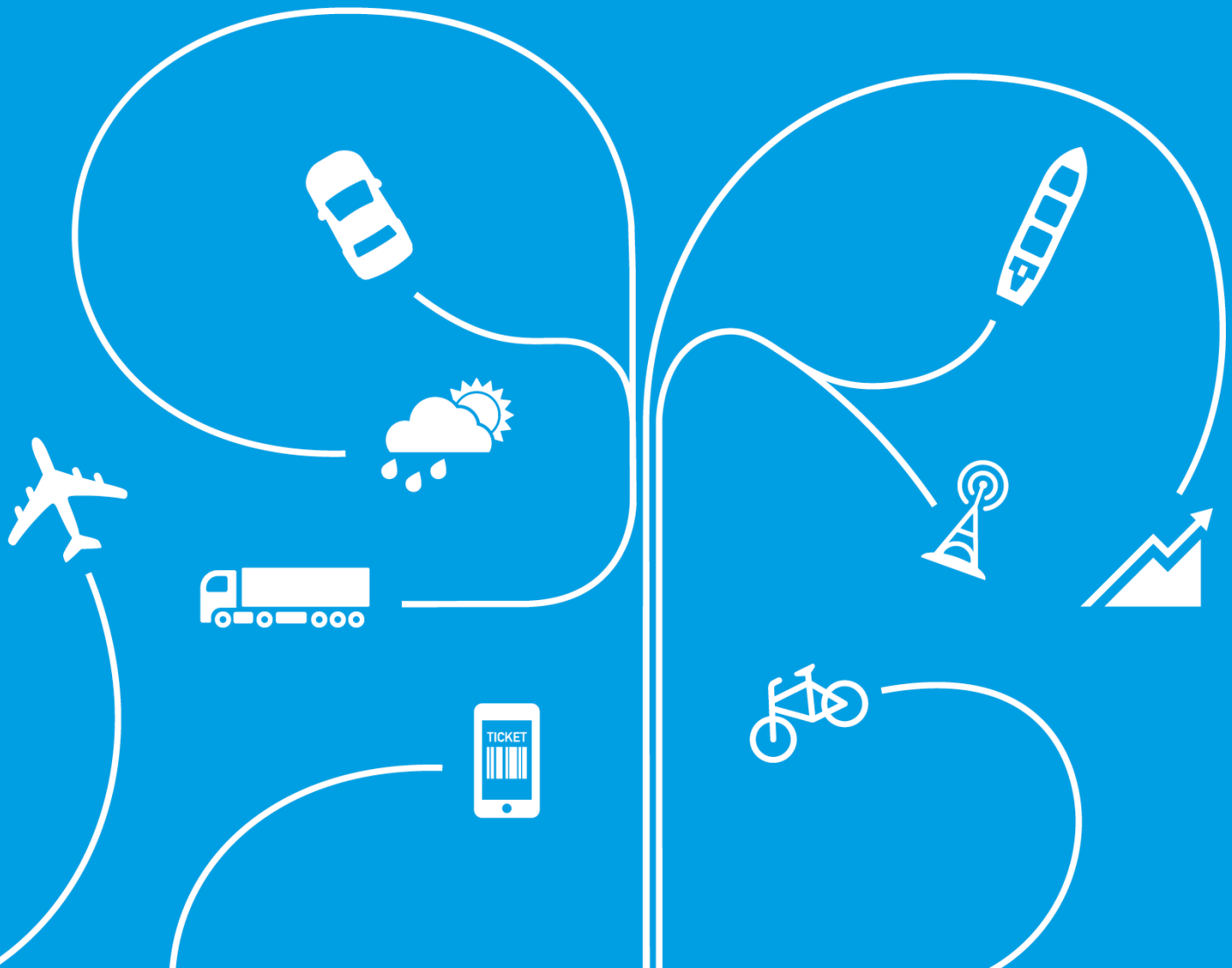
EXPLORING INTELLIGENT MOBILITY

CATAPULT Transport Systems

Taxonomy of Scenarios for Automated Driving

Technical Report

April 2017



Contents

Release Conditions

THIS DOCUMENT AND THE INFORMATION IN IT ARE PROVIDED IN CONFIDENCE, FOR THE SOLE PURPOSE OF USE BY THE TRANSPORT SYSTEMS CATAPULT, AND MAY NOT BE DISCLOSED TO ANY THIRD PARTY OR USED FOR ANY OTHER PURPOSE WITHOUT THE EXPRESS WRITTEN PERMISSION OF THE TRANSPORT SYSTEMS CATAPULT, NOT TO BE UNREASONABLY WITHHELD.

Disclaimer

This report has been produced by the Transport Systems Catapult under a grant from the Innovate UK. Any views expressed in this report are not necessarily those of the Department for Transport or the Centre of Connected and Autonomous Vehicles.

Acronym List

ABS	Antilock Braking System
ACARP	As Confident As Reasonably Practicable
ACC	Adaptive Cruise Control
ADAS	Advanced Driver Assistance Systems
A.I.	Artificial Intelligence
ALARA	As Low As Reasonably Achievable
ALARP	As Low As Reasonably Practicable
ALR	All Lane Running
AMOR	Asset Maintenance and Operational Requirements
ANPR	Auto Number Plate (licence plate) Recognition
APS	Assisted Parking System
ASIC	Application-Specific Integrated Circuit
ASIL	Automotive Safety Integrity Level
AV	Automated Vehicle
BCR	Benefit-cost-ratio
CAA	Civil Aviation Authority
CAV	Connected and Automated vehicle
CFMEA	Concept Failure Modes and Effects Analysis
CIHT	Chartered Institution of Highways and Transportation
CMOS	Complementary Metal-Oxide Semiconductor
CNN	Convolution Neural Networks
DFMEA	Design Failure Modes and Effects Analysis
DfT	Department for Transport
DMRB	Design Manual for Roads and Bridges
DSRC	Dedicated Short-Range Communications
DSIWG	Data Safety Initiative Working Group
DSP	Digital Signal Processors
DVSA	Driver and Vehicle Standards Agency
ECAP	European Car Assessment Programme
EGNOS	European Geostationary Navigation Overlay Service
Ego Vehicle	The vehicle under consideration
EPAS	Electronic Power Assisted Steering
ERAP	European Road Assessment Programme
ERF	European Road Federation
ESC	Electronic Stability Control
EU	European Union
EV	Electric Vehicles
FMEA	Failure Modes and Effects Analysis (FMEA)
FoV	Field of View
FPGA	Field-Programmable Gate Array
FRAM	Functional Resonance Analysis Method
FTA	Fault Tree Analysis

GPU	Graphical Processing Unit
GSN	Goal Structuring Notation
GNSS	Global Navigation Satellite System
HAZOP	Hazard and operability study
HiL	Hardware in the Loop
HMI	Human Machine Interface
I2V	Infrastructure to Vehicle
LDM	Local Dynamic Maps
LDW	Lane Departure Warning
LKA	Lane Keep Assist
Parc	Total number of vehicles in the country
QoS	Quality of Service
SFAIRP	So Far As Is Reasonably Practicable
SiL	Software in the Loop
SoC	System on Chips
SOTIF	Safety of the Intended Functionality
SRD	Systems Requirement Document
STPA	Systems-Theoretic Process Analysis
TSR	Traffic Sign Recognition
UAS	Unmanned Aerial Systems
VMS	Variable Message Signs
V2I	Vehicle to Infrastructure
V2V	Vehicle to Vehicle

1 Introduction

1.1 Overview

This document has been prepared by the Transport Systems Catapult (TSC) for the Department for Transport (DfT) and the Centre for Connected and Autonomous Vehicles (CCAV). This report represents the deliverable of the Project – 'Taxonomy of Scenarios for Automated Driving'. Any views expressed in this report are not necessarily those of the DfT or CCAV.

The majority of the driving task is relatively routine, but occasionally situations demand the driver to take action which is out of the ordinary or requires the driver to make an interpretation of the situation and act in a considered manner (common sense driving). Such situations could present challenges to Automated Vehicles (AVs) and their developers. AVs will need to adhere to rules governing their behaviour. If the rules and regulations governing vehicle behaviour within abnormal situations are not clear, then this could lead to unexpected or undesirable behaviour amongst AVs. Indeed, AVs may behave differently to the same abnormal situation depending on the AV manufacturer and the software algorithms that have been deployed.

Examples of abnormal situations that can be found on the UK highway network might include (but not limited to) responding to emergency vehicles, navigating temporary traffic management measures, responding to children that are near to the carriageway, overtaking an unsteady cyclist travelling up hill, overtaking broken down vehicles, etc. It could include circumstances under which vehicles would need to mount the footway, or cross solid white lines in order to make progress or make way. For instance, merging in turn could create a challenge, as could allowing gaps at junctions for other vehicles to pull out. There are numerous situations that can call upon the driver using 'common sense' to allow the traffic to flow freely or to avoid unnecessary road blockages.

This project investigates abnormal driving situations, and goes on to create a comprehensive 'taxonomy of scenarios for automated driving'. It is anticipated that the outputs of this study will start to outline possible strategies for addressing how to handle some of the more challenging automated driving scenarios.

1.2 Approach

The term 'taxonomy' refers to the branch of science concerned with classification. It is useful to classify automated driving scenarios so as to create a structured approach to dealing with them, and manage the complexity of understanding how AVs will cope with real world issues. Therefore, this study has created a structured approach to classify automated driving scenarios. This includes:

- A structured approach to help resolve the bewildering number of 'What ifs'?
- A structured approach to grouping the scenarios and issues in a taxonomy exercise.
- A structured approach to forming a model to describe (as an example) how the issues will be managed.

Ask one hundred drivers to list their challenging driving anecdotes then they may produce a list of one hundred issues. Ask another group of drivers and there would in all likelihood be a new list, with some

duplication as well as many new items. By performing a taxonomy on those scenarios, it is possible to reach an understanding of the broad classes of issues facing AVs in the hope of a common concept emerging to deal which each class of issue. Using this approach, it should be possible to understand the likely behaviours and risks emerging from any conceivable eventuality without having codified 'IF-THEN-ELSE' rules to explicitly handle each unique situation that presents itself.

There were two approaches to undertaking this project; the 'bottom up' brainstorming approach and the 'top down' modelling approach.

The brainstorming approach involved creating a list of issues for the taxonomy using conventional brainstorming. This forms a bottom up approach by anticipating the sorts of things that could be challenging based upon everyday driving experiences and reports of more obscure incidents. The modelling exercise has taken more of a top down approach by considering how the system will accommodate some of the challenges predicted by the taxonomy exercise. Ideally the two should meet in the middle, however the groupings of the issues don't necessarily need to map directly to the coping strategies, provided there is full coverage, i.e. there is a mechanism which handles each of the issues even if the strategy is avoidance, by not allowing the vehicle to be exposed to that issue. As an example, a Highway Pilot type feature would not normally be expected to cope with the arbitration at a rail level crossing, so there is coverage despite there being no explicit software algorithm designed into the system for this. How to ensure the feature is only used on the highway and not on more minor roads (which may have level crossing among many other things) will form part of the model's argument for that feature, and may include strategies from just relying on the driver to only use it where it has been designed to work, for instance through to the use of GNSS/GPS and geo-fencing to actively prevent it being used elsewhere.

The creation of the model's strategies was undertaken using brainstorming sessions. For this reason, the strategies will inevitably be non-exhaustive in at least some cases, and are provided as examples as much as anything else. The modelling technique used is known as Goal Structuring Notation (GSN) and is a diagrammatic from of argumentation which is explained in more detail in Section 4. It is important to stress (particularly considering stakeholder review feedback of a previous draft version of this report) that the use of GSN is not in and of itself a way to ensure full coverage of all potential issues associated with an automated driving feature. The creation of the GSN may be helpful with that thought process by structuring the thinking, however it remains just a tool to diagrammatically present the 'argument' that you have full coverage of issues rather than it actually doing the work of finding that coverage.

Beyond the use of just brainstorming, more rigorous techniques might be employed to deal with the complexities that may not be apparent. It was beyond the scope of this work to investigate those potential techniques since it began as an exercise to explore issues faced in everyday conventional driving in the context of AVs. However, two techniques for modelling and understanding accident causality, have been highlighted by a stakeholder since the first draft release of this report. The first is Systems-Theoretic Process Analysis (STPA)[1], a hazard analysis technique used to identify scenarios leading to identified

[1] STPA Primer (paper) and Engineering a Safer World (book)
http://psas.scripts.mit.edu/home/home/stpa-primer/

hazards so they can be eliminated or controlled. STPA generates a larger set of failure causes compared to traditional techniques. These causes may not involve component failures or unreliability. STPA was designed to also address increasingly common component interaction accidents, which can result from design flaws or unsafe interactions among non-failing (operational) components. The second, known as the Functional Resonance Analysis Method (FRAM)[2], is another technique which helps to model the interactions or resonance which can occur between normally benign actions and events which can sometimes lead to unexpected and disproportionate outcomes. Put simply, it does so by dividing the system up into functions then considers the potential variabilities on the interactions between those functions.

A final consideration of the approach taken in this work is that it largely considers the issues in isolation. This may be a hang-up from current automotive thinking of requiring robustness to single-point-failures which assumes that failures are rare therefore can be treated as occurring discretely in time upon which a safe condition can be entered (usually power down, deactivate or inhibit then rely on the driver to manage the situation). However, when dealing with issues of the real-world environment rather than component failures, we may find that multiple issues occur at any one time. It may be valid to assume that rare unusual events will not occur at the same time (to any realistic probability), such as an animal being loose on the highway will not occur in the same instant that a sink hole appears in front of a collapsing bridge while at the same time the pilot of a light aircraft is looking for a good place to perform an emergency landing. However, there is no reason to assume that any of these will not occur during poor weather conditions while people are also fleeing their abandoned vehicles on-foot on the highway due to the same event. Multiplicative complexity is a real issue that remains whilst dealing with issues one-at-a-time is still challenging enough.

1.3 Report Structure

This report is structured as follows:

- Section 2 summarises background information that was researched as part of the study, including a summary of relevant TSC studies and external reports, the Highway Code, and automotive safety related information which is of relevance to the challenges associated with AVs.
- Section 3 includes an initial long list of challenging driving situations for AVs.
- Section 4 introduces GSN, which is the tool used to structure the issues associated with challenging AV driving scenarios, and outlines some considerations with its use.
- Section 5 presents the GSN outputs associated with the Highway Autopilot use case.
- Section 6 presents the GSN outputs associated with the Urban Pilot use case.
- Section 7 provides concluding comments.

[2] FRAM Handbook
http://functionalresonance.com/how-to-build-a-fram-model/fram-handbook.html

2 Background Research

This section summarises some of the background research that was undertaken as part of this project.

2.1 Relevant TSC Studies

This document sits alongside other reports undertaken by TSC on behalf of DfT / CCAV. The following studies were completed within the 2015 / 16 financial year:

Project Name	Study investigates / provides:
Data recording guidelines for AVs	• Reasons for data logging • Importance of logging different types of data • Guidelines for AV data logging
Functional Safety (with respect to AV regulation)	• Current process for automotive safety • How safety cases could be developed for AVs • Considerations for regulators.
Pods on pavements	• Collated information with respect to automated vehicles operating in pedestrian environments
AV map data requirements	• Requirements of map data for AV systems
Planning and preparing for CAVs	• Stakeholder / literature recommendations on what public sector can do to accelerate CAV agenda.
Exploring the relationship between CAVs and energy usage and emissions	• Systems dynamics modelling tool and learnings for examining relationships
Investigating the case for safety database for AVs	• Database benefits and risks • Potential database features • Implementation options

Table 1: Relevant TSC Projects undertaken 2015 / 16 Financial Year

The following projects are scheduled to be completed by the end of the 2016 / 17 financial year:

Project Name	Study investigates / provides:
Safety Aspects of AVs	• How safe high level automation features need to be • Approaches to safety cases from technology developers • Societal attitudes towards risk from CAV technology
Retrofitting connected vehicle technology	• Minimum level of vehicle retrofit needed to facilitate selected connectivity applications
Standards for CAVs	• Builds on previous TSC work to map existing relevant standards in CAV domains.
Future Proofing Infrastructure for CAVs	• Opportunities for updating policy and guidance documents with respect to CAVs to help future proof UK infrastructure
Landscape review of CAV research	• Current work being undertaken by UK research organisations which could be relevant to CAVs

Table 2: Ongoing TSC Projects, 2016 / 17 Financial Year

The latest versions of the deliverables of each project have been sent to the project sponsors for each project. Interested parties can request further information from TSC.

Of particular relevance is '*Functional Safety (with respect to AV regulation)*' and '*Safety aspects of AVs*' projects, both of which consider the safety of AVs. The findings of this report will feed into 'Safety aspects of AVs', and vice versa.

2.2 The Highway Code

In order to start compiling a list of numerous real-world issues faced in everyday driving, the Highway Code was consulted, which represents the rules of the road for road users. Within the introduction to the Highway Code it is stated:

> *"Many of the rules in the Code are legal requirements, and if you disobey these rules you are committing a criminal offence. You may be fined, given penalty points on your licence or be disqualified from driving. In the most serious cases you may be sent to prison. Such rules are identified by the use of the words '**MUST/MUST NOT**'. In addition, the rule includes an abbreviated reference to the legislation which creates the offence.*
>
> *Although failure to comply with the other rules of the Code will not, in itself, cause a person to be prosecuted, The Highway Code may be used in evidence in any court proceedings under the Traffic Acts (see The road user and the law) to establish liability. This includes rules which use advisory wording such as 'should/should not' or 'do/do not'."*

The Highway Code has been written, naturally, for human drivers, and assumes an associated level of visual perception and intelligence. For example, 'Rule 144' states you MUST NOT:

- *Drive dangerously*
- *Drive without due care and attention*
- *Drive without reasonable consideration for other road users."*

It may be difficult to define to an AV control system what '*reasonable consideration*' for other road users is. Studying Rule 144 in conjunction with the paragraph above, leads one to surmise that failing to act with reasonable consideration for other road users represents a legal requirement.

Another example where an associated level of perception and intelligence is required can be found within Rule 200, which states:

> *"Choose an appropriate place to manoeuvre. If you need to turn your vehicle around, wait until you find a safe place. Try not to reverse or turn round in a busy road; find a quiet side road or drive round a block of side streets."*

An AV may struggle to determine the points within this rule, such as how to decide what is a safe place and what constitutes a busy or quiet road. It is also unclear what the vehicle should do if it must turn around but cannot get to a 'quiet' road to do so.

Rule 200 presents a different type of challenge for AVs:

> *"Older drivers. Their reactions may be slower than other drivers. Make allowance for this"*

To determine even the approximate age of other road users represents a considerable challenge for AVs. Even if it was possible to do so, it is then unclear how the AV expected to 'make allowance' for them.

2.3 Examples of Automated Driving Taxonomy Exercises

A relevant report, titled 'Use Cases for Autonomous Driving' was published by Walther Wachenfeld and Hermann Winner in 2014.[3] The report outlines a set of characteristics to describe automated driving use cases. These include:

- Type of occupant;
- Maximum permitted gross weight;
- Maximum deployment velocity;
- Scenery (type of road on which vehicle can operate);
- Dynamic elements (extent to which vehicle can mix with other road users, i.e. level of segregation);

3 https://www.daimler-benz-stiftung.de/cms/images/dbs-bilder/foerderprojekte/villa-ladenburg/Villa_Ladenburg_Use_Cases_English_Release_2.pdf

- Information flow between driving robot and other entities (for example, vehicle may receive path planning information or even direct control commands from outside sources);
- Availability concept (the extent to which the driver, or other entities, can take control of the vehicle within the operational envelope);
- Extension concept (who, if anyone, can take control of the vehicle at the boundary of the operational envelope);
- Options for intervention (on what basis can the occupant, or other entities, intervene in the driving task).

This study is useful in defining types of AVs, and their capabilities. In addition, the Adaptive Deliverable 2.1[4], which presents a systematic approach for the classification of automated driving and parking functionalities, was reviewed.

2.4 Stakeholder Engagement

The project and its approach was discussed with a number of key stakeholders. The themes discussed fed into the formulation of potential strategies for handling driving scenarios, discussed later in this report, but it revealed the variety of approaches being taken to the issues associated with automated driving.

2.5 Mind the Safety Gap

This section outlines general research and discussion into safety issues related to the automotive industry and AVs. The following is considered as important 'food for thought' prior to discussion of the taxonomy itself, and builds on information provided or being developed as part of previous and ongoing TSC studies, as outlined in Table 1 and Table 2.

2.5.1 Current Industry Practice

ISO 26262 is the automotive industry self-adopted IEC 61508 derivative functional safety standard which has the goal of avoiding potentially safety critical situations caused by hardware and software failures and is generally applied to systems under software control. A limitation of this standard which is not always recognised is that it only covers fault failures and not the so-called Safety of the Intended Functionality (SOTIF). Therefore, the functional insufficiencies of ADAS features are not covered and manufacturers are left to satisfy themselves that the systems are robust and reliable enough to sell in terms of the potential for litigation and damage to their reputation.

Weaknesses in estimation, interpretation and prediction steps can have consequences comparable to those of hardware and software failures, yet the means to formally ensure that equivalent safety to fault

[4] https://www.adaptive-ip.eu/index.php/deliverables_papers.html?file=files/adaptive/content/downloads/Deliverables%20%26%20papers/AdaptIVe-SP2-v12-DL-D2.1%20System%20Classification.pdf

failures is met has not yet been put in place or even a consensus reached as to how it should be undertaken. An attempt to address this issue for ADAS (SAE Level 0-2) has been made, currently in draft at the time of writing[5].

The main vehicle motion controls are formed of the familiar accelerator, brake and steering mechanisms. When these are brought completely under software control there are new risks associated with each of them. The principal is the same for all three, but it is perhaps easiest to visualise the potential hazards in the case of commanded steering, since a relatively small error in steering angle can result in a catastrophic head-on collision between vehicles.

Existing Electronic Power Assisted Steering systems (EPAS) have been developed to meet ISO 26262 Automotive Safety Integrity Level (ASIL) – D. ASIL is established by performing a risk analysis of a potential hazard by looking at the Severity, Exposure and Controllability of the vehicle operating scenario. Four levels, A to D, are used to represent ASIL with D representing the most stringent and A the least stringent level. ASIL D represents likely potential for severely life-threatening or fatal injury in the event of a malfunction and requires the highest level of assurance that the dependent safety goals are sufficient and have been achieved. This concerns both the process under which EPAS are developed as well as the design implementation. The ASIL-D rating comes because of the hazard arising from unintended or unintentional steering, being of the worst combination of exposure, lack of controllability by the driver, and severity of the consequences. When considering fully automated vehicle control (which is not within scope for ISO 26262) it is not clear what to do with the controllability rating *C,* except to set it to the case worst level C3 (90% or more of drivers would not be able to maintain enough control to avoid the hazard). The hazard of steering into the path of an oncoming vehicle cannot be understated as the closing speeds can be over double the speed limit (if both vehicles are slightly exceeding it) and the results of a collision are often catastrophic. If the steering command is passed from another software function hosted on another module, then the same hazard severity remains with the same potential consequences.

2.5.2 Fault Tolerant Fail Active Actuator Design Is Now Needed

From a component perspective, the existing design of EPAS has been done with the mechanical fall back left in place (the steering wheel is physically connected to the road wheels) so that it can be inherently failsafe just by deactivating the steering assistance. The very recent emergence of steer-by-wire into the market place still leaves a mechanical fall back using a clutch to reconnect the mechanical link to the wheels. If a fault is detected, then the system can power down and leave the driver with heavy but controllable steering. If a mechanical fault occurs such as a snapped drive belt between the drive motor and the steering rack, the driver is still able to steer the vehicle unassisted. If the driver is no longer expected to be in a position to immediately take back control at all times due to automation of the driving task, then it is likely that some redesign and re-evaluation of the fault failure cases will be needed. Because of the lack of human involvement, it may no longer be acceptable for power steering loss to occur from a single drive belt breakage or for a systematic fault to cause the steering system to just power down and go in to a fault mode. The same may also be true for the brake system which has a mechanical-hydraulic fall back, assisted or unassisted foundation braking is always available to the driver, so redundancy of the

[5] ISO/AWI PAS 21448 Road vehicles -- Safety of the intended functionality.

electrical systems for fail-active fault-tolerance behaviour has not thus far been needed for any of the primary controls.

2.5.3 Has Constrained Operation Slid into Full Authority?

Fault tolerance and fail-active behaviours are of limited concern when compared to the issue that commanded steering has the potential to cause. Pre-existing systems such as park assist can be offered without this being an issue, as their speed of operation is limited by the EPAS controller, this is known as ASIL decomposition. The high integrity module (EPAS) constrains the steering commands from the low integrity module and so ensuring safety. If the vehicle's speed exceeds a certain threshold, then the EPAS controller stops accepting external steering commands from the parking module. The available steering torque for commanded steering is also limited in the order of 3Nm, such that the driver can manually overpower and intervene if required. This functionality can be protected to ASIL-D as part of the EPAS controller's design and is generally covered by Type Approval for many regions in the world. Having this architecture allows system designers to implement automatic parking functions from a hardware module other than the EPAS controller which may have no ASIL rating, or usually ASIL-QM (quality management) and can be running arbitrary software acting on information provided from low integrity ultrasonic parking sensors. The driver is responsible for monitoring the environment around the vehicle and in the event of a failure.

Crucially, if we now remove the driver from the parking manoeuvre and require the EPAS controller to accept steering command potentially at any speed and full steering torque, then this changes everything as the ASIL-D protected constraints have been removed; the software issuing the steering commands has become responsible for preventing an incorrect steering decision whilst the EPAS controller is now responsible for executing it in the way requested. The change in steering authority means that the module issuing the steering commands has now also inherited the ASIL-D hazards whilst the EPAS controller has new additional responsibilities to guarantee that the steering commands are executed.

2.5.4 Tech Demonstrations are not Proof of a Safety Concept

What is currently being demonstrated, and even promised for production, is to produce steering commands from high performance, but general purpose, computing hardware which sits at the opposite end of the spectrum to the type of hardware currently used for ASIL-D applications, namely low performance highly reliable/resilient lockstep microcontrollers rated to AEC Q100 for suitability for operation in harsh automotive environments. A recent trend has seen the move towards high performance parallel computing. This can take the form of Graphical Processing Units (GPUs), Digital Signal Processors (DSPs), programmable logic Field-Programmable Gate Array (FPGAs) and other hardware variants such as custom silicon Application-Specific Integrated Circuits (ASICs). High processing performance is needed to resolve the complexity found in the real world, whether for image processing and object recognition, or using Artificial Intelligence (AI) techniques such as Convolution Neural Networks (CNNs) used for both image processing object classification, situation recognition and for decision logic. These processing units are normally incorporated with additional conventional CPU cores onto what are known as System on Chips (SoCs) such as those used in consumer electronics e.g. smart phones and tablets, and often lack the non-obligatory AEC Q100 rating. This heterogeneous silicon design does not lend itself well to current safety practises, both from a software and a hardware perspective, and fault failures may again become an issue. However, the real difficulty is that high performance processing may be required from the

perspective of resolving the complexity of the real world which leads us into the paradox of what it means to be safe, as discussed below.

2.5.5 The Design Complexity Paradox

There is an ongoing debate as to how many miles should be driven before AVs are '*proven*' to be safe. Existing EPAS controllers may have been '*proven-in-use*' as is said in the industry. However, they have been developed via a process to be safe-by-design rather than relying upon accumulated mileage to demonstrate that they are safe. It is possible to have several billion miles driven without issue as a fleet total, then in the next mile an area of memory or processor register is corrupted by events beyond the developers and designers immediate control that has not been corrupted before which then results in a catastrophic failure if this has not been explicitly accounted for by a failsafe design. This is a fault failure which is difficult to prevent, but the effect of such a failure can be guarded against at the design stage by accepting that random failures can occur at any time. This example just considers fault failures which are covered by ISO 26262. The systems currently being demonstrated for highly automated vehicles do not currently meet with this fault failure approach from both the perspective of the general-purpose hardware it is running on and the architecture of the software itself. It is conceivable that the hardware could be developed to meet the safety integrity requirements with significant time and effort, but is it less clear how this can be achieved for the software at its necessary complexity is entirely at odds with the normal processes for development of safety critical software. A relevant article was published by Dr Steven Shladover in the June 2016 edition of Scientific American:

> *"Software for automated driving must therefore be designed and developed to dramatically different standards from anything currently found in consumer devices. Achieving these standards will be profoundly difficult and require basic breakthroughs in software engineering and signal processing. Engineers need new methods for designing software that can be proved correct and safe even in complex and rapidly changing conditions. Formal methods for analyzing every possible failure mode for a piece of code before it is written exist—think of them as mathematical proofs for computer programs—but only for very simple applications. Scientists are only beginning to think about how to scale up these kinds of tests to validate the incredibly complex code required to control a fully automated vehicle. Once that code has been written, software engineers will need new methods for debugging and verifying it. Existing methods are too cumbersome and costly for the job."*

This is particularly true for CNN approaches where the input/output relationship is obfuscated by its nature of modelling complex processes and many of the normal techniques that provide checks and measures such as limiting design complexity and code reviews do not readily apply as the link between the code and the functionality is not clear. This leads us into the notion of functional insufficiencies.

2.5.6 Functional Insufficiencies

While sales volumes of CAVs remain a low proportion of the total vehicle parc, the limitations of algorithms are more likely to manifest in real-world incidents than hardware fault failures since these can take many cumulative years of fleet operation to manifest. Current ADAS features rely heavily on the driver to take over control in the event of an anomaly, so can err on the side of caution against positive actuation. They

can favour inaction over action and allow the driver to override in either case. Even stability control systems are there to assist the driver, not to guarantee stability under all conditions. The approach can often be leaning towards providing sufficient actuation to pass the tests (such as Euro NCAP), but be conservative in the real world to minimise false activations.

However, with the driver removed from the loop, both false positive (action when not required) and false negative (inaction when action is required) can be equally hazardous with potentially much higher consequences without a driver to handle the situation.

Sensing and perception limitations often mean that the trade-off between preventing false negatives and false positives is complex and difficult, as to improve the performance of one results in a significant worsening of the other, with the middle ground providing unacceptable performance of both. This detection trade-off is known as the receiver operating characteristic of a sensor or system.

2.5.7 Receiver Operating Characteristic Trade-offs

Often there may be no suitable compromise which prevents both false negatives and false positives with a system that performs some form of detection function, which for a CAV could be something such as detecting a pedestrian crossing the road. This problem is highlighted with Tesla's version 8 software update. The system on board the Tesla vehicle could not rely solely on either the radar or camera sensors, as there are times when both do not work effectively. Radar is not generally used stand alone for stationary objects, since there would be constant false activations instigated by environmental features such as passing gates, manhole covers, and low bridges. The camera cannot cope with all light conditions and atmospheric conditions such as fog, mist, heavy rain etc. The Tesla strategy to resolve the functional gap is to use a crowd sourced map of false-positive radar detection locations which can then be used to inhibit unwanted brake activations that would be based solely upon radar derived sensing. There is nothing to stop a legitimate hazard coinciding with an identified false positive location, although this is statistically less likely to occur and will inevitably improve confidence in the system without the underlying problem having been resolved. However, the driver is still expected to take ultimate responsibility for the vehicle and be ever vigilant, so this is an example of a fall-back measure which reduces the statistical likelihood of an inattentive driver crashing into a collision risk with the Autopilot feature active. For the current Tesla fleet size this may reduce the actual crash occurrence rate to zero, but does not actually remove the problem, only masking it whilst placing a new reliance upon external data for the safe operation of the vehicle.

2.5.8 Approaches to Testing

At the current time the only tool in the box seems to be on-road testing (and possibly simulation) to see what happens. The difficulty with an approach of endurance testing is that safety cannot be proven just by the absence of failures. The automotive industry has often used a proven-in-use argument (particularly prior to the adoption of ISO 26262) to cover the quality and safety of component design uncertainties and component reuse. This is only really a confidence measure which helps to ***assure*** component safety but does not always ***ensure*** it. Since road crashes and deaths are rare events when considered for individual vehicles, miles without crashes are a poor surrogate marker for the ultimate competency of safety of the software system. Some events happen at such infrequent intervals (examples could include bridge collapses, animals in the road) that are independent of distance travelled and exposure to these events cannot be accelerated just by accumulating more mileage on the road. Testing for determinism with system reaction and behaviour to these events can only be achieved by inducing them under controlled conditions rather than waiting for them to happen to a test vehicle. Another approach is to accept the system limitation and instead try to limit or prevent the real-world occurrences so that the system will not likely ever have to encounter them, thus circumventing the need to test.

> The benefit of the approach of this taxonomy exercise is that it may help to move from a performance based testing process to one of compliance. The method of identifying classes of issues, deciding upon a strategy to handle or mitigate them, and looking for compliance/conformance with that strategy, leaves far less to chance over a performance based approach. With a performance approach where miles are driven to increase the chances of encountering an unusual event which may or may not push the system beyond its limits, the absence of which is then offered as proof that the system is safe for deployment.

Finally, by always assuming that the system can and will fail under certain conditions, the performance during failure can be assessed by inducing the failure rather than waiting for them to happen. For example, if a Vehicle to Infrastructure (V2I) strategy is assumed for signalling at junctions, then the performance during a signal outage or corrupted message can be demonstrated by design and then tested to make sure that the vehicle still cannot run a red light, rather than testing to show that the communication system is robust and therefore can be relied upon (until the day it cannot be).

2.5.9 A Safer Nested Architecture?

Is there an architecture which can use simpler legacy control methods to constrain the functionality and ensure a basic minimum level of safety and predictability without preventing the expected action to be taken during unusual and challenging situations? It is hard to conceive how this could work without providing a means for the complex (lower integrity) system to override the simpler (high-integrity) one, which provides a weak link that defeats the original objective of providing high-integrity constraints. If this could somehow be achieved, then it may allow more complex systems to be supervised and 'plausibility' or 'sanity checked' by a simpler safety rated software system. Could it allow far more nuanced adaptive behaviours to be applied, provided they can live within enforced behavioural limits? Or will there always be complex scenarios which will require full control authority to resolve and mitigate hazards?

A similar idea is presented in a paper[6] presented at the 2016 SAE World Congress, which raises many testing and validation concerns for AVs along with the implied recommendation for a phased deployment. A monitor/actuator approach is described as well as a failover mission strategy, using a separate high integrity failover autonomy system capable of bringing the vehicle to a safe state using minimal redundancy. In this case the failover system might be a simpler redundant system capable only of bringing the vehicle to a safe rest position (the 'mission') when a problem or abnormality is detected in the main control system.

Other designs might be based around so-called *Byzantine* fault tolerance, where redundant controllers are used to vote to agree on the correct control action. To help prevent common mode failures (all controllers make the same mistake for the same reasons), the controllers should be of diverse design, unlike typical modern aircraft systems which expect identical outputs from homogenously redundant systems. Where two controllers are used and the outputs do not sufficiently agree, then a failover mission capability would be required to bring the vehicle to a safe condition. If three are used in a majority voting system, then at least two of them would have to agree, else the failover mission capability would again be needed. Even if this approach is found to work, then it remains a difficult challenge to ensure genuine independence between the systems, without creating an unusable system, where they are always *bickering*, failing to sufficiently agree, so that the vehicle is constantly being brought to a halt as part of its failover mission.

2.6 Automation vs. Autonomy

The terms autonomous vehicle and automated vehicle are being used interchangeably as synonyms of each other, and there is debate over definition of what full autonomy actuals means. Some dictionary definitions are as follows:

1. *"One who gives oneself one's own law"*
2. *"Freedom from external control or influence; independence"*
3. *"The right or condition of self-government"*
4. *"(philosophy) The capacity to make an informed, uncoerced decision."*
5. *"(mechanics) The capacity of a system to make a decision about its actions without the involvement of another system or operator."*

A comparison with the aviation industry and autonomous *cars* is made by a report[7] on the challenges facing an autonomous car's risk assessment; SCSC Developing Safe Systems' where the UK Civil Aviation Authority (CAA) are described as using the second dictionary definition above to define autonomy for Unmanned Aerial Systems (UAS) and states that no UAS currently meet the definition of autonomous, instead they are *highly automated* or *high authority automated systems*. The CAA requires that all UAS perform

[6] Challenges in Autonomous Vehicle Testing and Validation, Koopman et al, SAE 2016-01-0128
https://users.ece.cmu.edu/~koopman/pubs/koopman16_sae_autonomous_validation.pdf

[7]The challenges facing an autonomous car's risk assessment, Developing Safe Systems
Proceedings of the Twenty-fourth Safety-Critical Systems Symposium, Brighton, UK
http://scsc.org.uk/paper_131/18%20Spencer%20-%20The%20challenges%20facing%20an%20autonomous%20cars%20risk%20asessment.pdf?pap=999

deterministically with their response to any set of inputs being the result of pre-designed data evaluation output activation processes.

There has also been much recent discussion as to the miss-naming of features using the term pilot such as Tesla's Autopilot, Audi's Piloted Driving, and Volvo's Pilot Assist, with suggestions that these names are misleading. The term Autopilot is associated with aviation and marine sectors from which it does not refer to any single set of functionality, but rather refers to what is in effect an *'auto-mation'* pilot used to reduce the workload of trained operatives and not an *'auto-nomous'* pilot. Aside from the possible pubic misconception, the real limitation is that what is required and expected for so-called driverless cars is full autonomy. A problem arises due to the fact that the superposition of many automation features produces something which appears to approximate autonomy but in fact falls dangerously short of it, as the system is merely responding to inputs and lacks the contextual awareness that is expected of a human driver to overcome the challenges of everyday driving. Many unusual scenarios may appear complex or difficult to detect for automation software, but if presented to a human driver may result in a simple and obvious mitigation such as just slowing or stopping the vehicle. The word automation can suggest repetitive factory production line actuation, which has no feedback control, and requires constant monitoring and supervision. However, what is currently being demonstrated by the automotive and technology industries is a form of sophisticated automation and with that comes limitations and risks arising from unexpected scenarios. As with any automation system, these residual risks need to be scoped and planned for at the design stage so that they can be mitigated for (i.e. using ALARP principles). If society is prepared to take on these new risks, and be persuaded by the benefits of automation, then the risks should be evaluated and stated plainly, and perhaps put to some form of public consultation. The danger otherwise is that before the risks are realised, they are gradually introduced through increasing levels of vehicle automation being miss-labelled as autonomy and potentially played down by the vehicle manufacturers as they compete amongst each other for market share by offering the latest boundary/precedence stretching feature. Once the accepted risk tolerance level is understood or decided, then this can be used as a basis for setting target maximum acceptable occurrence rates of undesirable scenarios against which systems and procedures can be developed, tested against then monitored with the aim of continuous improvements.

3 Initial List of Challenging Driving Scenarios

In order to start to classify potentially challenging driving scenarios, it was first necessary to list them. A brainstorming exercise held amongst TSC staff produced a list, and they were then grouped into categories. The resulting list (see Table 3) is not exhaustive, but provides an indication of the type of events considered and the categories into which they were initially placed:

Category	Abnormal / Challenging Driving Event	Issues involved
Obstructions	Parked vehicles	How to ensure they are parked and have not momentarily stopped. How to allow for possibility of doors opening?
	Disabled (broken down / crashed) vehicles	Passing may lead to compromising other rules of road such as crossing solid white lines
	Pedestrians	How much space to leave? Different clearances for different types? Pedestrian behaviour can be unpredictable. Should vehicle slow down when passing pedestrians on footway?
	Passing cyclists	How much space to leave? Different clearances for different types? Cyclist behaviour can be unpredictable.
	Road flooding	Difficult to sense the depth? Could lead to loss of control of vehicle or splashing of pedestrians.
	Animals in road (either shepherded or loose)	For smaller animals, it can be difficult to decide whether to pass over animal or attempt to stop or swerve.
	Ridden horses	Determining appropriate speed and overtaking strategy.
	Negative obstructions such as pot holes or road / bridge collapse	Could be difficult to sense.
	Load shedding from other vehicles	Action could depend on density / mass of objects being shed, but might not be possible for machine to determine.
	Vehicles in process of becoming disabled, e.g. tyre blow out, lorry jack knifing, tall vehicle overturning in wind etc.	Challenging for machine to detect and interpret subtle clues that provide indications.
	Traffic calming measures	Speed humps, chicanes, etc. need to be detected and negotiated appropriately.

	Fallen power cables / branches in carriageway	Could be challenging to detect with sensors.
	Level crossing	Similar to traffic signal but consequences of stopping on rail line could be catastrophic.
	Overtaking	Challenging to detect oncoming vehicles
Lane reallocation / rerouting	Temporary lane closure on highway	Road layout may differ from map being referred to by vehicle
	Temporary contraflow	Automated driving feature may be designed for highways and not for two-way traffic operation
	Lane designations (e.g. bus lanes, high occupancy vehicle lane, hard shoulders)	May need to clarify under what circumstances vehicles can enter.
Adverse weather / environmental conditions	High winds	Loss of control
	Snow either falling or on carriageway	Sensor visibility compromised, road markings and kerb lines obscured, loss of vehicle control
	Heavy rain	Sensor visibility compromised, road markings obscured, loss of vehicle control
	Ice	Loss of control
	Fog / bright sunshine etc.	Sensor visibility compromised
Road Etiquette	Emergency vehicle in vicinity	How to avoid impeding whilst obeying rules of road
	Crossing white lines	Under what circumstances can vehicle do this?
	Interpreting gestures from other road users	Challenging to detect and interpret the meaning of hand gestures, flashing of headlights, etc.
Traffic flow arbitration	Police or authorised persons intend to stop AV	How does AV recognise what is an authorised person, and then interpret commands?
	Two lanes merge into one	Often involves interaction between human drivers
	Merging onto highway	Often involves interaction between human drivers
	T junction	Poor lateral field of view from AV
	Cross-roads	Poor lateral field of view, right turn stale-mate
	Temporary speed limits	How to ensure location is communicated to AVs
	Temporary traffic signals	How to ensure location is communicated to AVs
	Temporary stop-go sign	Can AV interpret?
	Giving way to oncoming vehicles on narrow section of road	Often required communication between human drivers to decide who proceeds first
	Roundabouts	Detecting correct lane allocations, ‘give way to right’ standoff

	Zebra crossings	Giving way to waiting pedestrians
	Traffic signal failure	Junction reverts to priority based, or interaction between human drivers

Table 3: Initial long list of challenging driving scenarios for AVs

4 Goal Structuring Notation

4.1 Introduction

This section introduces concepts and terminology associated with the Goal Structuring Notation process.

> The resulting diagrams are presented in full in Sections 5 and 6, but it is important to first understand the methodology, as described here.

This section also includes a discussion as to how the GSN can be used to define a safe and working system, and tolerance of risk which will need to be considered with the move to automation.

Discussion of the outcomes of the GSN is included within Section 5 and Section 6.

4.2 Goal Structuring Notation Concepts and Terminology

Goal Structuring Notation[8], or GSN, is a goal model technique that is can be used to make safety cases to satisfy the regulator in safety-related industries, but can also be applied just as well to a more generalised structured argumentation. In simple terms, one generally starts with a high-level goal such as "My system is safe" given the context of "here's how and where it will be used" and "here's why it's safe…". The idea is to form a structured argument around something, usually why a system will be safe in operation, hence GSN is a structured form of argumentation. Since it is an argument it is subjective and the reader is invited to scrutinise it so as to become convinced as to its validity and therefore that the top-level goal or assertion is true, such as the system described will in fact be safe in use.

The GSN Community Standard defines GSN as the following:

> *GSN is a graphical argumentation notation that can be used to document explicitly the individual elements of any argument (claims, evidence and contextual information) and, perhaps more significantly, the relationships that exist between these elements (i.e. how claims are supported by other claims, and ultimately by evidence, and the context that is defined for the argument). Arguments documented using GSN can help provide assurance of critical properties of systems, services and organisations (such as safety or security properties).*

An advantage of GSN is it helps make complex issues and relationships clear and understandable, without the need for large quantities of descriptive text.

The automotive industry is currently a largely self-regulated industry which collaborates in some ways and competes in others. As it lacks a regulator there is often no safety case written for external review, but the individual companies satisfy themselves of quality and safety through internal standards and adherence to some notion of industry best practise. Techniques such as Failure Modes and Effects Analysis (FMEA) and more recently design (DFMEA) and concept (CFMEA) FMEAs and variants of them are used as documentary evidence that due thought and process has been applied to potentially injury causing

[8]GSN Community Standard Version 1: http://www.goalstructuringnotation.info/documents/GSN_Standard.pdf

failures, and wider quality issues which can lead to financial and reputational losses for the company. FMEAs are conceptually similar in some ways to GSN, but lack the overall narrative for the whole system. They highlight areas of potential weakness, attempt to quantify the seriousness of those weaknesses and describe the mitigating step or actions put in place to address the weaknesses.

For the purposes of this work, GSN has been used to construct a general argument towards the safe and practical deployment of a limited set of automation features (highway and urban automation) for road vehicles in the United Kingdom (UK), although the issues addressed will inevitably be applicable more widely than just the UK. Where appropriate it references when an FMEA would form a basis of support for the argument to show that all plausibly lower level failures have been considered. It is not intended to be a detailed and complete safety case for AVs, however it does try to lead the reader towards the recognition of the practical (and sometimes inconvenient) issues faced on the way to formulating a full safety argument.

It is felt that at the time of writing the industry is tackling the challenges faced in placing an AV onto UK roads in a siloed manner, i.e. by considering the vehicle as an individual entity rather than the vehicle as part of a complete system including infrastructure, other vehicles and the users. As is the case in other industries, each new deployment is considered a new case (change of context) which may invalidate previous assumptions and therefore some re-evaluation is usually required.

The following sections describe some of the terminology peculiar to the GSN techniques. This includes:

- Goals
- Contexts
- Justifications
- Assumptions
- Strategies
- Solutions
- Modules

4.2.1 Goals

A goal is a basic assertion of something which is deemed or required to be true. It maybe that the system is required to perform in a certain way or to a certain performance level to uphold that goal, so goals are supported by further goals until there is nothing further to add except the evidence that the goals will be met.

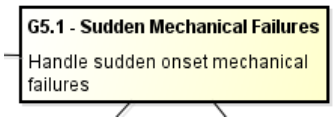


Figure 1 : Example of a GSN goal

4.2.2 Contexts

Context provides the landscape or clarification for a goal or strategy. The highest-level goal should always be provided with some context, as a goal cannot always be met. The system may be safe being tele-

operated on the surface of Mars, but perhaps not when close to a group of pedestrians crossing a road on Earth. Contexts have been used to provide an expanded explanation of goals and strategies to keep the goals and strategies as concise as possible. Changes to context will potentially invalidate the rest of the model without further re-evaluation to ensure that the effects of the changes do not have any new consequences.

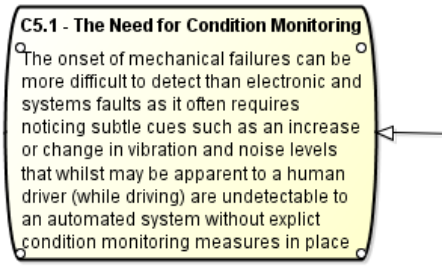


Figure 2 : An example of a Context within GSN

4.2.3 Assumptions

Assumptions are unproven points of contention that are assumed to be true. The reader needs to agree with them in order to agree with that branch of the argument. Not agreeing may cause the reader to negate a particular strategy in favour of another.

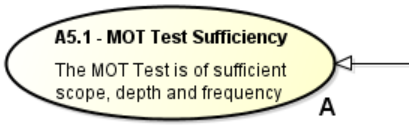


Figure 3 : An example of a GSN assumption

4.2.4 Justifications

Justifications provide an explanation or rationale why an approach has been taken, or why another seemingly obvious or better approach has not been adopted. They can be used to explain goals and strategies.

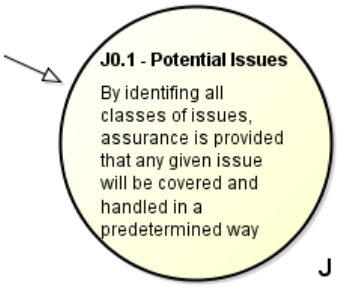


Figure 4 : An example of a GSN justification

4.2.5 Strategies

A strategy in GSN is a declaration of how the argument will be supported, the strategy that is being adopted to provide further reasoning to the preceding assertions. However, for the purposes of this exercise, they

have been used also to propose candidate implementation 'solutions' (to real-world operational issues) for the system itself. Further use of the term solution is avoided in the sense of technical and procedural solutions to try to reduce confusion with the GSN terminology of an argumentation evidential solution described in the next subsection. Implementation strategies have been provided since the specifics of how the system will operate have not been decided at this conceptual stage and it does not make sense to take the normal GSN approach and strategize over just the argumentation of a system that has not yet been defined. Strategies that are coloured yellow are singular to the particular issues which they are trying to address. Blue-purple pastel colour has been used to denote when there is a choice of options available to address the same issue. These options might be used in combination or may be mutually exclusive. This is left to the reader to decide in each case. The next stage of evolution would require pruning the unused or eliminated strategies to define the actual system and expand the detail of what would be deployed. This strategy selection process is described in more detail in a following subsection.

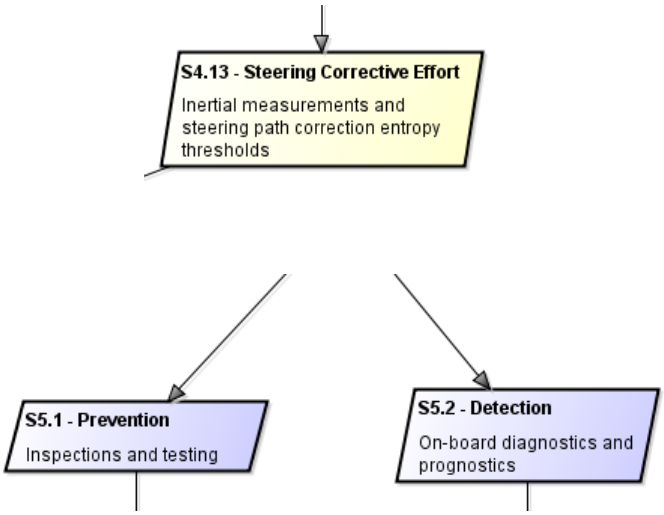


Figure 5 : An example of GSN strategies

4.2.6 Solutions

Solutions are intended to be the evidential means of proving the assertions which proceed them. They can take the form of background studies, user trials, detailed design documents, test specifications and their results or anything else which can be conceived as supporting the arguments above them. Since the GSN argument produced for this report is at the concept level, other detailed safety cases produced in GSN or otherwise could potentially be cited in their entirety, or this argument model could be refined to become those safety cases.

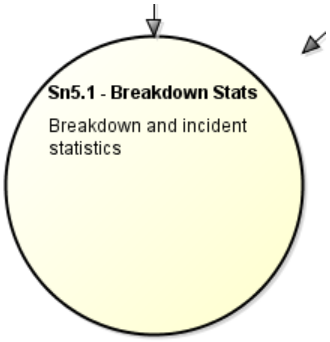


Figure 6 : An example of a solution in GSN

4.2.7 Modules and Naming

Modules have been used to separate the model into sections, each of which relate to a particular class of issue. Each of the modules have been assigned a number which is used to label the items with. This is to assist with traversing and referencing the model. In some cases sub-modules have been used to keep the module from becoming unwieldly. As a worked example to illustrate for the Highway Pilot GSN model, the question could be posed:

"*What will happen if a lane needs to be closed, how do I know it will actually work*?"

In response to this question the module on lane reallocations is found, then within the goal for lane closures G1.2 is found, assuming a V2I strategy is preferred, then this is supported by the goals "G1.4 - Lane Closure Initiation is Failsafe" and "G1.5 - Closure Notifications Are Failsafe" and their supporting argumentation and a discussion around those strategies, how they will work in practise and what proof is needed could then ensue. Conversely if the goal "G8.28 - Target Incursion Rate Met" is referenced, then from its numbering, this can be located within Module 8 for obstructions under the branch for limiting the occurrence of objects on the road to which the wider debate of prevention versus cure (detection) of people and animals roaming on highways could ensue.

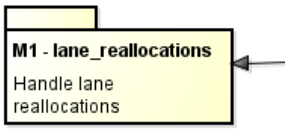


Figure 7 : An example of a GSN Module

4.3 Using the GSN to Define a Safe and Working System

The GSN is used to set out a conceivable range of candidate technical and procedural strategies. However, just because various strategies are proposed, this does not necessarily mean the authors agree that all of the strategies will be able to achieve the desired level of safety and practicality. However, only strategies that have been deemed to be worthy of inclusion have been considered, even if their inclusion is only merited to the extent that they should be formally discounted rather than just ignored. It is the job of the GSN model structure to demonstrate that particular strategies may not really work through the evidence it requires to support them, or lack thereof. The model presented requires *pruning* of the used strategies and the argumentation around each strategy is intended to help with that by showing the evidence that should be required to give confidence for each strategy.

To assist with the pruning of candidate strategies the following section of the US military standard MIL-STD-882 provides a useful statement for a design safety in section 4.3.4:

> *"The goal should always be to eliminate the hazard if possible. When a hazard cannot be eliminated, the associated risk should be reduced to the lowest acceptable level within the constraints of cost, schedule, and performance by applying the system safety design order of precedence. The system safety design order of precedence identifies alternative mitigation approaches and lists them in order of decreasing effectiveness:*
>
> *a. Eliminate hazards through design selection*
> *b. Reduce risk through design alteration*
> *c. Incorporate engineered features or devices*
> *d. Provide warning devices*
> *e. Incorporate signage, procedures, training, and Personal Protective Equipment.*

Often the simplest approach is to solve a problem such as: *"tell people not to step into the road in front of cars"*. However, the MIL standard section advises in effect to reverse that thinking and attempt where possible to eliminate the hazard through design. Only when this is not possible due to real-world constraints, then start to consider softer measures such as training and education, in effect to treat these as a last resort instead of a first option.

The candidate strategies that are proposed are to cope with the issues arising from scenarios which otherwise the automation may not cope with, resulting in hazard or just failure of proper operation, that is; failure to make sufficient progress along a journey route. Generally, there are some similar reoccurring themes depicted in the solutions. These are usually choices between:

- Solely vehicle based sensing and perception ***against*** having dedicated infrastructure which may have its own sensing or remote monitoring;
- Implementing a failsafe approach ***versus*** a sense and react approach;
- Constrain or control the chaotic elements within the environment ***or*** rely upon system control reactions;
- Rely upon the driver and occupants (when present) to assist with unusual situations ***or*** use more systematic approaches which leave less to the chance of abuse, miss-use or other human shortcomings.

Even more generally these can be further reduced to just a choice between:

- Use assured infrastructure;
- Use only vehicle systems;
- Rely on the driver/occupant to override/intervene.

The factors which would ultimately determine the solution selection in each case are a function of the cost and practicality of the 'best' solution against whether the 'lessor' solutions are still good enough. It should be remembered that some of the strategies propose the addition of infrastructure or measures which

could be reused for other solutions from which the cost and benefit would be shared which may modify the resultant practicability.

The GSN approach requires evidence that minimum standards of operation are met.

To assist with this determination, the term As Low As Reasonably Practicable (ALARP), has been used in the GSN text. ALARP is broadly equivalent to SFAIRP (So Far As Is Reasonably Practicable) and ALARA (As Low As Reasonably Achievable) that are used in some countries. However, SFAIRP could be noted for being subtly but importantly different to ALARP in that it requires a positive demonstration of due diligence over and above an outlay of capital. This can help when considering events that are said to be of high consequence but of low probability (those which often result in court cases), so possibly ALARP might be better substituted for SFAIRP, but both have the same kind of sentiment that needs to be conveyed. From an Australian regulatory perspective, they have been summarised as follows[9]:

- ALARP asks what is the risk associated with the hazard and then can that risk be made as low as reasonable practicable
- SFAIRP asks what are the available practicable precautions to deal with the identified issue and then tests which precautions are reasonable based on the common-law balance (of the significance the risk vs the effort required to reduce it)

ALARP is derived from the Health and Safety at Work Act 1974, which requires:

> *"Provision and maintenance of plant and systems of work that are, so far as is reasonably practicable, safe and without risks to health".*

ALARP supports the notion that given sufficient time, money, resources and effort, a risk could be reduced to zero, or in other words, eliminated entirely. The idea behind it is that sufficient measures should be taken to eliminate risk until the next proportionate step would be in *gross disproportion* between the costs and benefits of doing so. If the candidate strategy before the *ridiculously too expensive or impractical* option still leaves too much residual risk, then it should be considered that there is currently no viable proposed option and the system should not be deployed unless or until that changes. This report does not attempt to postulate what that residual risk level should be, except to advise caution in certain areas where the risk can be reasonably assumed to be too high. This leads us next to the need to quantify the acceptance level for risk tolerance.

4.4 Tolerance of Risk

There is also a confidence argument to be made when selecting a particular strategy or approach based upon ALARP. How good is the knowledge supporting that decision? In the hindsight of a post-accident investigation or adversarial court case, decisions to limit safety measures, testing and general effort that

[9] SFAIRP vs ALARP, Richard Robinson and Gaye Francis, R2A Due Diligence, CORE2014 Conference
http://www.r2a.com.au/wp-content/uploads/2015/02/CORE-2014-paper-SFAIRP-vs-ALARP.pdf

appear to have been made mainly on the grounds of cost can appear callous, therefore the notion of ACARP[10] should also be considered:

> *"…it is important to provide reasons not only for action but also for doing nothing. When an accident has led to an inquiry or a prosecution, inaction may, in retrospect, give the impression of negligence, and negligence may be hard to refute. There needs to be evidence, first that it was not negligence and, second, that inaction was justified. More than that, the engineer should be able to claim to have been 'as confident as reasonably practicable' (ACARP) in each ALARP decision. To increase the likelihood of favourable SFAIRP judgments, ALARP decisions should be supported by ACARP arguments."*

Much of the strategy pruning and selection exercise when adopting the GSN technique and applying to the subject of AVs is a matter of the appetite for risk and the tolerance of road fatalities and serious injuries. Society accepts that roads are a dangerous part of everyday life, which is balanced against the convenience of personal travel. However, it is unlikely to be acceptable to suffer any degradation in the safety record due to the application of automation, particularly as one of the touted advantages of automation is the reduction in road casualties with the removal of human error. Incidents due to driver inattention and misjudgement are prevalent and may be significantly reduced by automation. However, there are new risks emerging from what will appear as seemingly random *accidents* that are in fact incidents caused by the insufficiencies of automation. These may be far less palatable to society, seeming as a fait accompli caused by bad system design rather than a force majeure of circumstance. These new incidents caused by systemic failings should be considered additive to the existing incidents that take place. This leads to an important acknowledgement:

Accidents involving automation may involuntarily involve road users who would not have normally been involved in an accident through their direct actions.

It is strongly suggested by the authors that where possible failsafe measures are implemented, and where not, the residual risks are considered against the likelihood of hazard occurrence (which may have to be determined through study once the system has been partially developed) and the wider benefit to society of realising the full benefits of the feature.

Whilst the cost of a safety measure may at first seem prohibitive, it may be possible to recover the deployment costs by offsetting them from other areas such as a reduction in crash clean-up and the wider economic impact of reduced road closures and delays.

Another example is the use of virtual digital signage which may diminish the need to provide and maintain physical signage and gantries/Variable Message Signs (VMS). This may also lead to changes to signage location, and content changes becomes more trivial and cost effective with more potential for dynamic time-based changes.

[10] ALARP Explored, Felix Redmill, Newcastle University Computing Science Technical Report Series http://eprint.ncl.ac.uk/pub_details2.aspx?pub_id=161155

A report[11] from the H2020 EPCOS project that is looking at the modularisation of aviation safety cases, has the following to say about the trading of one sets of risks against another:

> *It should be noted that a change which increases safety risk in one domain is usually difficult or impractical to justify, even when it significantly decreases safety risk overall.*

A debateable ethical issue is that when applied to a feature such as Highway Pilot, a complete inability to cope with sudden rare events (e.g. an articulated lorry jack-knifing, sink holes, bridge collapses, objects dropped off bridges on to roads, or animals on the road), cannot be justified by a reduction in general crashes just because one is more prevalent than the other. Whilst it must be accepted that conventional manual driving is not without its risks, the limitations of automation in effect enters the vehicle occupants into a game of chance, which unlike general driving, they have little or no influence over once they have disengaged from the driving task. The redistribution of risk from one group of would-be manual drivers who may have been inattentive/incapacitated/incapable when faced with a high risk situation, to a new group who are selected by fate to suffer a collision, is an ethical issue which requires further thought and research.

[11] Improving European Aviation Safety Approvals, Developing Safe Systems
Proceedings of the Twenty-fourth Safety-Critical Systems Symposium, Brighton, UK
http://scsc.org.uk/paper_131/06%20Bull%20-%20Improving%20European%20Aviation%20Safety%20Approvals.pdf?pap=987

5 GSN Outputs – 'Highway Pilot' Use Case

5.1 Introduction

This section presents the results of the GSN for the Highway Pilot use case.

5.2 Overview

The Highway Pilot feature is considered for a traditional M1 vehicle that is normally manually driven. The feature can be turned on for motorways and certain dual carriageways which do not contain junctions or crossing places and which exclude non-motorway traffic. The context explains that when the feature is activated the driver should always be present and in a position ready to take back control, however not at short notice. Recent human factors research has shown that it can take between four and forty seconds for a driver to take back control after an unscheduled Human Machine Interface (HMI) handover request. Stéphane Feron, a HMI expert at PSA Peugeot Citroën, has coined the term 'drissenger' for the role of a driver who is also part passenger in a supervisory role. He describes the difficulty with determining a *sufficient time margin* as[12]:

> *The main challenge for level 3 is the take over request with a 'sufficient time margin' as the vehicle's reaction time is highly dependent of multiple factors and there is no definitive value for the 'sufficient time'.*

It is felt that most of the automotive industry is now moving towards a consensus that sudden handovers are potentially dangerous for the majority of drivers and therefore are not a realistic strategy moving forwards for increased automation. The GSN does not provide a time-based definition for how long the notice period should be, but instead this forms part of the argumentation that a demonstrably adequate notice period should be given which in turn will have scaling consequences for other system design choices. In addition, it is assumed that in some instances a handover will be impossible (even if not allowed or permitted by law) due to the driver falling asleep or becoming incapacitated for other reasons such as a medical emergency. The main outcome for the feature is that the system must be capable of deterministically handling anything that can plausibly happen to a vehicle whilst travelling under highway conditions described by, at the very least, bringing the vehicle to a safe stop in a place which does not place any road users in danger.

The unintended consequence of this seemingly small requirement change from a short to longer handover period is that the system (which includes the vehicle, its control software, and any supporting infrastructure, and procedures) must take on at least an order of magnitude more responsibility for safety and the resulting system will require far more resilience and assurance surrounding it, which generally implies more complexity.

The Highway Autopilot GSN has been divided into eight modules as shown in the table below. The following sections provide a summary of each module.

[12] http://blog.carandus.com/2016/03/the-levels-of-autonomy-for-a-car-and-hmi-according-to-peugeot/

No.	Module type	Description
HP1	Lane Reallocations	This covers lane closures (e.g. due to a stranded broken down vehicle) and changes in lane usage restrictions (bus lanes, car share lanes, hard shoulder running).
HP2	Road Etiquette	This covers allowing emergency vehicles to pass and appropriate speed regulation. It could be extended to include traffic merging and other issues which currently require a degree of discretion and common sense.
HP3	Lane Rerouting	This is for when a lane position needs to be temporarily or permanently changed, usually during and after roadworks.
HP4	Adverse Weather	Strategies to restrict the use of the feature during bad weather, particularly during the sudden onset of challenging weather when the system is already active.
HP5	Mechanical Failure	To make sure that mechanical failures do no go undetected while the feature is in use.
HP6	Intervention	When external invention is needed either to stop a vehicle or for speed reduction due to an incident or roadworks.
HP7	Operational Envelope	Ensuring the feature is only active on the roads it is intended for. It does not cover other aspects of the control envelope such as stability and speed regulation.
HP8	Obstructions	Strategies for handling collisions with static and dynamic obstructions in the road.

Table 4 : Highway Pilot module type and short description

The structure of the Highway Pilot GSN is confirmed as shown in Figure 8:

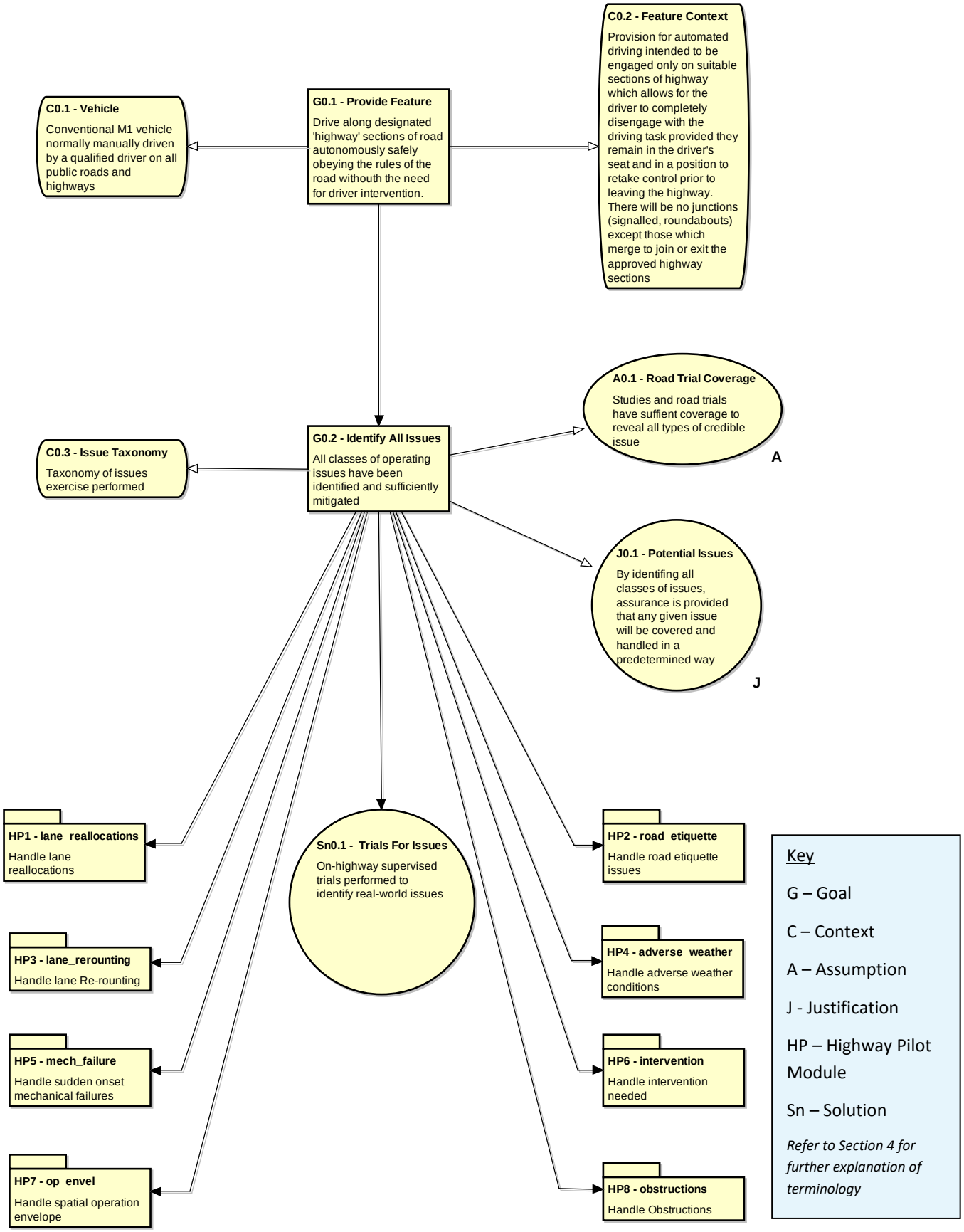


Figure 8 : Highway Pilot GSN – Top Level Structure

5.3 Lane Re-allocations (HP1)

This module looks at lane closures and when the permitted use of a lane changes with time, based upon vehicle type or occupancy. These may seem very different things but they both involve a sudden change in whether it is acceptable to proceed in the current lane.

Lane closures make a lane unusable to traffic and can be needed with sudden and immediate effect, as well as for planned activities such as road maintenance. Currently a lane could be closed via a variety of means, from overhead gantry VMS, or more physical means such as cones placed on the carriageway, or a large strobe lit mobile signs mounted on highway maintenance vehicles. Lane closures need to be patently obvious to drivers and the consequences of remaining in the closed lane at full highway speed can be very high. There is some inevitable overlap with lane rerouting dealt with in another module, however a distinction has been made between a lane that ceases to exist or be open and a lane which is gradually shifted laterally to a new position. When considering how this should be done systematically without the direct assistance of a driver (which must be discounted due to the feature context allowing for that driver may fall asleep even if not allowed to do so) it quickly becomes challenging once the consequences of failure are imagined. It can be divided into a matter of directly sensing the environment or providing some external means which equates to some form of infrastructure. Visual perception can always fail, so merely sensing cones on the road will not be sufficient in most cases without some other protection if the vehicle penetrates the coned area. Having the maintenance lorry in-situ is a plausible alternative since it provides a very recognisable visual reference, and if that fails then it is in effect a large physical metal barrier which is easily detectable by conventional on-vehicle (Adaptive Cruise Control (ACC) enabling) radar systems. In this case the vehicle should at the very least stop or change lane even if it fails to provide a pleasant experience for the vehicle occupants. There may be conditions when this logic is defeated; what happens after the lorry has been overtaken and there are only cones to indicate a lane should not be used? Further consideration would need to be given to any additional complexities resulting in consequential hazards once the vehicle has stopped such as the reaction of following vehicles and the likelihood of them suffering the same issue. However, it is unfortunately (probably) impractical to deploy a lorry to the point of every lane closure with almost immediate effect, so another solution must be considered.

Mapping providers are keen to present map based solutions which often involve what have become termed Local Dynamic Maps (LDMs). LDMs capture dynamic and transient information about changes to the road which once detected are then disseminated to other vehicles using V2V or V2I. The idea is that once a suitably equipped vehicle encounters a change, such as road works or a stranded vehicle, that information can then be given to following vehicles to allow them to take appropriate evasive action. The issue is that this approach leaves many things to chance, and assumes that the first vehicle to encounter the lane closure can itself detect the problem and behave appropriately, so in effect this does not solve the problem beyond providing non-dependable advisory to other vehicles. However, the general notation of having a dynamically updated map may be a valid one provided it can be made failsafe. Following this line of thought leads to an infrastructure approach with some form of synchronous handshaking such that it will be possible to at least know when the latest information may not have been received. Wireless communications with infrastructure (e.g. V2I) must be assumed to be able to fail, however unlike perception (where it is possible to not know what you don't know) if a transmission is expected, based on time and/or location, but not received then the system will know it is at risk of there being an unreported

change. With this knowledge, preventative action can be take such as attempting a handover to the driver, or when that fails, attempting to stop in a safe location. This in turn requires that the data loss is detectable far enough in advance to safely stop before encountering the lane potential lane closure, but these details can be resolved as part of the requirement cascade of the system design and the definition of its constraints from which things like optimum safe harbour spacing and capacity can be determined. With safety ensured, secondary issues such as maintaining communications minimum Quality of Service (QoS) at peak times requires consideration.

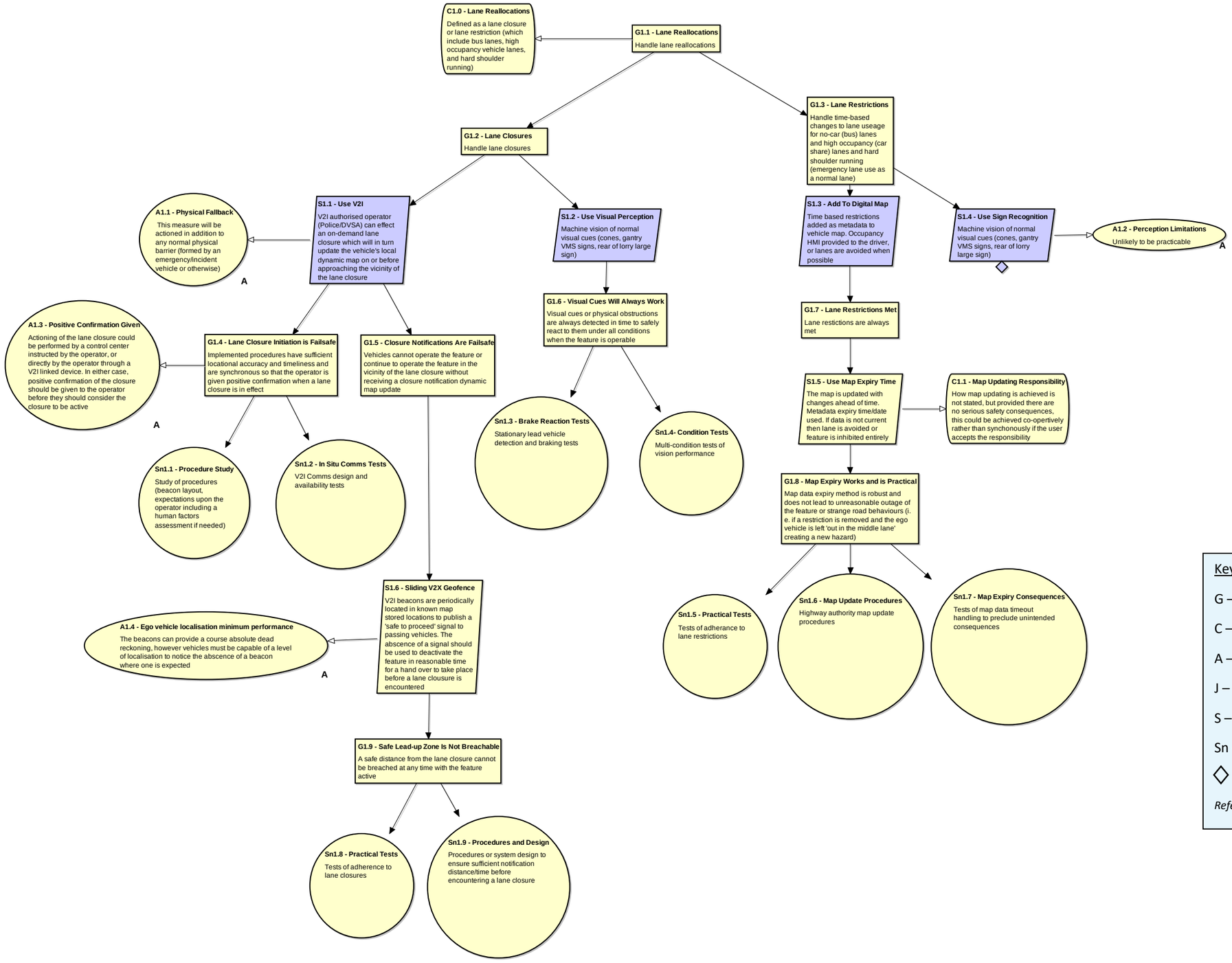
In summary, what this amounts to is having either

- *safe to proceed* radio beacons which are noticed in their absence;
- **or** a dynamic map layer with a short expiry time.

Tests would then be needed to ensure that both conceptually and in practise the protection system can never be defeated, to a level of confidence which is commensurate with the risk of (potentially fatally) harming road workers, and people in stranded vehicles are well as the occupants of the vehicle hosting the feature.

The next consideration within this module is when the allowed use of lane changes with time such as for bus lanes, high occupancy lanes or hard shoulder running. The means to handle lane use changes divide in a similar way to lane closures. Machine vision of conventional road signs could be relied upon (requiring optical character recognition), but this is likely to be prone to failure. A map based solution is again a natural solution as time restrictions could be added as easily as speed limits to existing maps as an attribute. The difficulty arises at the point where changes are made to the lane regulations. This can be managed using map expiry times and aligning regulatory changes to map expiry times to ensure the changes are fully propagated. A map update failsafe might not be needed in this case (perhaps with the exception of hard shoulder running) but if it has been implemented for other reasons it could be reused to prevent the feature being used without the latest information for the intended route.

The Highway Pilot Lane Reallocations module is shown in Figure 9:



Key

G – Goal

C – Context

A – Assumption

J – Justification

S – Strategy (yellow = singular, blue = one of several options)

Sn – Solution

◇ Undeveloped

Refer to Section 4 for further explanation of terminology

Figure 9 : Highway Pilot GSN – Lane Reallocations Module (HP1)

5.4 Road Etiquette (HP2)

This module has been used to address emergency vehicle procedures and the normal regulation of vehicle speed.

5.4.1 Emergency Vehicle Passage

Allowing emergency vehicles under blue light operation to pass is currently handled somewhat awkwardly by current practise as different drivers will react differently and with the best of intentions. Due to a lack of overall coordination this can lead to emergent situations that can make the progression of the emergency vehicle worse rather than better. Many drivers in an urban setting choose to mount the footway or on a highway (through an emergent consensus) may enter the central reservation zone to help provide a central corridor for the responding vehicle to pass through. This may be technically illegal, or a breach of the rules of the road, but it is unlikely that any court would see it as being in the public interest to prosecute. The Highway Code Rule 219 gives further guidance about predicting the path of the emergency vehicle and pulling over to the side of the road without breaking any rules of the road or endangering other road users.

How this is handled does probably not have any direct safety implications for road users, but there are indirect implications if for instance the ambulance is delayed from arriving at the scene of a person with a medical emergency which could mean life or death for that individual. The lack of an explicit set of rules presents a difficulty in writing software to take action to let emergency vehicles pass. It could be left to the discretion of the driver to take back control or provide instruction through a user interface, or if that is ruled out then a systematic approach is needed. The system can be left to 'notice' the situation either though the normal cues which a human driver uses, or directly through a wireless V2I / V2V notification. Once the approaching vehicle has been recognised, it could still be left to the driver to take action once prompted, or further left to the system to decide what action to take. The latter option is difficult to develop without refinement to the Highway Code or better definition of what the procedure should be. Scenarios can be enacted under controlled conditions to develop and test the resulting system.

5.4.2 Speed Regulation

Speed regulation is a seemingly obvious and necessary condition for an AV to function, but one which has not received much attention in open forums. Not exceeding the speed limit is probably a given, and something that the DfT has stated as an expectation of AVs, but the speed limit is just that, a limit, not a target speed appropriate for all conditions and sections of a road. Driving at or below the speed limit at a rate which ensures passenger comfort and vehicle stability seems to be the natural solution. How this is determined is less obvious. Setting a target speed for normal conditions which could be applied to a digital map as an attribute is one strategy, or varying the speed limit could be another. A vehicle specific offset could then be applied to the target speed to allow for model specific capabilities and driver/user preferences. Handling larger speed offsets for degraded weather conditions presents a more challenging issue. This could be left to the individual system implementers to decide how to sense conditions and apply offsets, but the burden of proof that their methodology works should be maintained prior to the deployment of the vehicles rather than left to a future court case after an incident.

It should also be noted that current Electronic Stability Control (ESC) systems are provided to assist the driver in not losing control of the vehicle and are not an absolute guarantee of stability under all conditions. Continental state[13]:

> *Without enough friction between tyres and road (good "grip"), there is no way of keeping a vehicle under control. Active safety systems such as ABS and ESC will help in situations when emergency braking or imminent skidding put a driver's skill to the ultimate test. However, even such valuable systems as ABS and ESC will only react, when the vehicle is already in danger of getting out of control. Realistically it is entirely up to the driver to estimate whether there is enough grip to allow for the vehicle's current speed.*

Once the driver is removed from the loop, the bar is substantially raised for stability to be guaranteed, which is quite difficult to achieve for all road conditions. Driving at a suitably slow speed such that the vehicle is not close to the limits of its traction is one way, but can be a self-defeating argument since the limit of traction is not always known in advance and its estimation is a reactive part of the stability control system, so the 'suitably safe' speed is not always known in advance of a problem. If transitioning from a salted section of road to a section of snow covered road, then this estimation may be incorrect until instability (wheel slip or yawing) is sensed. The European Sixth Framework Programme Project FP6-2004-IST-4 called FRICTI@N has made some progress with improved road surface friction coefficient estimation[14]. However, although this improves greatly upon what is currently in production, it still only offers measures which are reactive to the road surface which are currently 'under foot' to the vehicle with no anticipatory ability of what is ahead to allow pre-emptive slowing down before hitting a patch of ice.

The Highway Pilot Road Etiquette module is shown in *Figure 10*:

[13] http://www.continental-corporation.com/www/pressportal_com_en/themes/press_releases/3_automotive_group/interior/press_releases/pr_2010_10_12_sensorfusion_en.html

[14] FRICTI@N Deliverable 13 - Final Report: http://friction.vtt.fi/FRICTION_FinalReport_D13.pdf

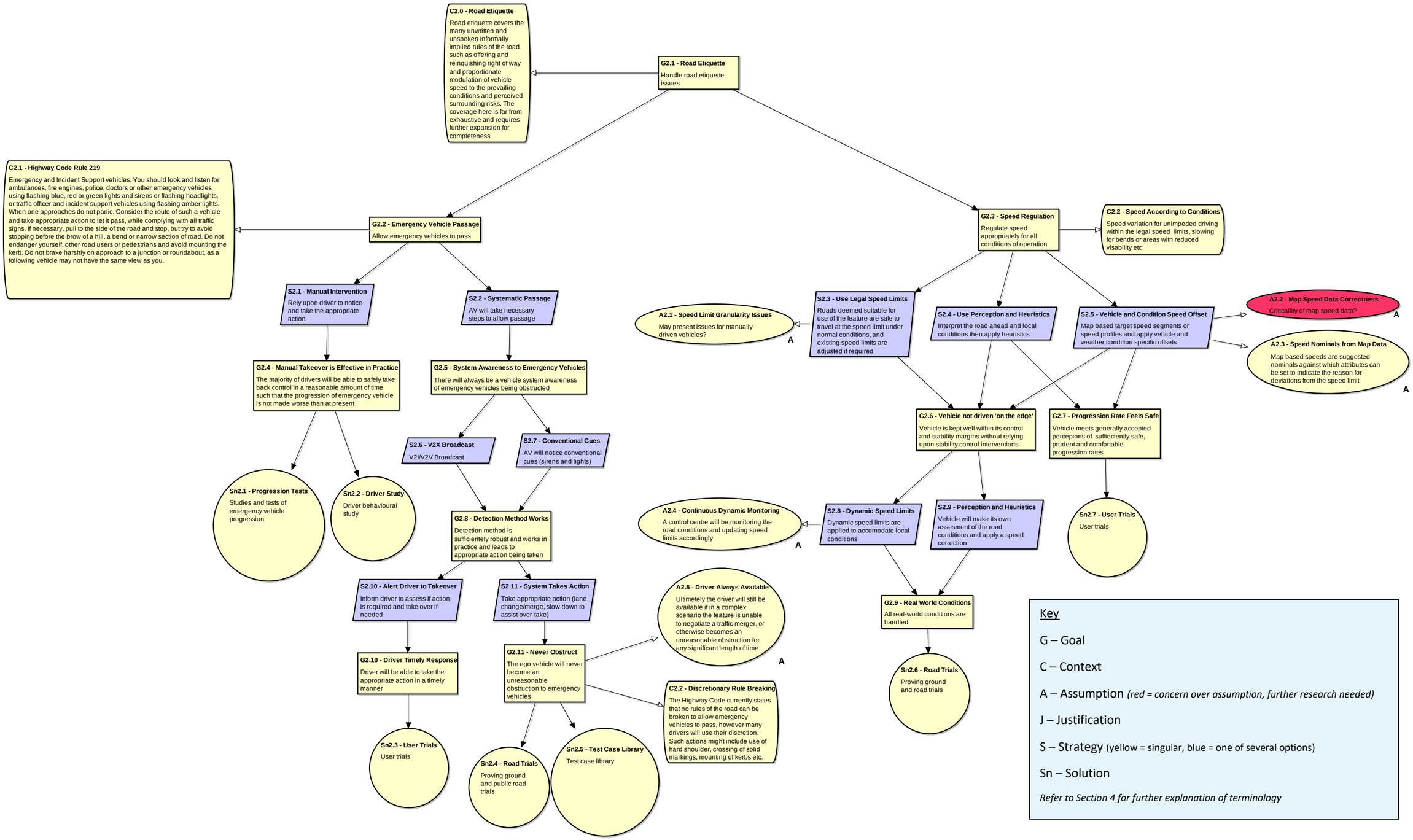


Figure 10 : Highway Pilot GSN – Road Etiquette Module (HP2)

5.5 Lane Rerouting (HP3)

Lane rerouting is when a lane position needs to be temporarily or permanently changed, usually in and around roadworks. How this affects the vehicle depends largely on how the vehicle normally tracks its lane position. Currently two methods are prevalent firstly the use of image recognition of the lane marking lines as is currently used for ADAS features such as Lane Keep Assist (LKA) and Lane Departure Warning (LDW). This can work well enough as a driver advisory on highway sections of road since the lane markings are usually well maintained and clearly visible due to sufficient vehicle separation (from the required braking distances). When traffic queues form, the traffic tends to close up and the lines are not always visible. Tracking the vehicle ahead under these conditions is a partial solution which is used, but that assumes that the vehicle ahead does not have the same problem and is following the same route. There have consequentially been reports of vehicles unintentionally and unexpectedly following the lead vehicle off the highway onto another road due to this deficiency.

A second approach is to use medium to high fidelity digital maps and by locating and orientating the vehicle on the map to a high degree of precision, the vehicle is kept in the real-world lane by driving in a virtual lane on a map. Obviously, this relies on ensuring the virtual lane on the map is kept in alignment with the physical one on the road. This has placed a high degree of responsibility on the map being up-to-date and accurate as well as the real-time positioning on the map being equally up-to-date and precise. Whether this can be done to a level of confidence high enough to be assured for high speed driving remains an open issue. Failsafe mapping and localisation interlock are both required unless substantial risks to all road users are to be tolerated. They don't need to always work, but the system does need to know when the confidence in either has dropped over time in order to take action such as handing over control to the driver or stopping the vehicle somewhere safe. If the safe stopping location is in question, as a consequence of map obsolescence say, then knowing where a safe location is may become a self-defeating argument without a fall-back procedure.

Similar arguments to lane reallocations can be made using maps with expiry times or safe-to-proceed wireless beacons which can provide a basic level of localisation coupled with the assurance that the latest map information is being applied. It could be argued that precise localisation is more critical for keeping in lane than for lane closures as these can be set further ahead of physical obstructions than the width of a lane. However, it is unlikely that one would be needed without the other, this is if localisation is needed to know where a lane closure is then it will also be needed for lane locations and hence the latter sets the requirement for localisation performance. Many different localisation strategies are being proposed by researchers and technology companies. However, no clear candidate has emerged so far. Rather, it may be being assumed that some fusion of all the available techniques will provide full coverage. This may not be the case and it may need to be explicitly resolved such that there are no coverage gaps that could result in a perfect storm. This should be done by design and not just left to a *'try it and see'* approach to testing.

The Highway Pilot Lane Rerouted Lanes module is shown in Figure 11:

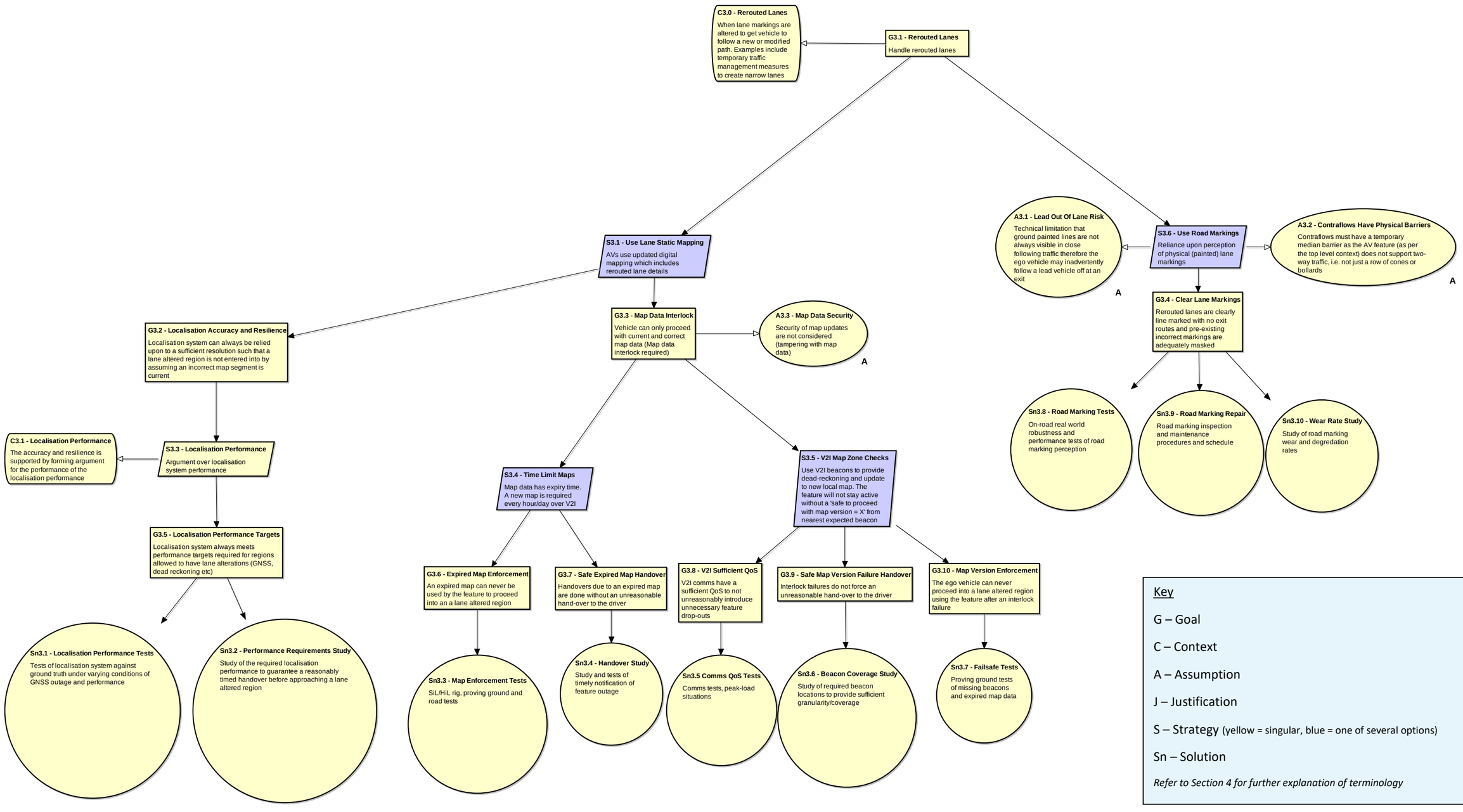


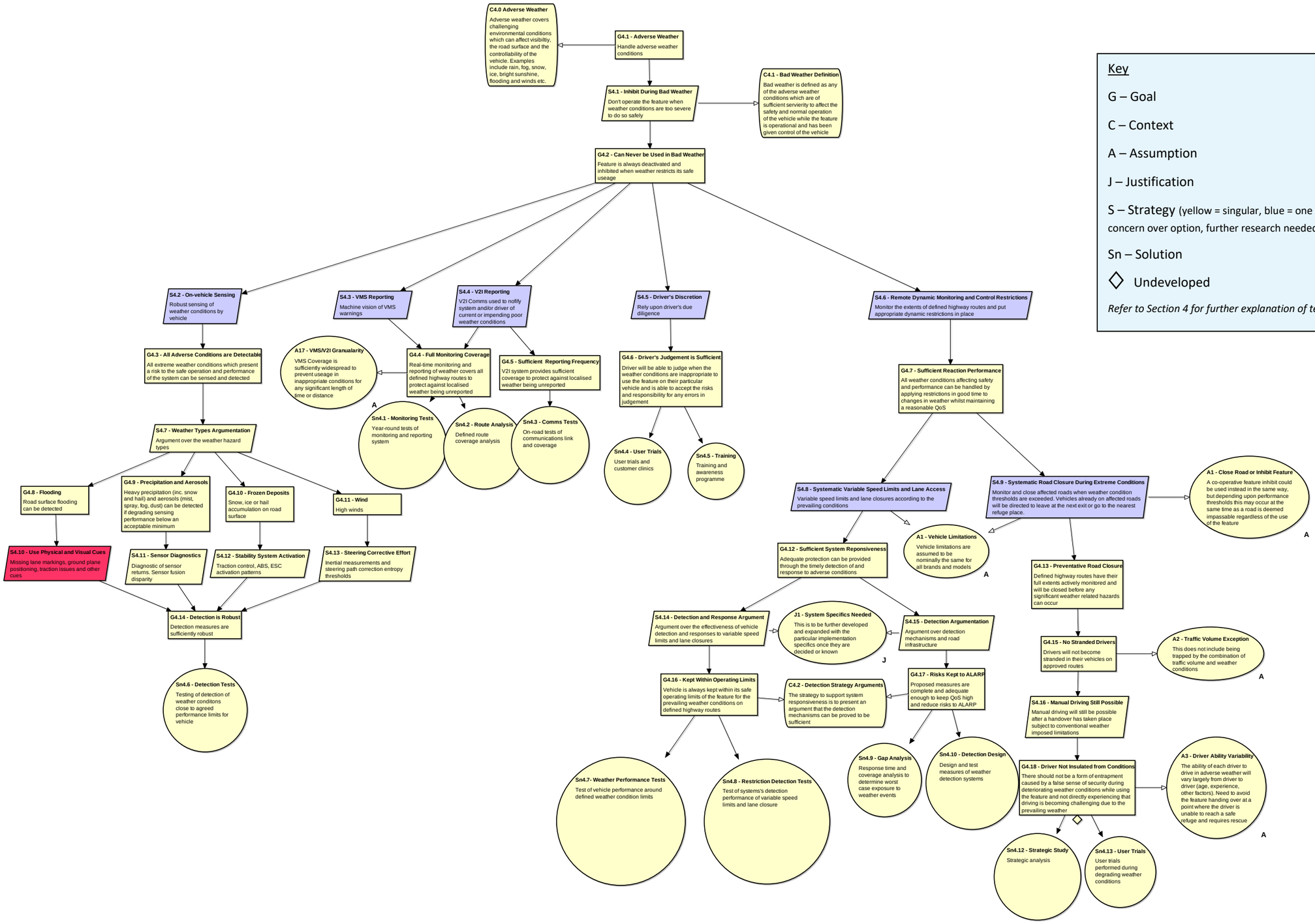
Figure 11 : Highway Pilot GSN – Rerouted Lanes Module (HP3)

5.6 Adverse Weather (HP4)

The basic principal is to inhibit the feature during bad weather. The focus in this module is on the ability of the system to continue sensing and understanding its environment rather than keeping control of the vehicle due to poor traction, but the containment measures could be similar for both cases. What *bad weather* actually means and how bad it has to be depends upon the capability of the system to cope with various conditions that may affect its performance, particularly during safety critical situations. It might be conceptually simple just to prevent the feature being turned on during bad weather, but consideration is needed for when the feature is already in use and the weather starts to suddenly deteriorate. Further to this, it should not be forgotten that the feature context provides for the fact that the driver may have fallen asleep or have become incapacitated whilst the feature is active, so a handover may not be possible and just stopping might be particularly dangerous especially in road conditions where there is reduced visibility and reduced traction.

Again, the choices emerge for preventive measures (weather forecasts), relying on the driver's discretion, or a fully systematic approach whether using a standalone infrastructure support method. Part of the problem rests with detection. Infrastructure monitoring and support would be one way, with the benefit of being able to see the conditions far ahead in time to stop or handover gracefully, but there is also the risk that very localised weather events go undetected. If the driver retains responsibility and remains alert and awake, they will still have been physically disconnected from driving the vehicle so their proprioceptive feedback from the vehicle controls is not in place to know that conditions are making driving difficult, and they may also not be able to judge when the system is or is not able to cope with the prevailing conditions. Some people would be cautious and others would use their desire to arrive at their destination in good time as cause to just take a chance and proceed with their journey. Prosecuting after the fact may not be helpful or enough of a deterrent given the potential for uncertainty around the decision to stop the feature and the burden of proof with determining who/what was driving. Fully systematic detection and reaction for all plausible adverse weather conditions is quite a challenge, but if it can be achieved there are still practical considerations to be made. There is the possibility of stranding the driver (and passengers) on the road in a safe harbour or in remote areas in poor weather conditions, if the feature has been permitted to drive them into conditions that are beyond their own ability to continue driving, and may require their subsequent rescue, so suitable provisions or warnings may need to be made.

The Highway Pilot Adverse Weather module is shown in Figure 12:



Key

G – Goal

C – Context

A – Assumption

J – Justification

S – Strategy (yellow = singular, blue = one of several options, red = concern over option, further research needed)

Sn – Solution

◇ Undeveloped

Refer to Section 4 for further explanation of terminology

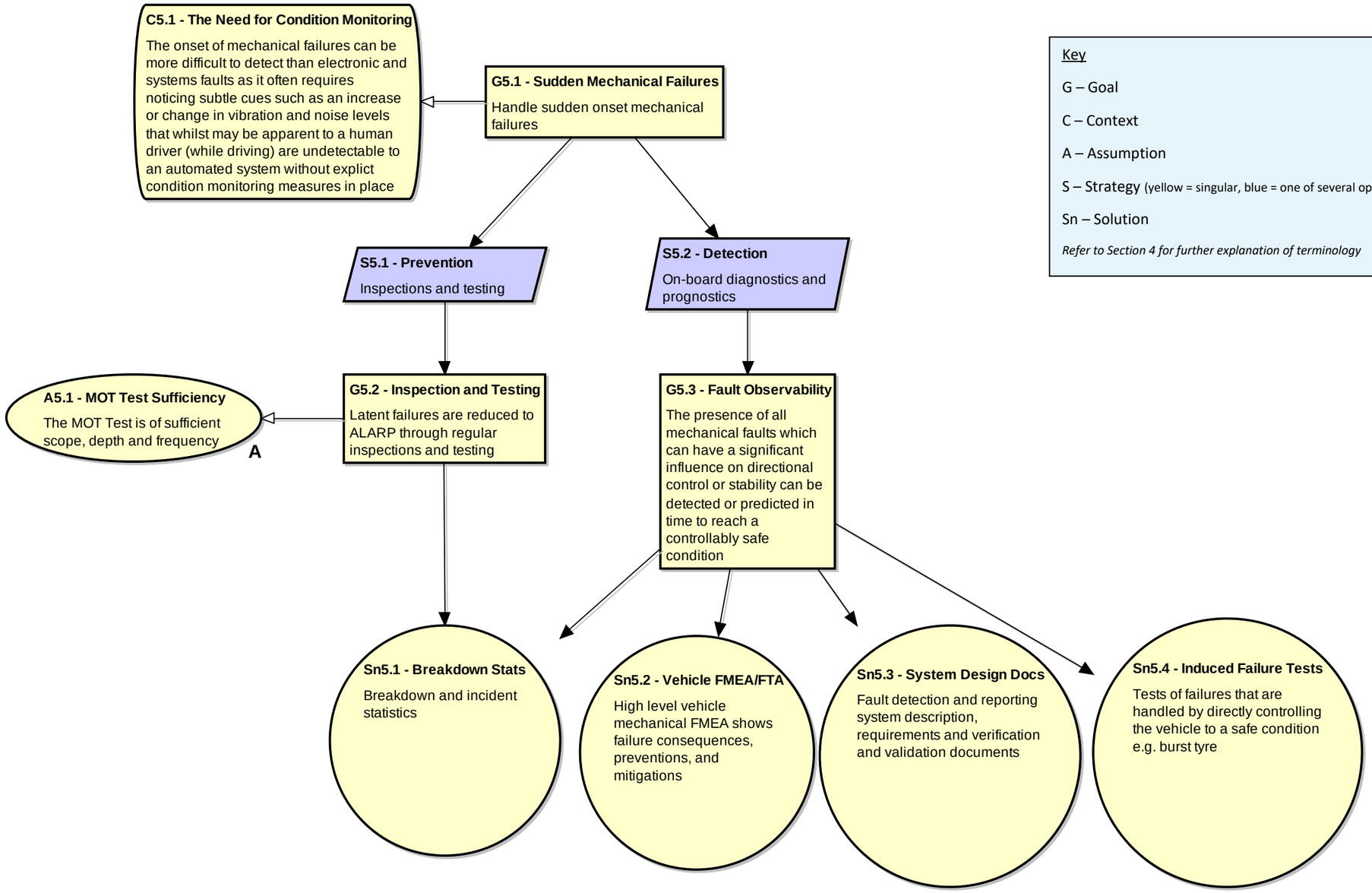
Figure 12 : Highway Pilot GSN – Adverse Weather Module (HP4)

5.7 Sudden Mechanical Failure (HP5)

This module deals with mechanical failures. The reason for limiting to mechanical failures is that electronic and software failures are dealt with during the development and testing of the system. The base vehicle is otherwise assumed to be well maintained and not an issue which may not always be the case. As with the rest of the vehicle, existing vehicles have been developed with the reasonable assumption that there is a driver in control and mechanical failure modes analysis will concentrate on being able to bring the car safely to rest in the event of a component failure. It will also be assumed that the driver will notice the failure from degraded performance or increased noise/vibration, strange handing etc. If the driver is not physically driving anymore for extended periods of time, then a mechanical fault could develop and go unnoticed, not least due to the lack of proprioceptive feedback to the driver when disconnected from the main driving controls.

What the module aims to achieve is to eliminate failures that could be dangerous and go unnoticed to both the driver and the system when the system is in control. Vehicles currently have limited or no ability to self-diagnose or perform pre-failure prognostics of mechanical faults, with tyre pressure monitoring being the only possible exception. Vehicles are generally highly reliable compared to previous decades, so there may not necessarily need to be any action taken, but it should also not be left to chance and component failure rates should be reconsidered against the automation system's ability to recognise them and respond appropriately when needed. There are two basic approaches, the first is simply mitigation by regular inspections, and the second is to increase/improve the on-board detection capabilities, which may require additional instrumentation on the vehicle which is not currently needed. Taking the first approach, the current MOT check may still be adequate, but the MOT interval may need adjusting and depending upon the analysis this could be extended in the extreme all the way to something equivalent to aircraft pilot pre-flight checks. In practise, expecting people to perform their own checks may be unrealistic, even if they were as trivial as checking tyre pressures once per day. Competency would inevitably lead to neglect of these checks in everyday life, but it may also lead to a market pressure for vehicles to be sold which can self-check to a higher degree. Fleet owned and operated vehicles could be checked as part of the operational procedures as part of an operating license, but for this would be more appropriate and applicable to a run empty service such as the Urban Pilot use case described in section 6.

The Highway Pilot Sudden Mechanical Failures module is shown in Figure 13:



Key

G – Goal

C – Context

A – Assumption

S – Strategy (yellow = singular, blue = one of several options

Sn – Solution

Refer to Section 4 for further explanation of terminology

Figure 13 : Highway Pilot GSN – Sudden Mechanical Failures Module (HP5)

5.8 Intervention (HP6)

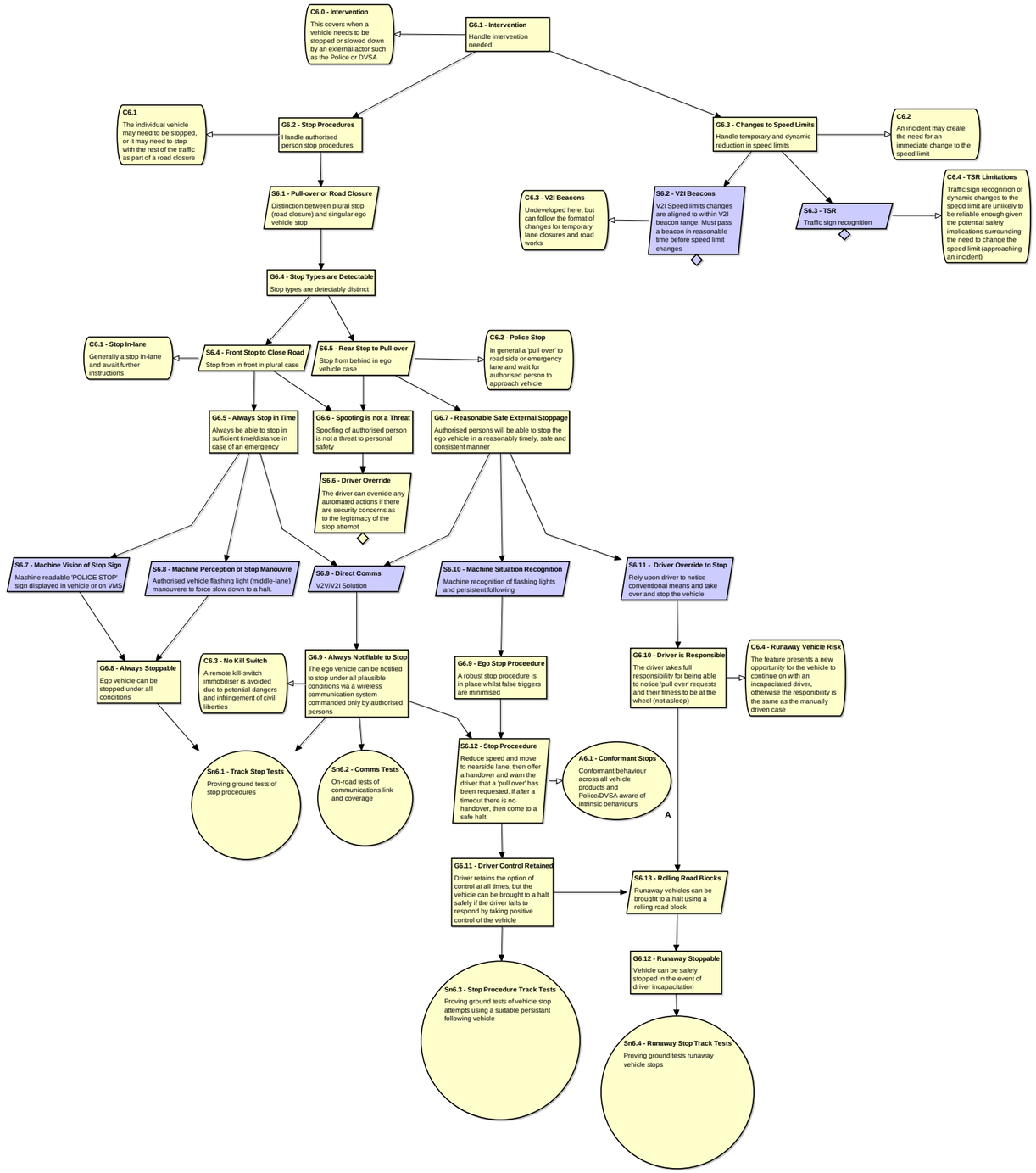
There will at times be a need for external intervention to change the speed of a vehicle (for roadworks) or to bring it to a complete stop. Dynamic changes to speed limits can be dealt with in a similar way to lane closures and lane rerouting, whether by visual perception or through an infrastructure means.

A vehicle may need to stop as part of a road block for all vehicles or as part of Police and Driver and Vehicle Standards Agency (DVSA) stop procedures for individual vehicles. An assumption has been made that a rolling road block or thereafter a full road block will be initiated from in front of the vehicle. Conversely, in the first instance, individual stops will be initiated from behind by a following vehicle. For road blocks, it could be treated as a multi-lane closure for approaching vehicles, but it will first need to be initiated to break the existing flow of traffic using a stop manoeuvre. When the stop manoeuvre is performed, it will be up to the vehicle systems to recognise this by some means.

Stopping just one vehicle is more challenging than an indiscriminate road block as even if some remote mechanism is provided other than the usual visual signals (flashing of lights or a Police Stop sign), the initiator (Police/DVSA) still needs to identify the particular vehicle they wish to halt either via its location or registration or some other means. Authentication to prevent malicious spoofing of the stop mechanism remains a concern, particularly for visual signs which could easily be replicated, remembering that just those signs and not a convincing replica of a full Police car would be needed to trick the system in to stopping. For this reason, the author suggests that a guiding principal should be to leave ultimate control with the driver as far as possible to allay people's concerns where law enforcement could just incapacitate any vehicle on a whim. The converse situation where the driver may become incapacitated or somehow trapped in a moving vehicle needs to be addressed, along with the genuine occurrences of uncooperative criminals who need to be stopped for law enforcement and public safety reasons. A set of strategies are required which cater for these opposing concerns.

The decision to stop could be left as it currently is: entirely with the driver. However, due to the previously mentioned concerns about a sleeping, otherwise incapacitated, or just unobservant driver, there still needs to be a way to stop a would-be 'runaway' vehicle. The vehicle could be stopped in theory by just driving in front of it and slowing down, but there is always the risk it may just overtake, so two Police vehicles may then be needed to box it in. A better solution is for the system to be notified of the stop 'request' and try to inform the driver via a user interface. If the driver remains unresponsive, then the vehicle will attempt a progressive stop which can be overridden by the driver at any time. A driver who refuses to stop is treated in the same way as existing drivers who fail to stop. Provided a reliable means to initiate a stop request is provided with full coverage on roads where the feature is to be used, then the needs of all stakeholder should be met by following this approach.

The Highway Pilot Intervention module is shown in Figure 14:



Key

G – Goal

C – Context

A – Assumption

J – Justification

S – Strategy (yellow = singular, blue = one of several options)

Sn – Solution

◇ Undeveloped

Refer to Section 4 for further explanation of terminology

Figure 14 : Highway Pilot GSN – Intervention Module (HP6)

5.9 Operational Envelope (HP7)

The operational envelope refers to the spatial (geographic) envelope and does not extend to cover all operating conditions which are covered by other modules (such as adverse weather). The vehicle must handle anything which can reasonably occur within an operating area which it has been designed for.

As with other modules, there is the high-level choice between relying upon the driver to only use the feature where it is appropriate to do so or using a systematic approach (in effect some form of geo-fencing[15]). If geo-fencing is required, then it must be relied upon in proportion with the potential risks involved if the vehicle was to leave the intended highway and find its own way onto minor or unapproved roads. An example of this might be if the feature was to be used on a section of unsuitable dual carriageway for which it may appear to function correctly to the user until a junction or other limiting factor is encountered with potentially disastrous consequences.

When the journey route meets a natural boundary such as needing to leave on a junction slip road (off ramp) or the end of a motorway then thought needs to be given to what will happen at that point. The natural course of action will be to offer to hand over to the driver (either rolling or stationary), but more thought needs to be given to when the handover is not accepted. If the vehicle simply stops at a threshold (roundabout give way line at the bottom of a slip road, or signalled junction at the end of a motorway) then this can present its own danger. The vehicle could continue past the junction until the end of the motorway is reached or the vehicle runs out of fuel, but this defers the problem rather than solves it. Using safe harbours is another approach, but it may not be feasible to add safe harbours everywhere where they might be needed.

The Highway Pilot Operational Envelope module is shown in Figure 15:

[15] A geo-fence is a virtual perimeter for a real-world geographic area.

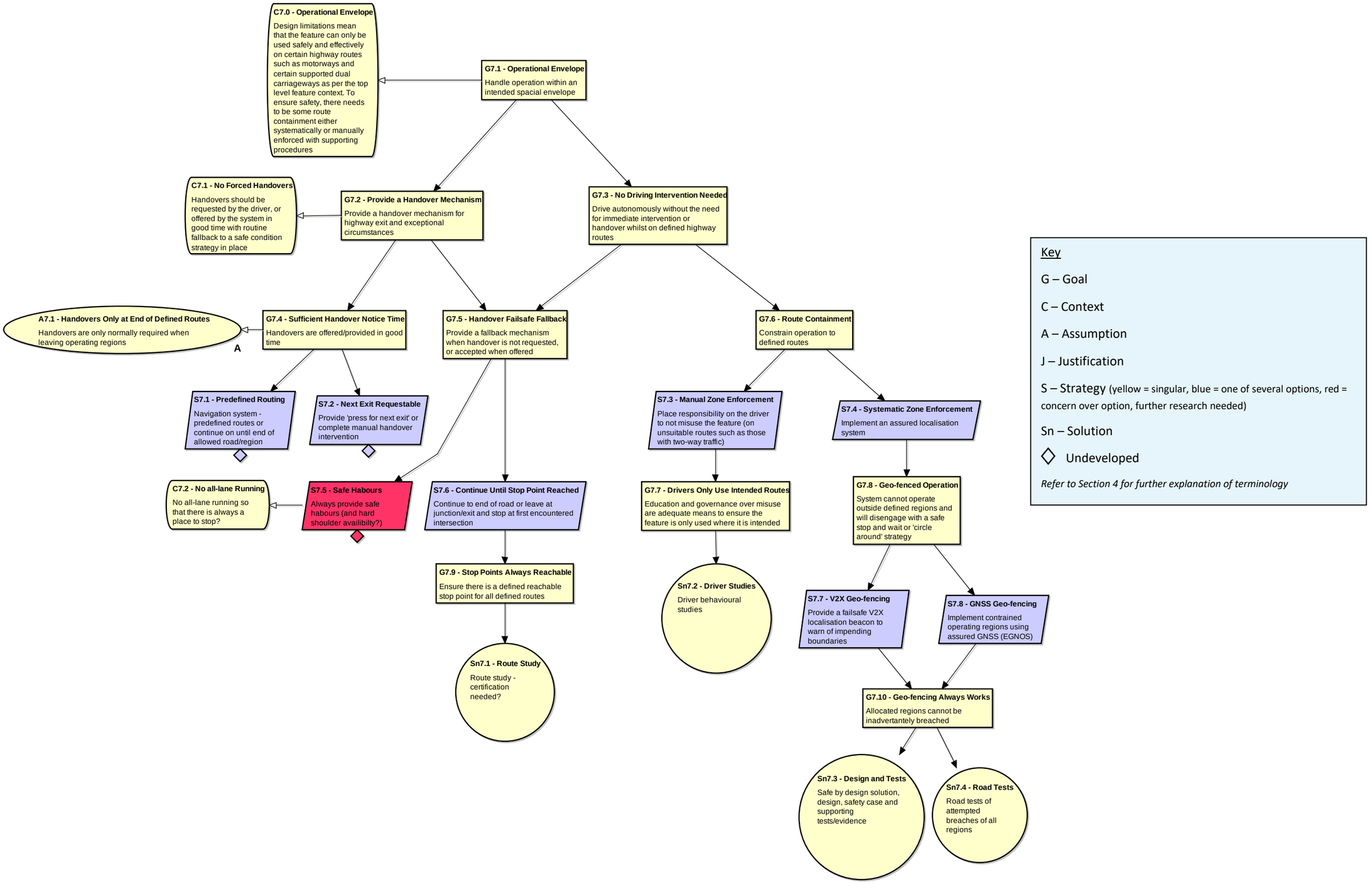


Figure 15 : Highway Pilot GSN – Operational Envelope Module (HP7)

5.10 Obstructions (HP8)

For a feature with the remit of Highway Pilot, obstructions on the road are probably the greatest challenge. Motorways are the most regulated of roads, having exclusions for pedestrians, cyclists, horse riders and slow moving vehicles. This provides for a more controlled environment to operate the feature within. However, the difficulty emerges when considering breaches to these regulations and other anomalies such as goods vehicles shedding their loads or objects falling from bridges over the carriageway. Human drivers are not impervious to the effects of these rare events, but they do have the benefit of perception and recognition of unusual situations as they emerge. These situations are of low probability but high consequence and need to be addressed so that they are handled as well or better than a human driver (which in itself needs further definition) or they are prevented from occurring when it is known that the system will not be able to respond appropriately to them. Prevention measures (such as fencing and monitoring) will still leave a residual risk, but this risk can be set low enough to be palatable to society.

The approach taken has been to divide the argument into static and dynamic obstructions. Objects falling from a lorry are initially dynamic obstacles but once they come to rest will transition into a set of static obstructions, which in time could be handled by remotely closing the affected lanes. It is the interim response period between the load leaving the lorry and the potential for impending collisions that needs to be addressed. Vehicles can be prevented from driving into objects by directly sensing the obstruction. We cannot rely upon any classification of the object since it is possible for almost any object to end up in the road. Generally, it could be argued that static obstructions would be physically connected with the ground plane (road) so that anything which is touching the road which should not be there is a threat. However, the logic starts to fall down when considering items such as plastic carrier bags which may or may not be on the ground and may or may not be stationary.

For low density objects (e.g. plastic carrier bags and soft packaging) it is not normally sensible to brake abruptly or at all. Sudden braking for phantom objects is in its self a collision risk and may unreasonably disrupt normal traffic flow. A measure of density is needed to detect these low risk objects in the absence of human contextual understanding. A sensor which can measure object density of a wide variety of materials in front of a moving vehicle is not currently feasible. With object classification using image recognition techniques, receiver operating characteristics come into play, particularly when trying to decide if any random object presented in the path of a vehicle is something to brake for or not. For some situations, the risk of false braking events may be viewed as being too high, whilst for others not braking when needed is equally serious and it may not be possible to satisfy both conditions at the same time. This is where preventative action is preferred. If many classes of objects can be eliminated by monitoring and governance, the remaining less plausible occurrences could be more easily met by a conservative braking strategy i.e if in doubt then brake. As market penetration increases for AVs then the risks from braking should reduce as result of improved reaction times and advanced warning with back propagation of a vehicle braking far ahead.

Dynamic obstacles represent even more of a challenge, particularly in the case of livestock or animals in general. The fact that they are moving implies they have a trajectory and hence this needs to be known and predicted in order to avoid a collision. This maybe intuitive for a human, but is challenging for systematic prediction since it first involves classifying the object and understanding what it is, then deciding what action may be needed. The combination of rarity, with the difficulty in classification and

prediction at enough distance to take precautionary action means that there is a reasonable chance that any measures developed will not work as intended at the point in time when they are needed. This could be accepted as another risk of automation that must be accepted, but a study should be performed to check that the actual likelihood of occurrence is sufficiently low and that the reaction of the system is known and deterministic as far as possible so that a potentially dangerous situation does not translate into tragedy.

Overall whilst such events as collapsed bridges and sink holes are extremely rare, they will be perceived as being obvious to a human driver and so the automation system should not *plough* into a rare or unusual but avoidable hazard due to sensing or perception limitations that could have been avoided if a human had been driving.

The Highway Pilot Obstructions module is shown in Figure 16:

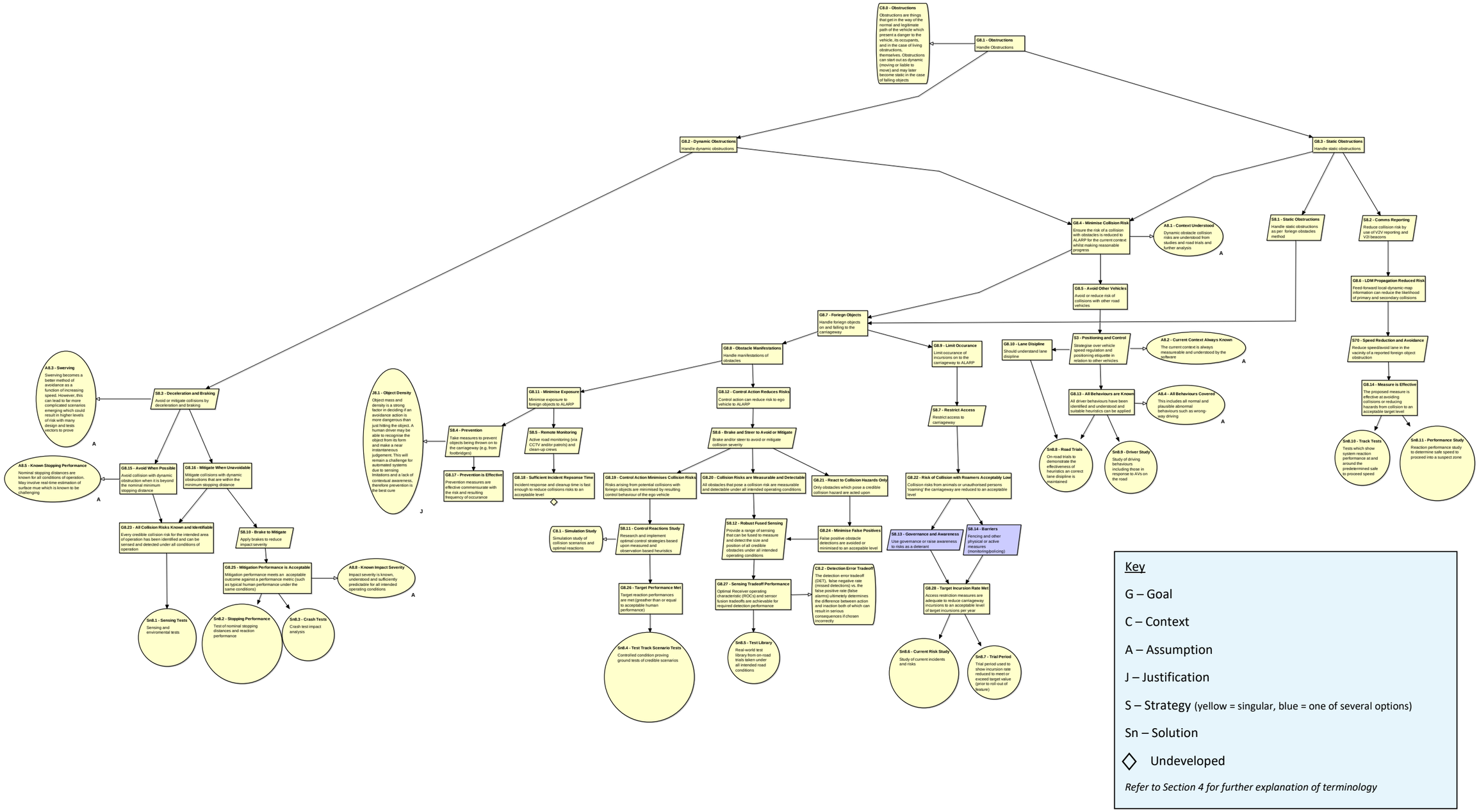


Figure 16 : Highway Pilot GSN – Obstructions Module (HP8)

6 GSN Discussion of Outputs – 'Urban Pilot' Use Case

6.1 Introduction

This section presents the results of the GSN for the 'Urban Pilot' use case.

6.2 Overview

The Urban Pilot use case is considered as an M1 vehicle with the modification that it has **no provision to be manually driven**, and therefore handover to a human driver within the vehicle is not an option. The vehicle is to be operated in designated urban zones up to 30 mph either as part of a managed fleet or by individual owners. The significant difference over Highway Pilot, other than the removal of manual controls, is the ability to run without occupants. Despite being slower, urban environments are inherently more chaotic and feature two-way traffic with a much broader mix of road user types. This places considerable additional complexity on the Urban Pilot compared to the Highway Pilot. Much of urban road travel requires ad-hoc arbitration between individual drivers in addition to the need to negotiate junctions and crossings of various types. By its nature this represents a significant challenge to automate to a level which requires no direct supervision. True autonomy for such a vehicle is difficult to define never mind to realise and as such what inevitably must be considered is a higher level of automation with some accepted limitations. These limitations can be met by a variety of means, from a simplification of the environment using new rules of the road (e.g.no jaywalking), an acceptance of new types of accidents (possibly traded against a reduction of inattentive driver related accidents), through to supporting infrastructure to widen the perception of the vehicle beyond what is possible for it to sense from its own location and sensor field of view.

Human 'common sense' hazard anticipation, preconception or perception can to some extent be emulated by emerging A.I. techniques (such as off-line trained CNNs), but we should be realistic as to what can be achieved with these A.I. techniques and not allow them to result in automation bias. In general, we advise that they should only be used where they can only add benefit and their failure does not significantly worsen a situation, such as triggering unnecessary emergency braking or other evasive action which in itself could be hazardous.

Some modules from the Highway Pilot GSN are considered broadly applicable to the Urban Pilot use case, and have not been altered. These include:

No.	Module type	Description
HP1	Lane Reallocations	This covers lane closures (e.g. due to a stranded broken down vehicle) and changes in lane usage restrictions (bus lanes, car share lanes)
HP3	Lane Rerouting	This is for when a lane position needs to be temporarily or permanently changed, usually during and after roadworks
HP4	Adverse Weather	Strategies to restrict the use of the feature during bad weather, particularly during the sudden onset of challenging weather when the system is already active
HP5	Mechanical Failure	To make sure that mechanical failures do no go undetected while the feature is in use.
HP8	Obstructions	Strategies for handling collisions with static and dynamic obstructions in the road

Table 5 : Modules from Highway Pilot Use Case which are considered applicable to Urban Pilot Use Case

For the 'Obstructions' module (HP8), there may be differences in strategy associated with the differences in design speed, but the issues are largely the same. Passive containment measures (such as fencing and monitoring) are probably impractical in the urban environment and more reliance on systematic vehicle-based measures may be appropriate (i.e. sense obstruction and stop). Pedestrians in the road are handled in separate modules for Urban Pilot.

The following table outlines new modules that were constructed as part of the Urban Pilot GSN.

No.	Module type	Description
UP2	Road Etiquette	This covers allowing emergency vehicles to pass and appropriate speed regulation. Also pick-up/drop-off and emergency stopping, traffic merging and prudence when passing pedestrians on pavements
UP7	Operational Envelope	Ensuring the feature is only active on the roads it is intended for. It does not cover other aspects of the control envelope such as stability and speed regulation
UP9	Junctions and Crossings	Handling of signal and priority controlled junctions and level crossings
UP10	Pedestrian Crossings	Handle Zebra, signalled and uncontrolled pedestrian crossings
UP11	Overtaking	Handle being overtaken (overtakee), facing an on-coming overtaker, overtaking and undertaking
UP12	Antisocial Behaviour	Resilience to mistreatment and misuse of the vehicle such as pranks, use of the vehicle for criminal activity, entrapment of vehicle occupants, and general poor treatment from other road users

Table 6 : Urban Pilot Use Case GSN Modules

The structure of the Urban Pilot GSN is confirmed as shown in Figure 17:

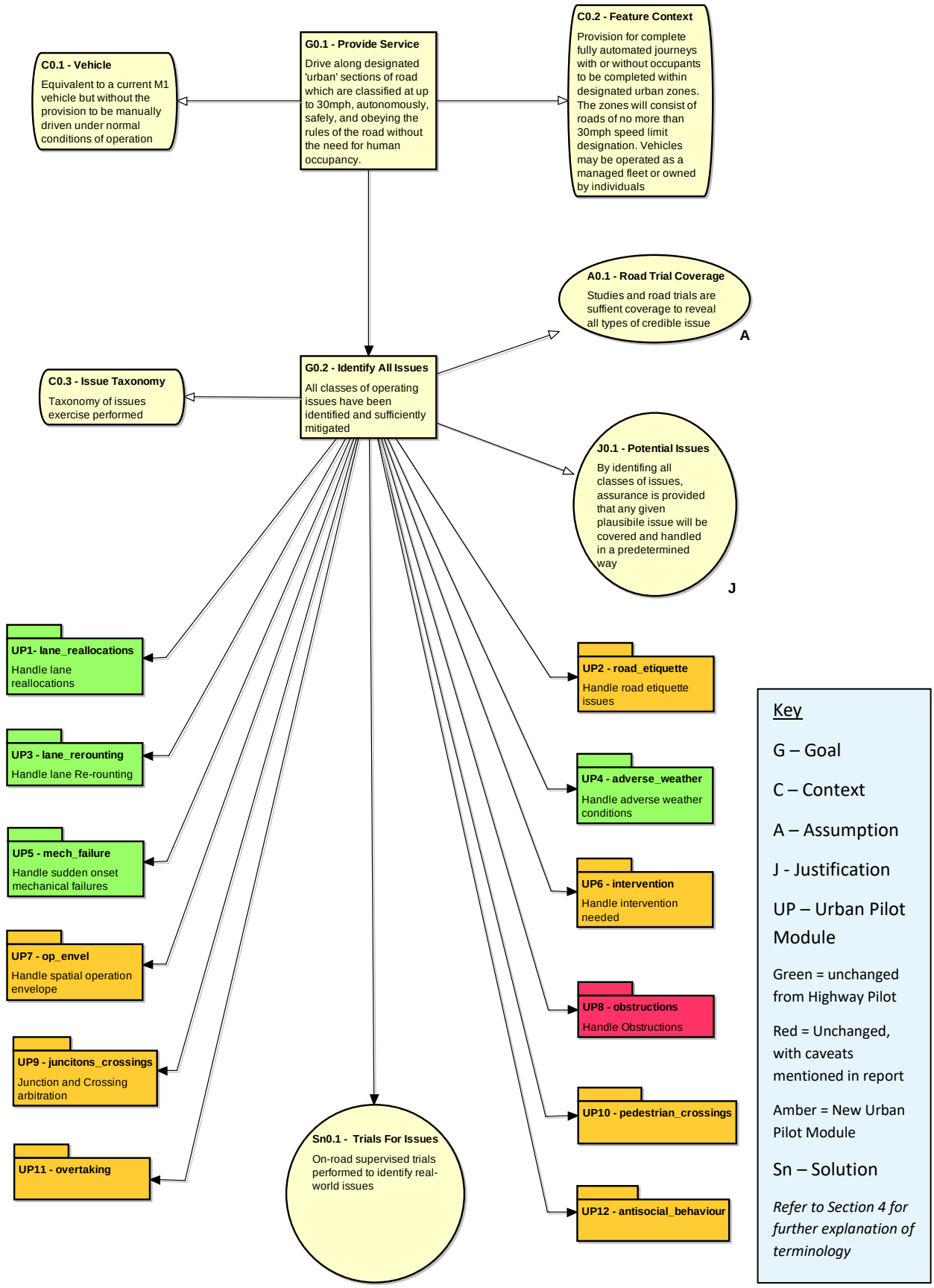


Figure 17 : Urban Pilot GSN – Top Level Structure

Details of each module that is unique to the Urban Driving use case are provided as follows.

6.3 Road Etiquette (UP2)

In addition to emergency vehicle passage and speed regulation covered in the Highway Pilot road etiquette, this module includes coverage of passing parked vehicles, passing pedestrians, appropriate stopping, and merging obstruction etiquette which will be explained in the sections below.

The road etiquette modules have been broken down further to sub-modules, as shown in Figure 18:

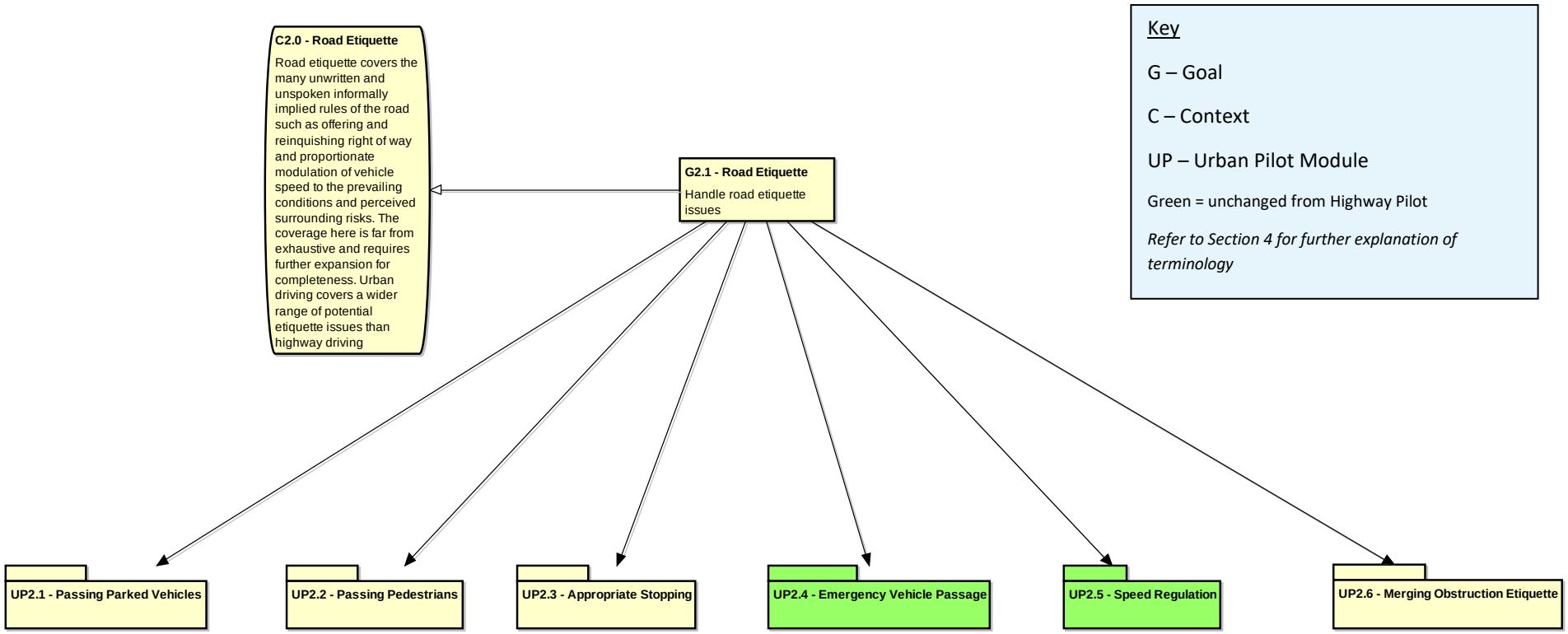


Figure 18 : Urban Pilot GSN – Road Etiquette Module Structure

6.3.1 Passing Parked Vehicles (UP2.1)

Parked vehicles present a risk due to the potential for doors to be opened as well as obscuring pedestrians who may be attempting to cross between parked vehicles. Human drivers may be much better at detecting the signs (visual cues) that a vehicle occupant is present and may be about to open their door. It is probably unrealistic to expect machine vision and perception to be able to see inside each car and make a judgment whether there is a risk someone is about to try to leave their car. The mitigation options range from doing nothing and accepting the risks, through to slowing down and/or increasing the clearance by moving across the lane if space allows.

The Passing Parked Vehicles sub-module is shown in Figure 19.

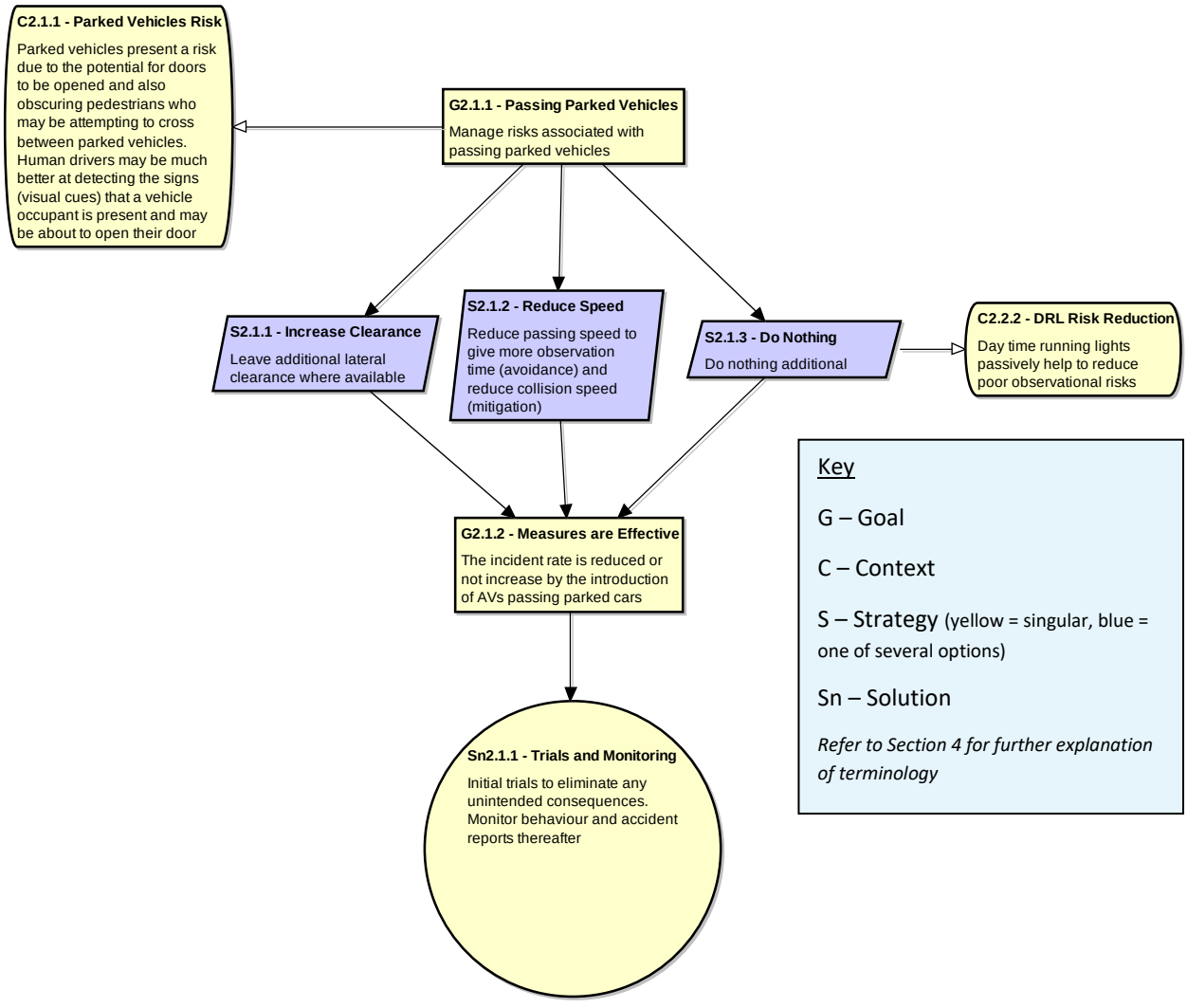


Figure 19 : Urban Pilot GSN – Passing Parked Vehicle sub-module (UP2.1)

6.3.2 Passing Pedestrians (UP2.2)

Detecting pedestrians in the road is covered by obstructions and pedestrian crossings modules. However, one issue of etiquette is anticipating when pedestrians who are not waiting to cross may suddenly step on the road either deliberately or unintentionally. Pedestrians can present a risk if they enter the carriageway at short notice and a pre-emptive action may need to be taken. Examples might include an unaccompanied young child, a person under the influence of alcohol or a crowd of rowdy teenagers waiting for a school bus on a narrow footway. One option is to do nothing anticipatory. Real-world incidents usually result from a combination of events aligning in an unfortunate way. Automation will not suffer from human driving deficiencies such as driver inattention and breaking of speed limits so even if no direct action is taken there may be sufficient net reduction and mitigation of many pedestrian step-out related collisions. The disadvantage of automation is that an attentive human driver may be able to spot pedestrians who represent a risk of entering the road and apply due diligence by slowing down, moving the vehicle further away from the kerb, or sounding the horn. This is one occasion where A.I techniques may prove beneficial since they have ability to evolve an expert judgement to detect risk cases and the consequences of false positive detections are minimised (e.g. unnecessary slowing down and moving out, pre-charging the brakes and emergency stop etc).

The Passing Pedestrians sub-module is shown in Figure 20.

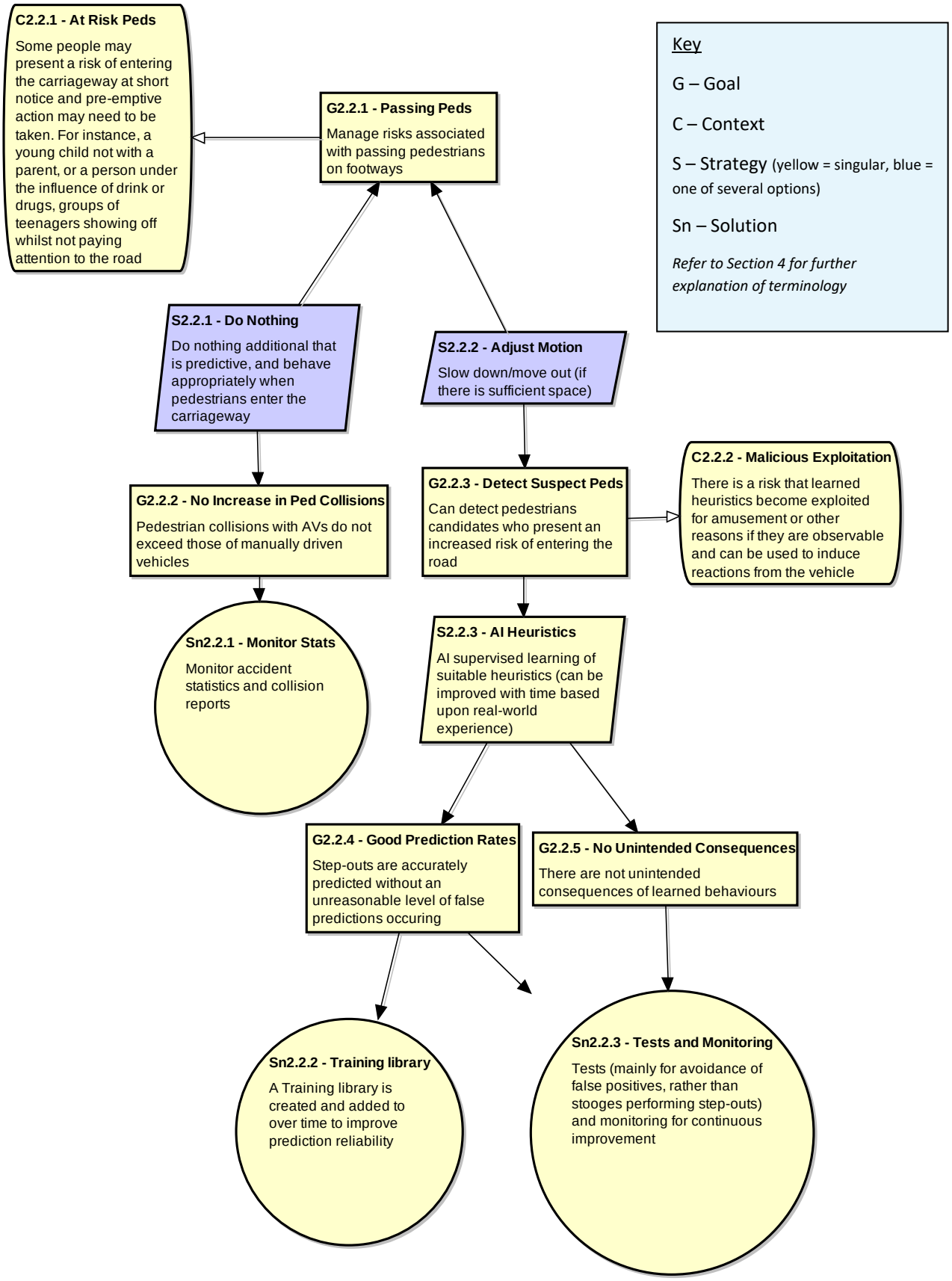


Figure 20 : Urban Pilot GSN – Passing Pedestrians sub-module (UP2.2)

6.3.3 Appropriate Stopping (UP2.3)

Appropriate stopping considers how and where to appropriately permit the vehicle to pick-up and set down passengers, as well as the provision for an occupant initiated emergency stop (E-Stop). Much of this is procedural and could be resolved as part of a system deployment and it could be for the vehicle manufacture themselves to address the needs and preferences of their users. However, some consideration needs to be given to the fact that automated stopping for egress and ingress may lack the finesse and diligence that a human driver may use, and just following the Highway Code as it stands may not be sufficient in practice. Consideration of an E-Stop should be given for unforeseen and even foreseen circumstances which cannot be gracefully handled in any other way. The strategy presented in the GSN model only offers a controlled deceleration and not a full braking force stop or complete power-down of the system. Either of these two latter cases could themselves be dangerous in use. The counter risk is that a progressive E-Stop might misused, and passengers could start to use it only for the purposes of stopping to alight prior to the preprogramed destination of the vehicle.

It might be worth considering having a kill-power switch (similar to those fitted to busses) on the outside of the vehicle which can only be activated when the vehicle is stationary. The reverse argument to this is that it would also present a significant security risk (for example if someone wanted to disable the vehicle to mug the vehicle occupants) and existing passenger vehicles do not generally have this albeit that they are currently unable to run-empty.

The Appropriate Stopping sub-module is shown in Figure 21.

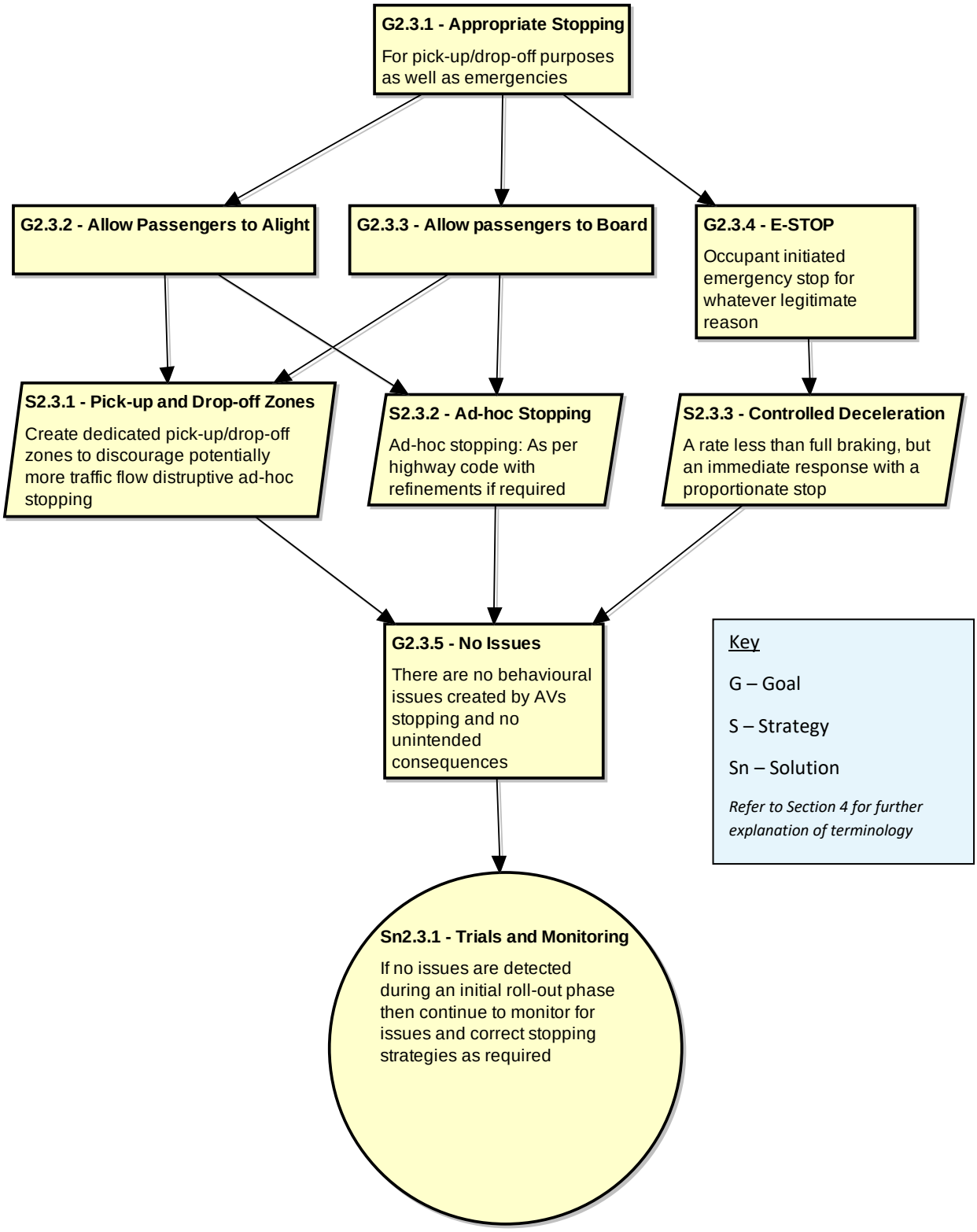


Figure 21 : Urban Pilot GSN – Appropriate Stopping sub-module (UP2.3)

6.3.4 Merging Obstruction (UP2.6)

At time of peak traffic volumes and in slow moving queues, often the absolute right of way is relinquished to allow others to merge in the interests of common sense and to keep the traffic moving or just due to empathy in the hope that others will do the same for you. There are no rules for this (except for certain cases, such as allowing busses to pull out), and some rules may need to be developed if AVs are to follow the example of their human driven counterparts. Doing nothing might be an option particularly while CAVs have a low market penetration and make up only a small fraction of the UK parc. This would also allow time to wait for managed connected infrastructure to emerge which may intelligently control traffic flow at pressure points for example, rather than attempting to solve the problem from the outset. Otherwise a strategy of leaving gaps or a merge in turn approach could be taken. Whatever approach is taken, it should be monitored for undesirable effects that may worsen traffic congestion or driver frustration, rather than improve it.

The Merging Obstruction sub-module is shown in Figure 22.

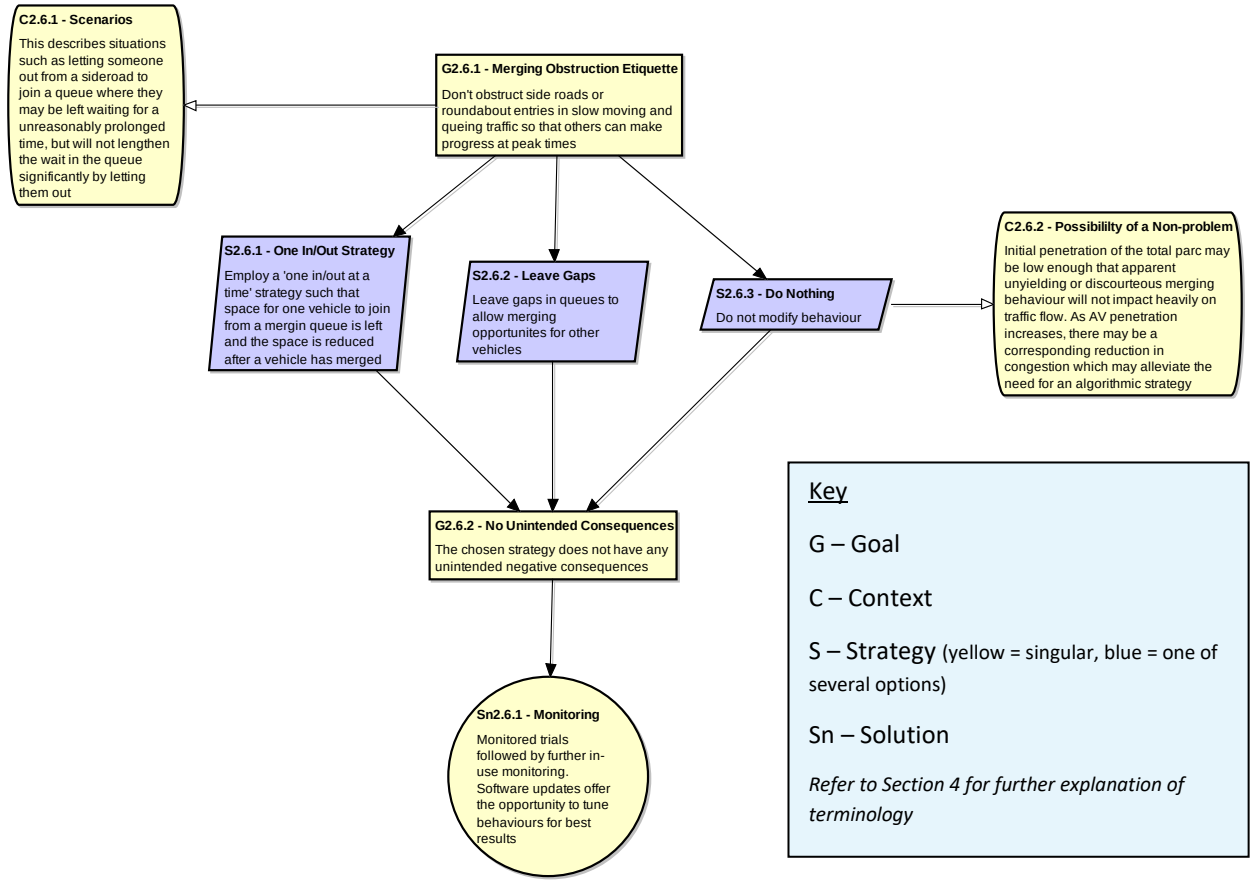


Figure 22 : Urban Pilot GSN – Merging Obstructions sub-module (UP2.6)

6.4 Operational Envelope (UP7)

This module is largely the same as its counterpart for Highway Pilot. The main difference is that there is no option to trust a human driver to constrain automated operation to the intended regions and road types. For this reason, only systematic methods of geo-fencing are considered. The Urban Pilot Operation Envelope module is shown in Figure 23.

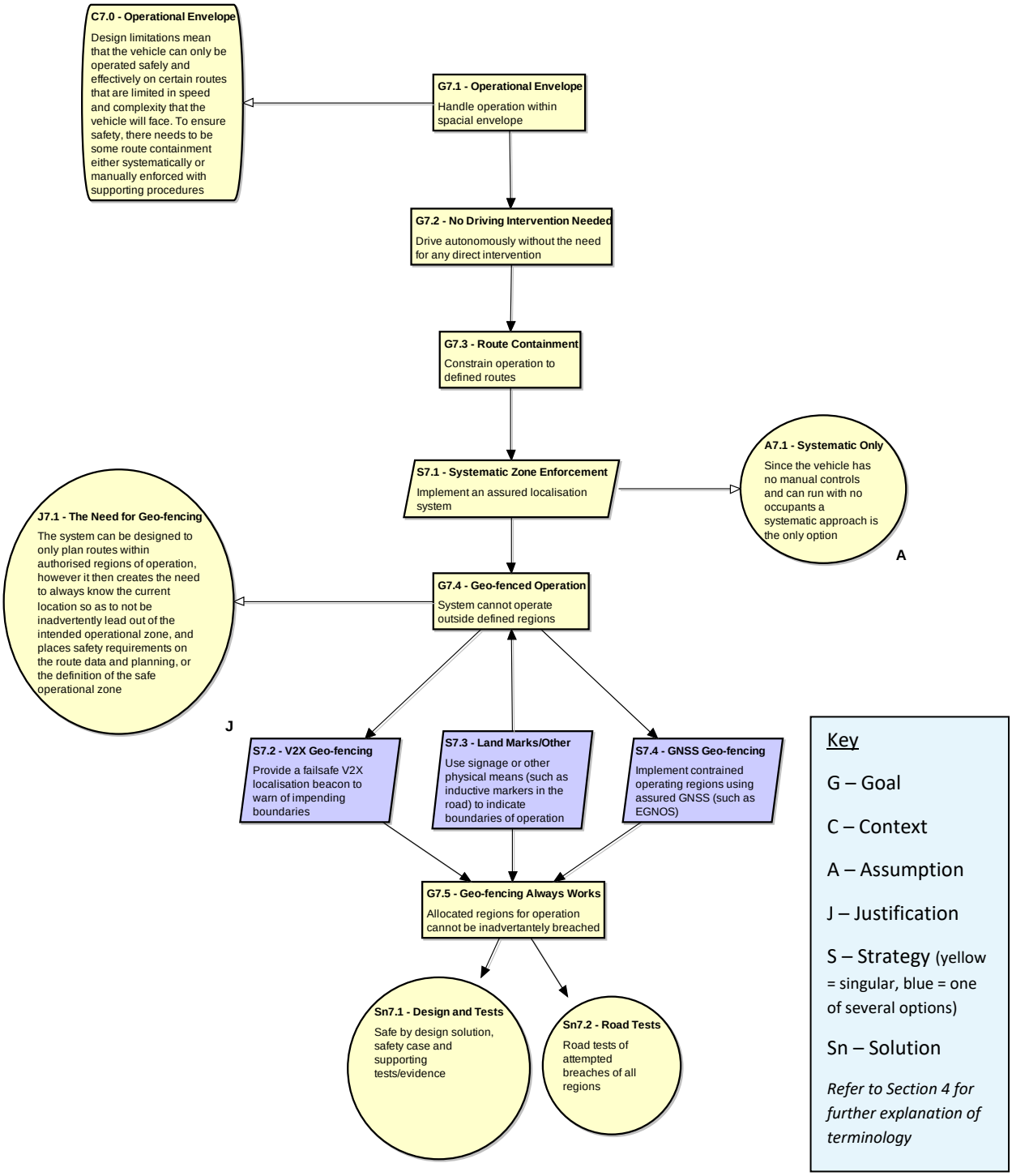


Figure 23 : Urban Pilot GSN – Operational Envelope-module (UP7)

6.5 Junctions and Level Crossings (UP9)

The junctions considered for this module have been divided in to:

- signal controlled junctions;
- priority controlled junctions;
- level crossings;
- signalled controlled junctions with a fault failure.

Starting with signalled junctions, the main concern is being able to detect the signal phase both in time to stop for a red signal and to not pass through a red signal at any time. The means of detection falls into two categories. The first is to use machine vision (normally Complementary Metal-Oxide Semiconductor (CMOS) cameras) to see the signal lamps in the same way as human drivers do. The second is to use the more direct method of Infrastructure to Vehicle (I2V) radio communications.

It is suggested by the authors that whilst it may at first appear to be convenient strategy to use vehicle mounted cameras to view existing traffic signals, it is conceptually flawed and problematic. Humans are capable of vision fixation and tracking of a visual target through a variety of means such head and body movement, as well as saccadic eye motion to give which is known as *directed* or *active* vision. The human eye also has variable aperture and focus so together can cope with resolving the detail from very complex scenes in challenging light conditions. Beyond this, humans are capable of contextual perception so that for instance a green balloon held in front of a signal by a pedestrian waiting to cross would not be mistaken for a green traffic signal[16]. The windscreen mounted cameras used on vehicles are fixed focus and aperture and cannot be pointed. This is primarily for keeping cost, complexity and reliability at their optimums, but adding variable focus, aperture, and a mechanic servo means of directing the camera would tend to add more than it solves. All of these items take time to actuate and you need to know in advance where to direct the area of interest to which requires perception. Notwithstanding that these technical challenges could perhaps be solved with time, it is still an inferior and unnecessarily challenging strategy. The signal phase is known to the traffic signal controller to a high level of integrity. The notion of converting this to an analogue optical signal then attempting to convert that back to a digital signal using optical sensing from a moving vehicle under any possible light condition does not hold up to reason when the outcome of that process is preventing potentially life threatening collisions.

Various V2I schemes are emerging around the world (e.g. Dedicated Short-Range Communications (DSRC)) to make traffic signal phasing available wirelessly to vehicles. However, this information is being transmitted as a driver advisory, not as a mission critical piece of information. This exchange of information needs to be developed to be failsafe. The source of the information is already failsafe but how it is propagated may not be and further analysis and development may be needed. In addition, the communication needs to be allowed to fail (be absent when expected) without harmful consequences to the approaching vehicle. For this to be allowed, the vehicle needs to know when to expect a transmission. For this to work the vehicle system needs to know where it is against a digital map to realise that it is approaching a signalled junction and needs to have the real-time signal status for the junction so that it can stop the vehicle if this information cannot be obtained for whatever reason. This places a high level of

[16] http://www.theregister.co.uk/2016/08/24/google_self_driving_car_problems/

dependence on digital map integrity and on the performance of localisation. Both must be *'correct',* commensurate with the risk of *running* a red light. This is also true where cameras are used since a missed signal needs to be noticed in its absence.

Some consideration to signal failure is given also. Currently signalling systems cannot fail in such a way as they show a green light to two or more conflicting directions. This is covered by TR 2500 and BS EN 12675:2001. Generally multiple red signals are provided to ensure redundancy. If there is a severe fault, such as two or more red signals failing, all signals are extinguished to avoid ambiguity. Some other countries use a flash red-amber sequence to indicate a fault. Problems for CAVs may still arise for partial failures of certain signals. This is another benefit to using a wireless radio V2I approach where rather than trying to infer a fault, the exact status can be transmitted to vehicles so that they may proceed with caution.

Priority controlled junctions are another significant challenge for which it is difficult to offer a general strategy beyond that of proceeding with caution. A move towards wirelessly managed junctions, possibly signalled junctions would assist with resolving the technical challenge, but may be impractical in practise. Level crossings are out of scope for the feature, but cannot be simply ignored either. The strategy is to avoid them by choosing alternate routes or having them removed, but if they are to be avoided then this needs to be enforced to the same level of integrity as the risk of encountering one and an incident occurring. If this cannot be guaranteed, then additional monitoring and support at level crossings may be needed in urban areas where CAVs may encounter them. The Junctions and Level Crossings module is shown in *Figure 24*.

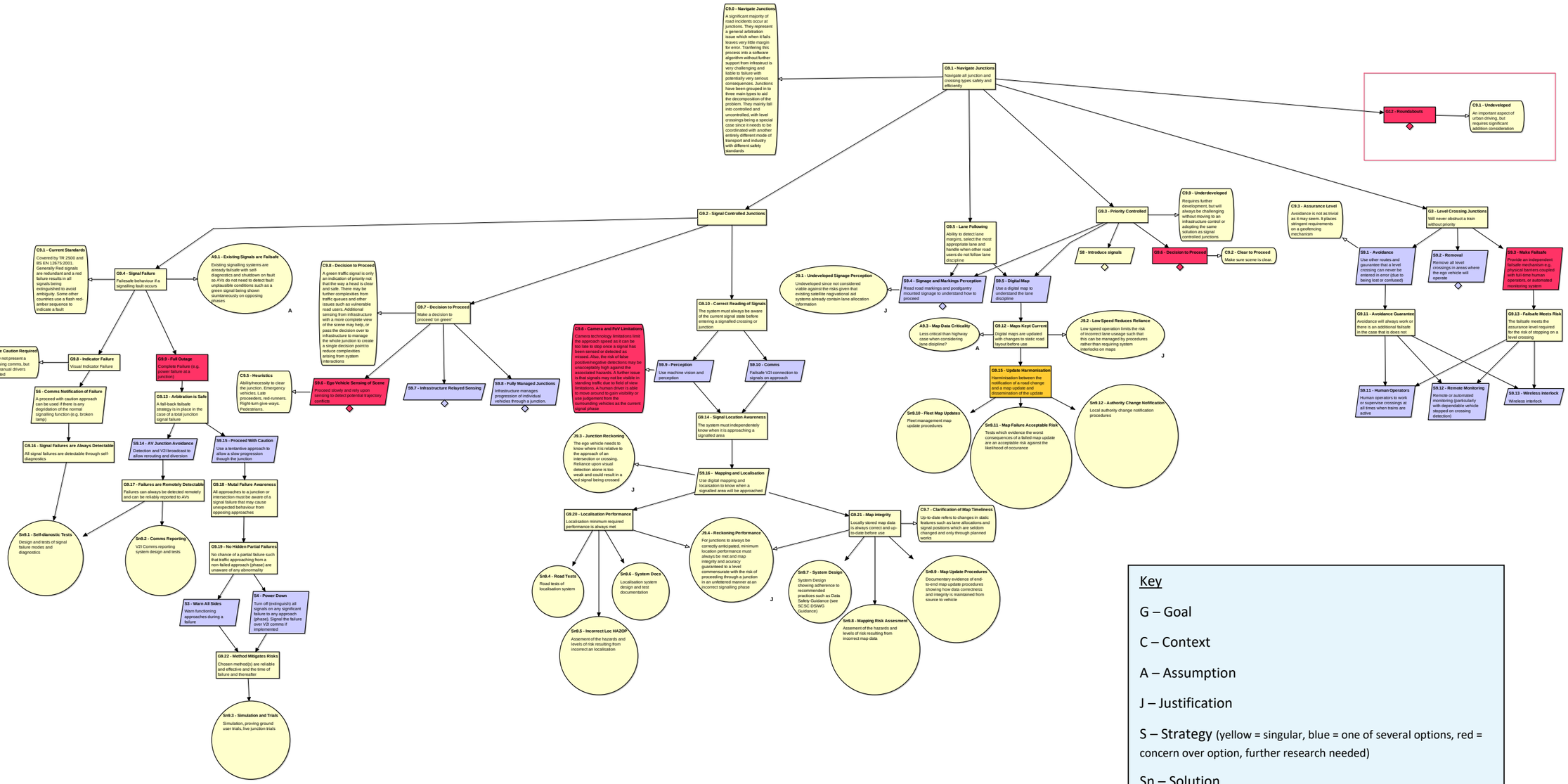


Figure 24 : Urban Pilot GSN – Junctions and Level Crossings-module (UP9)

6.6 Pedestrian Crossing Arbitration (UP10)

Pedestrian crossings are perhaps one of the most challenging everyday aspects of operating a CAV in an urban area. The pedestrians are not a compliant part of a system who can be directed and controlled and will exercise free will which will be experienced as a random and chaotic variable to the automated system. Unlike many of the potential obstructions on motorways, which are quite rare but still need to be addressed, pedestrians are vulnerable and will be frequently and routinely encountered by CAVs operating in urban areas, yet offer many behavioural and detection challenges as described by Colin Sowman in his paper for ITS International in October 2016 and referenced in this article[17]:

> *Also relevant are the figures obtained by the UK's Institute of Advanced Motorists which show that despite the potentially fatal consequences, almost half of pedestrians knocked down by a vehicle did not take enough care before stepping into the road. If pedestrians and cyclists know vehicles (ADAS or driverless) will not hit them, they will walk or cycle across roads at will which could bring city-centre traffic to a near standstill. While the increased safety is to be welcomed, additional measures or legislation will be needed to control pedestrians and cyclists in order to keep the traffic flowing.*

This is a form of automation bias. The thresholds of expectation may move so that by trial and error people may step out in front of moving vehicles that are moving at greater speeds with much less physical clearance than before. This may be accelerated when coupled with the fact that people won't feel the need to worry about disrupting an empty running vehicle by forcing it to slow down or stop. There is also the malicious aspect to defeat any cautious strategy which is put in place, also referencing Colin Sowman:

> *By simply walking out in front of a driverless vehicle, braking sharply ahead of it or placing cones across a road, criminals could divert a driverless vehicle to hijack it, steal the cargo or rob the passengers. This is particularly the case where the 'driver' cannot assume control.*

From mischievous teenagers looking for a prank to impress friends through to people with more malevolent intentions, knowing when to stop whilst keeping vehicle occupants secure is of concern. C-ITS solutions are emerging now which may offer significant steps forward with the pedestrian detection challenge for controlled crossings. The follow extract is from an article published by Neavia Technologies[18]:

> *…vehicles equipped with V2X technology can automatically receive alerts when pedestrians are crossing or about to cross a road. This represents a significant step forwards for road safety: In many situations, pedestrians are not visible by to car drivers. They can be hidden due to the road configuration, or by other vehicles. They can also be less visible in case of fog, or under poor lighting conditions. In those situations vehicles' ADAS systems (Advanced Driver Assistance Systems) are less relevant, or reacting extremely late.*

The main concern for these approaches is their potential for inconsistency, as they are being offered to bolster the vehicle's own performance, rather like ADAS does for a human driver. If they are to be relied

[17] http://www.itsinternational.com/categories/utc/features/the-downside-of-driverless-vehicles/
[18] http://www.neavia.com/2016-11-neavia-unveils-worlds-first-v2x-pedestrian-warning-solution/?lang=en/

upon then they need to be designed with this in mind, so that the vehicle's systems can place the same level of reliance on the infrastructure support each and every time a crossing is approached. This means that both the pedestrian detection and the communication mechanism need to meet agreed minimum performance standards rather than taking the current approach of '*it will help when it can*.'

The main theme of the module is that there is much which can be done with infrastructure to lower the risks on dedicated crossings. Informal crossing on arbitrary sections of road will still incur risks from sensing limitations, but these could be mitigated by improving public understanding and updating the Highway Code and laws in respect of these technological limitations. In the same way that it is accepted that you do not have a right to roam on railways, rather it is a form of trespass, some consideration to obstructing traffic flow by being in the road may have to be given on the grounds of both safety and disruption. The module considers three forms of crossing from fully signal controlled, through the rule controlled Zebra crossing to general uncontrolled crossing which can take place almost anywhere.

The Pedestrian Crossings module is broken down into sub-modules, as shown in Figure 25.

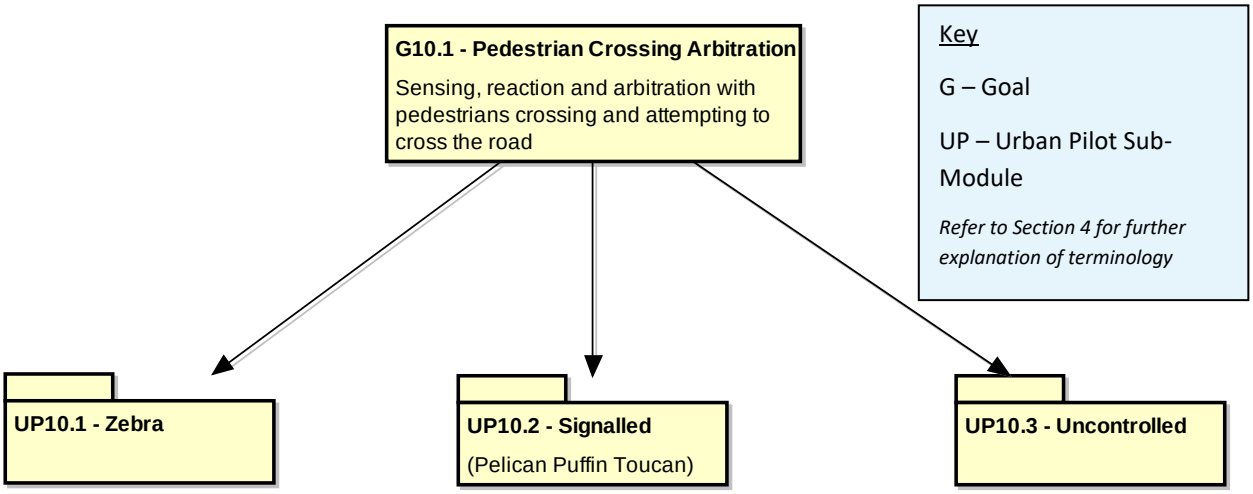


Figure 25 : Urban Pilot GSN – Pedestrian Crossing Arbitration Module Structure

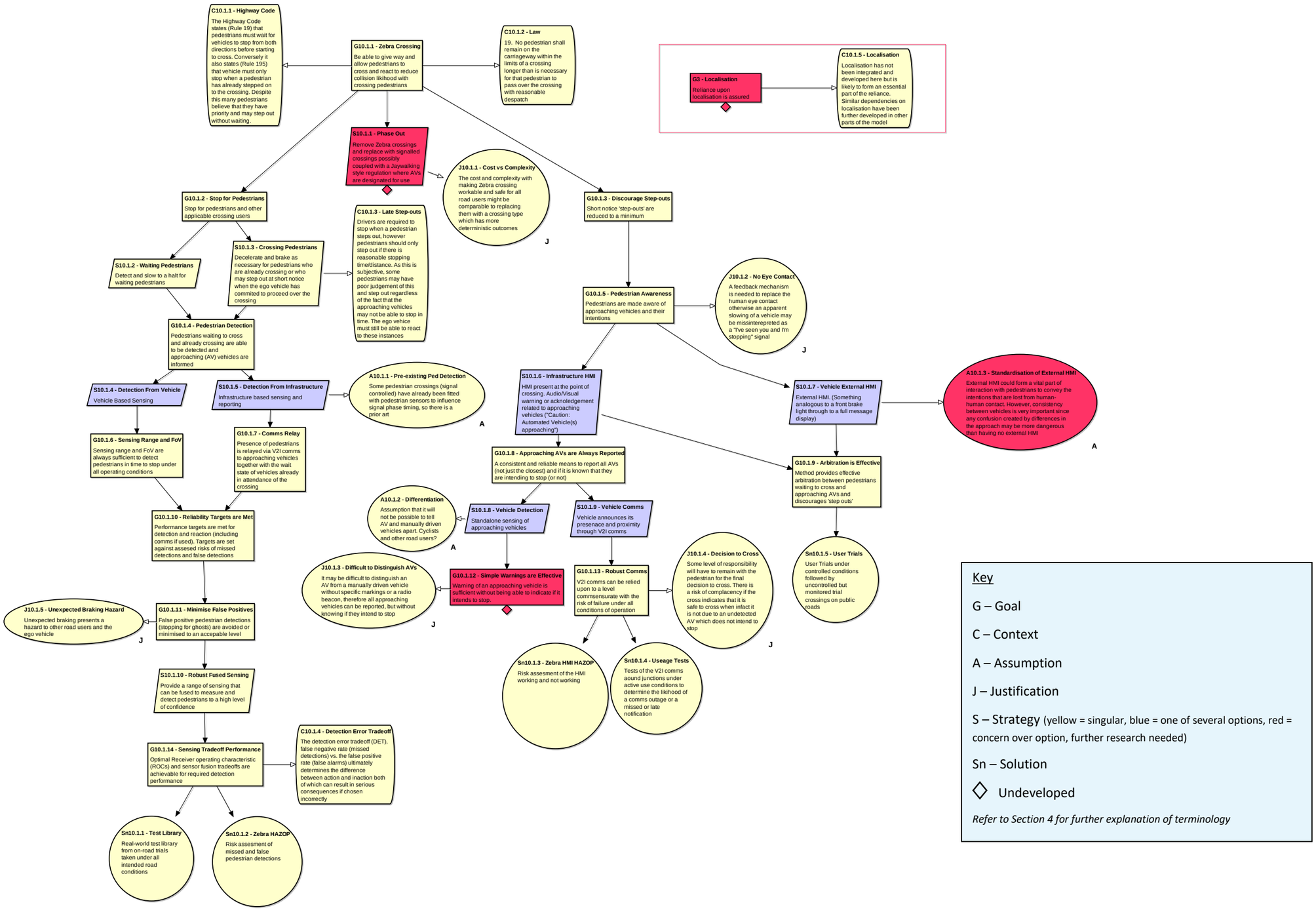
6.6.1 Zebra Crossings (UP10.1)

Zebra crossing are the main example upon which the other types of crossing have been derived from in the GSN model. The technically simplest approach to zebra crossings would be to replace them with signalled junctions, which are far more deterministic and do more to discourage people from just stepping out or even running out on to the crossing. However, it may not be practical or cost effective to remove or upgrade them all. One issue with them is understanding who has right of way. Vehicles are mandated to stop when a pedestrian steps out on to the crossing, however pedestrians are supposed to wait for an approaching vehicle to stop (to make sure it can and does stop) before crossing. If these rules were applied literally by all parties, then either pedestrians would never be able to cross on busy roads or there would be a stalemate whilst one waits for the other. Pedestrians are not compelled, as motorists are, to read the Highway Code and there exists an option that pedestrians have an automatic right to cross no matter what and only a sense of self-preservation prevents some people from walking out in front of a vehicle in a stream of traffic with insufficient time for it to stop. It is left to the judgement of the pedestrian (who may be a child with limited experience) whether a vehicle has enough time and distance to comfortably stop without hitting them on the crossing.

The first challenge is to discourage pedestrians stepping out when there is not time to stop. It should be borne in mind that any pedestrian detection system, no matter how good, may fail at some point so prevention is better than cure in that pedestrians need to wait for the vehicle to demonstrate that it is going to stop for them in sufficient time to allow them to cross. This could be achieved via an external visual/audible HMI on the vehicle itself, or via the same HMI but mounted on or near the crossing. There needs to be some conformance in the approach from vehicle manufacturers such that pedestrians become accustomed to one clear message rather than various brand defining techniques, sounds and/or graphics. For this reason, the crossing infrastructure approach has some appeal, but it adds the additional complexity of requiring a failsafe mechanism for cars which do not announce themselves when approaching the crossing. This would likely employ a similar mechanism to signalled junctions where the vehicle must always know when it is encountering a crossing and expect a handshake before continuing though at full speed. Further confusion may result from the mixture of legacy vehicles with CAVs. Just because every approaching CAV has announced it will stop does not mean a manually driven vehicle's driver has noticed a pedestrian is on the crossing. A proceed to cross with caution message may help with this as a reminder to those crossing the road that they should make the final visual check themselves. More thought would need to be given to people being assisted by guide dogs.

The other part of the equation is pedestrian detection. This could be done from infrastructure, or from the approaching vehicle. The issue with a vehicle based strategy alone is that it may fail to detect pedestrians under some conditions, not least due to the limited field of view from the vehicle itself. A technically superior solution would be to detect via infrastructure mounted sensors, with vehicle based detection also used as a final resort in case a pedestrian runs out onto the crossing. Time-of-flight and infrared camera pedestrian detecting sensors are being developed for infrastructure mounting and these offer the reliability and robustness to adverse weather and light conditions and well as a good field of view when appropriately mounted. Also, infrastructure based vehicle detection is already common practice. Having the crossing itself as the one source of truth to perform the crossing arbitration seems an attractive option for an assured system.

The Zebra Crossings sub-module is shown in Figure 26.



Key

G – Goal

C – Context

A – Assumption

J – Justification

S – Strategy (yellow = singular, blue = one of several options, red = concern over option, further research needed)

Sn – Solution

◇ Undeveloped

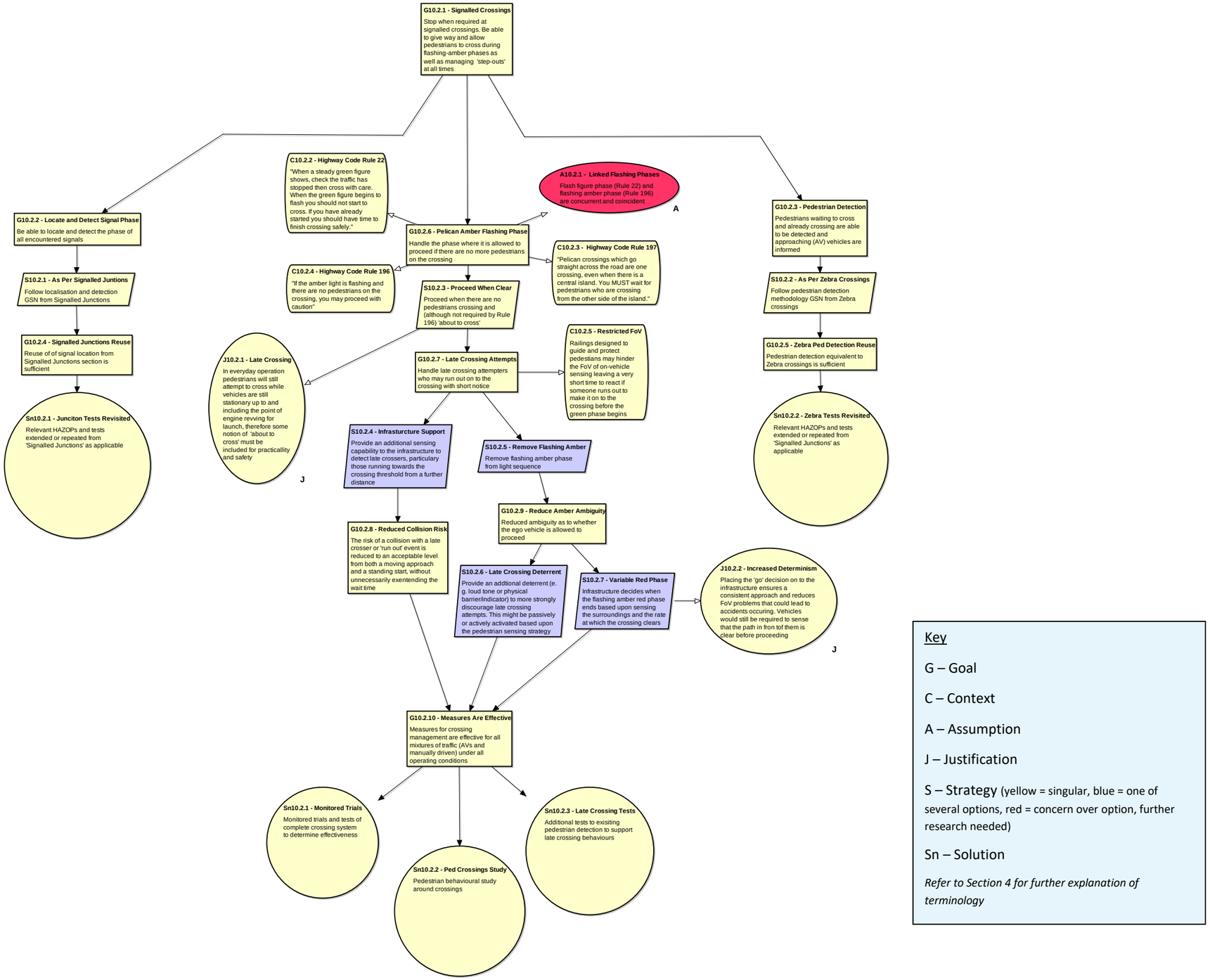
Refer to Section 4 for further explanation of terminology

Figure 26 : Urban Pilot GSN – Zebra Crossings sub-module (UP10.1)

6.6.2 Signalled Crossings (10.2)

This covers Puffin, Pelican and Toucan crossings and their other derivatives. The module borrows from and links to zebra crossings and signalled junctions as aspects of both are required. The approaching vehicle needs to be sure of the signal phase so the same solution should be adopted. Beyond that, pedestrians may not wait for the correct signal phase and may cross late or early and run out to try to make it in time before cars pull away. Electric Vehicles (EVs) will not have the running engine revving or restarting as an early indicator that vehicles are starting to move off. For these reasons, much of the same logic is needed as for zebra crossings, where pedestrian will attempt to cross at will so it can be treated as a traffic signalled Zebra crossing. It is also suggested that the flashing amber phase is removed entirely since with pedestrian detection the red phase is made longer. Having the ambiguity of the flashing amber phase only provides opportunity for problems, leading to the vehicle needing to detect itself if there are any stragglers on the crossing which may be prone to error. Different vehicles would potentially have different implementations of this and it would be better left to the infrastructure with wider and dedicated sensing capabilities to make the final judgement to proceed if clear.

The Signalled Crossings sub-module is shown in Figure 27.



Key

G – Goal

C – Context

A – Assumption

J – Justification

S – Strategy (yellow = singular, blue = one of several options, red = concern over option, further research needed)

Sn – Solution

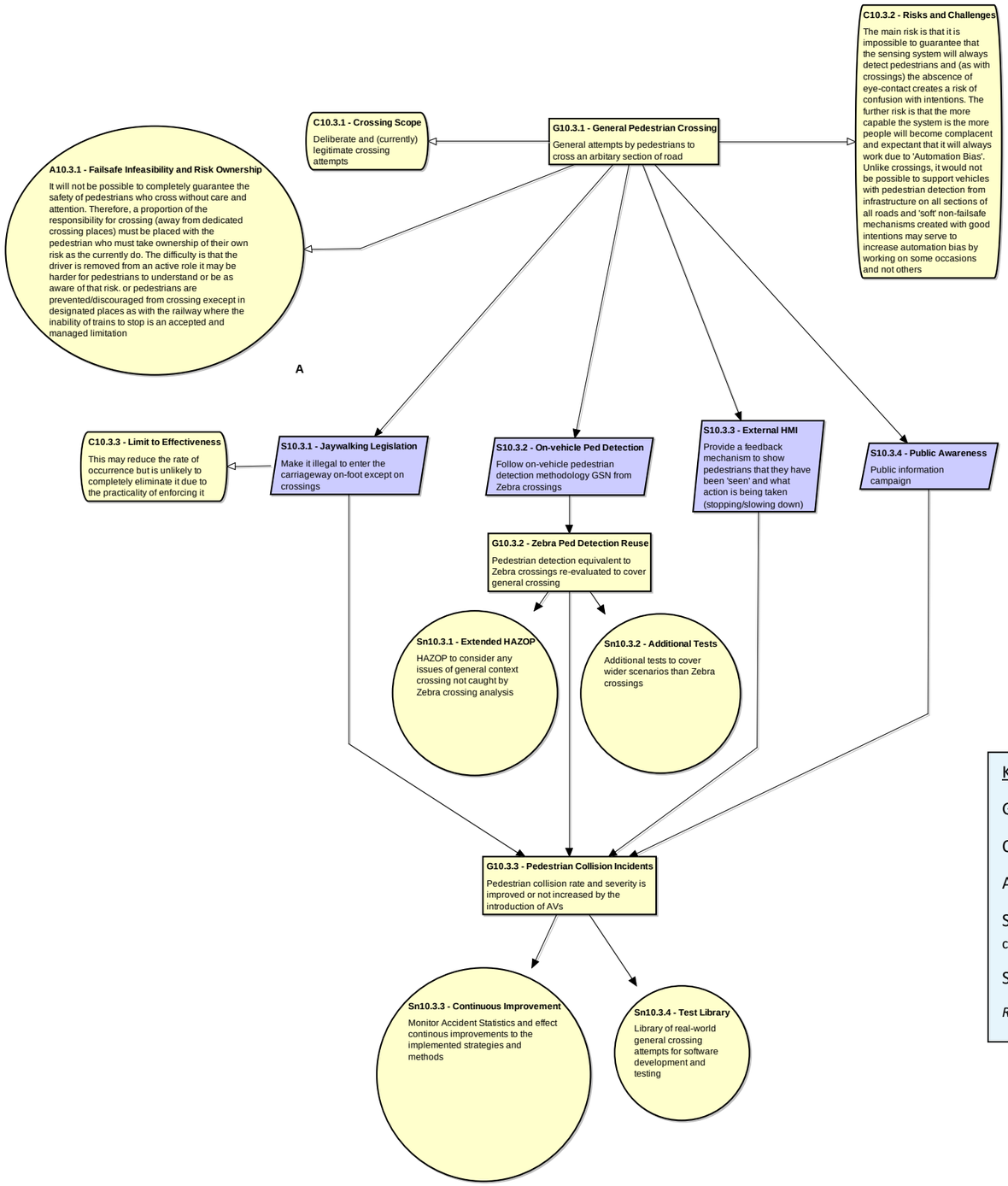
Refer to Section 4 for further explanation of terminology

Figure 27 : Urban Pilot GSN – Signalled Crossings sub-module (UP10.2)

6.6.3 Uncontrolled Crossing (10.3)

This presents the most challenging in that as with other infrastructure crossings, there are no guarantees the vehicle will sense every pedestrian every time. However, the lack of infrastructure means it will either need to be handled from the vehicle or pedestrians are restricted from crossing away from designated places. That would imply a Jaywalking law which from experiences from the USA, are often not adhered to or rigorously enforced. There is also the issue of public acceptance of such a law as pedestrians hold on to the notion of a right to cross anywhere, except motorways and railways. This is perhaps one area where a drop in absolute safety (human to automated) may have to be conceded and might be addressed by a public education programme to raise awareness. The real danger is one of automation bias because in the clear majority of cases the automated system may perform as well or better than a human driver. This will nurture a false sense of security and the public may start to assume an unrealistic margin of safety when attempting to cross. As the threshold is pushed with regard to stopping distance, the risk of being hit due to a sensing and perception insufficiency will increase. Ultimately the pedestrians hold their fate in their own hands, but it may help to clarify the legal position away from the unrealistic position of expecting a software based system to detect pedestrians 100% of the time under all conditions, and strongly encourage the use of official crossings where they are provided.

The Uncontrolled Pedestrian Crossings sub-module is shown in Figure 28.



Key

G – Goal

C – Context

A – Assumption

S – Strategy (yellow = singular, blue = one of several options, red = concern over option, further research needed)

Sn – Solution

Refer to Section 4 for further explanation of terminology

Figure 28 : Urban Pilot GSN – Uncontrolled Pedestrian Crossings sub-module (UP10.3)

6.7 Overtaking (UP11)

The overtaking module covers two way roads where vehicles cross over partially or fully into a lane intended for on-coming traffic in order to, pass another vehicle or stationary obstruction. It covers more than just the ego vehicle performing an overtaking manoeuver since it can itself be overtaken or face an on-coming vehicle which is overtaking from the opposing direction. There are also situations where a so-called undertake manoeuvre may legitimately need to be performed. For the ego vehicle the basic message is to not attempt to overtake since the benefits do not outweigh the technical risks. Also in practice some temporary infringement of the speed limit is often required to minimise the on-coming collision exposure time, but this as an unlawful action and is difficult to sanction from an algorithmic perspective. However, in certain circumstances it may be impractical for the ego vehicle to not overtake in order to make reasonable progress such as when following behind slow moving road users (cyclists, road maintenance machinery) or stationary obstructions such as rows of parked cars.

When being overtaken it may become necessary to slow down to allow the overtaking vehicle to move over early if it meets an on-coming vehicle unexpectedly. This may be the fault of the vehicle performing an unsafe overtake, but some drivers do so expecting to be able to force their way back in if needed. In the interests of trying to prevent serious head-on collisions there may need to be some situation recognition and response which could be described as road etiquette, but this may diminish its potential seriousness. An option is to not do anything and hope drivers learn or are informed of this new potential risk from CAVs when performing overtakes.

Whichever overtake is being considered, the sensing performance needs to be capable of detecting and responding to the higher closing speeds involved which could as high as 70 mph for a 30 mph speed limited road. Sensing and sensing fusion techniques (where used) must be shown to be able to handle the scan and frame rates needed for situation recognition at these speeds otherwise the ego vehicle is not fit for operation in unsegregated two-way traffic.

For passing a stopped vehicle, understanding when a vehicle has actually stopped, with the intention of being passed may be much more difficult that it first seems. A vehicle may be have stopped to set down passengers, or to allow another larger vehicle to pass from the opposite direction. Simply assuming it has stopped and is not waiting for something in the road ahead and pushing past might cause considerable disruption. Some visual or V2V indication of its intentions would help with this, but may require modification of the entire existing vehicle parc to be effective.

Undertaking of stationary vehicles which have stopped waiting to turn right or for loading reasons on one-way streets may be necessary on some routes. Determining the correct context and ensuring sufficient space to do so are challenges to be overcome.

Additional consideration to passing cyclists is needed as they have some of the elements of vulnerability and unpredictability in common with pedestrians.

The Overtaking module has been split into sub-modules, as shown in Figure 29.

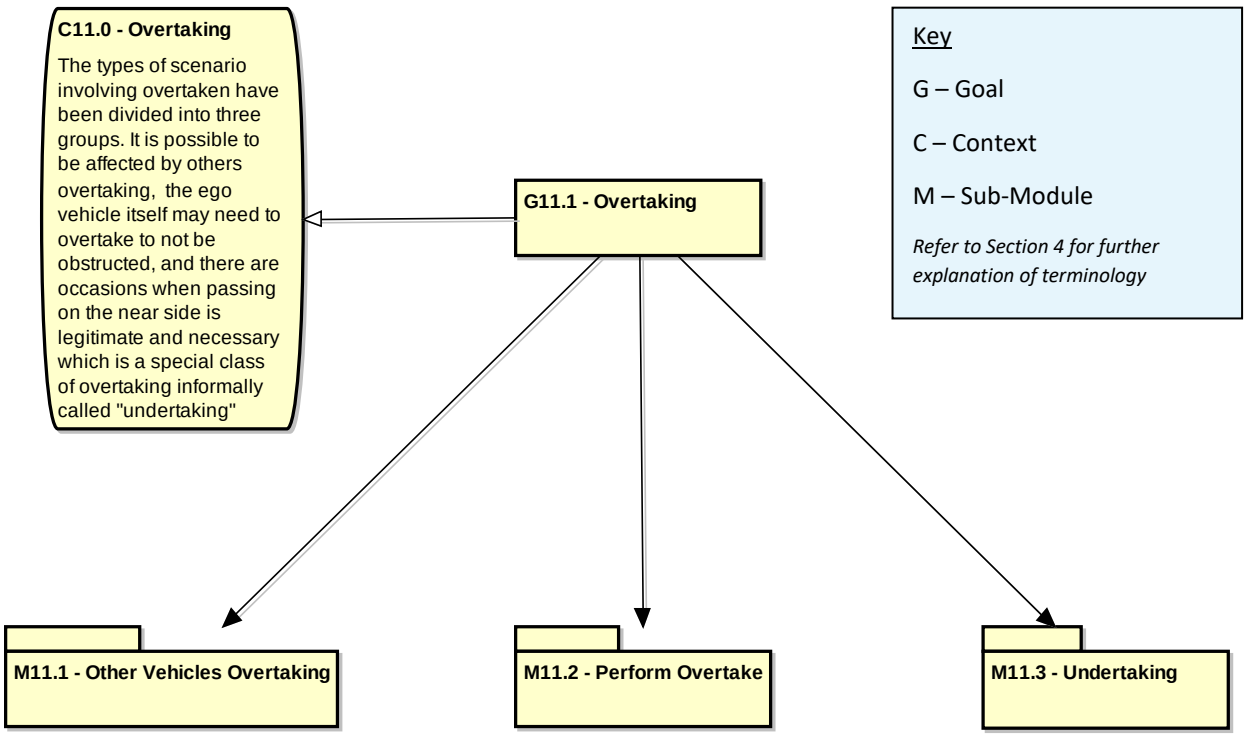


Figure 29 : Urban Pilot GSN – Overtaking Module Structure

The sub-modules are shown in Figure 30 to Figure 32.

Key

G – Goal

C – Context

A – Assumption

J – Justification

S – Strategy (yellow = singular, blue = one of several options)

Sn – Solution

Refer to Section 4 for further explanation of terminology

G11.1.2 - Other Vehicles Overtaking
The ego vehicle being overtaken where it involves temporarily crossing over or encroaching into a lane normally intended for travel in the opposite direction

C11.1.3 - Urban Overtaking
Overtaking may be less of an issue in an urban conext than with other higher speed routes, but still needs to be considered

G11.1.5 - Being Overtaken
Being overtaken where it involves temporarily crossing over into a lane normally intended for travel in the opposite direction

C11.1.4 - Interaction Risks
The new risks due to AV interaction with existing road users may be twofold. The first is AV behaviour (possibly more prudent slower progression rate) may induce more overtaking in general from manually driven vehicles. Secondly, the response to being overtaken (even if that is no response) may increase the risks to the overtaker, making it harder to abort or pull in for instance

G11.1.6 - Head-to-Head Scenario
Overtaker travelling in the same path (head-to-head) from the opposite direction.

J11.1.4 - Limit Head-on Closing Speeds
Limit the closing speed to on-coming traffic. This reduces the criticality and performance upon sensing requirements, but will increase the exposure time as the manouvre will take longer

S11.1.2 - Restrict Overtaking: Absolute
Where the manoeuvre requires encroaching upon an lane intended for use by vehicles travelling in the opposite direction at >20 mph, for all vehicles, restrict overtaking to passing stationary and slow moving obstructions (<=10 mph) and enforce that a slow speed (e.g. max 20 mph) manoeuvre is performed

J

S11.1.3 - Attempt to Respond Appropriately
Attempt to behave appropriately with the caveat that the 'overtaker' takes responibility for the additional risk of overtaking a vehicle which may act with a lower contextual awareness

S11.1.4 - Take No Action
Take no action with the acceptance that the number of head-on collisions may increase until the AV fleet significantly out numbers the manually driven fleet

S11.1.5 - Situation Recognition
Recognise when the vehicle in front is approaching in the same lane and reduce speed accordingly to mitigate a collision

A11.1.1 - Breaking the Rules
As with all rules of the road, many motorists will intentionally break them and emergency vehicles may also do so legitimately. Consideration would need to be given to the risks (frequency and severity) in these cases

A

S11.1.8 - Restrict Overtaking: Relative Speed
Restrict overtaking to where an unreasonably (to be defined quantitatively, e.g. 10 mph) large differential speed exists between the lead vehicle and the speed limit

J11.1.5 - System Performance Reliance
The ego vehicle's detection system must be able to locate itself and estimate the position of the approaching vehicle on its map to understand if they are using the same lane and are potentially on a collision course. This places requirements on the performance of mapping, localisation, range and velocity estimation of the approaching vehicle

J

G11.1.12 - Sensing Capability
The performance of the sensing and localisation performance is sufficient to work with the braking strategy

G11.1.13 - Braking Strategy
The braking strategy adopted is effective and workable along side the sensing capability

Sn11.1.5 - Track Tests
Use proving ground controlled conditions to verify and validate the response performance

G11.1.9 - Sensing Performance and Criticality
The sensing performance and functional integrity of the sensing and processing system on the ego vehicle for detection of on-comming traffic meets that required for the selected strategy (and the resulting hazard exposure) for handling other overtaking vehicles

G11.1.10 - Collisions Remain Within Limits
Collisions resulting from AVs being overtaken do not reach unacceptable levels

Sn11.1.1 - Sensing Design
Sensing system design and test documents, (SRD, FMEA)

Sn11.1.2 - Overtaken HAZOP
Sensor system HAZOP section for being overtaking

Sn11.1.3 - Track Tests
Proving Ground tests where a testable strategy is used

Sn11.1.4 - Statistical Monitoring
Monitoring of accidents to ensure AVs are not inducing an unreasonable increase in overtaking related accidents

Figure 30 : Urban Pilot GSN – Other Vehicles Overtaking sub-module (UP11.1)

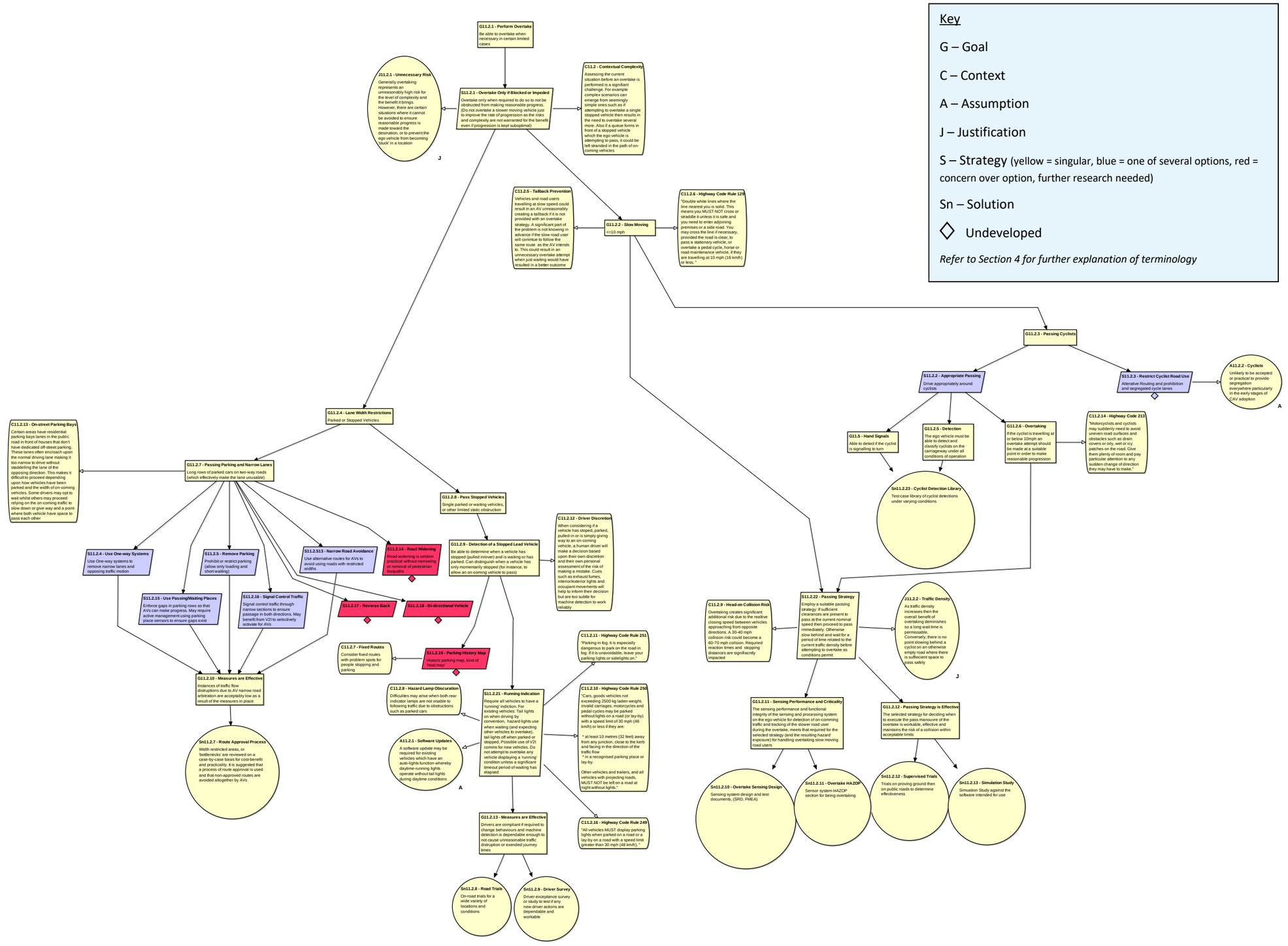


Figure 31 : Urban Pilot GSN – Perform Overtake sub-module (UP11.2)

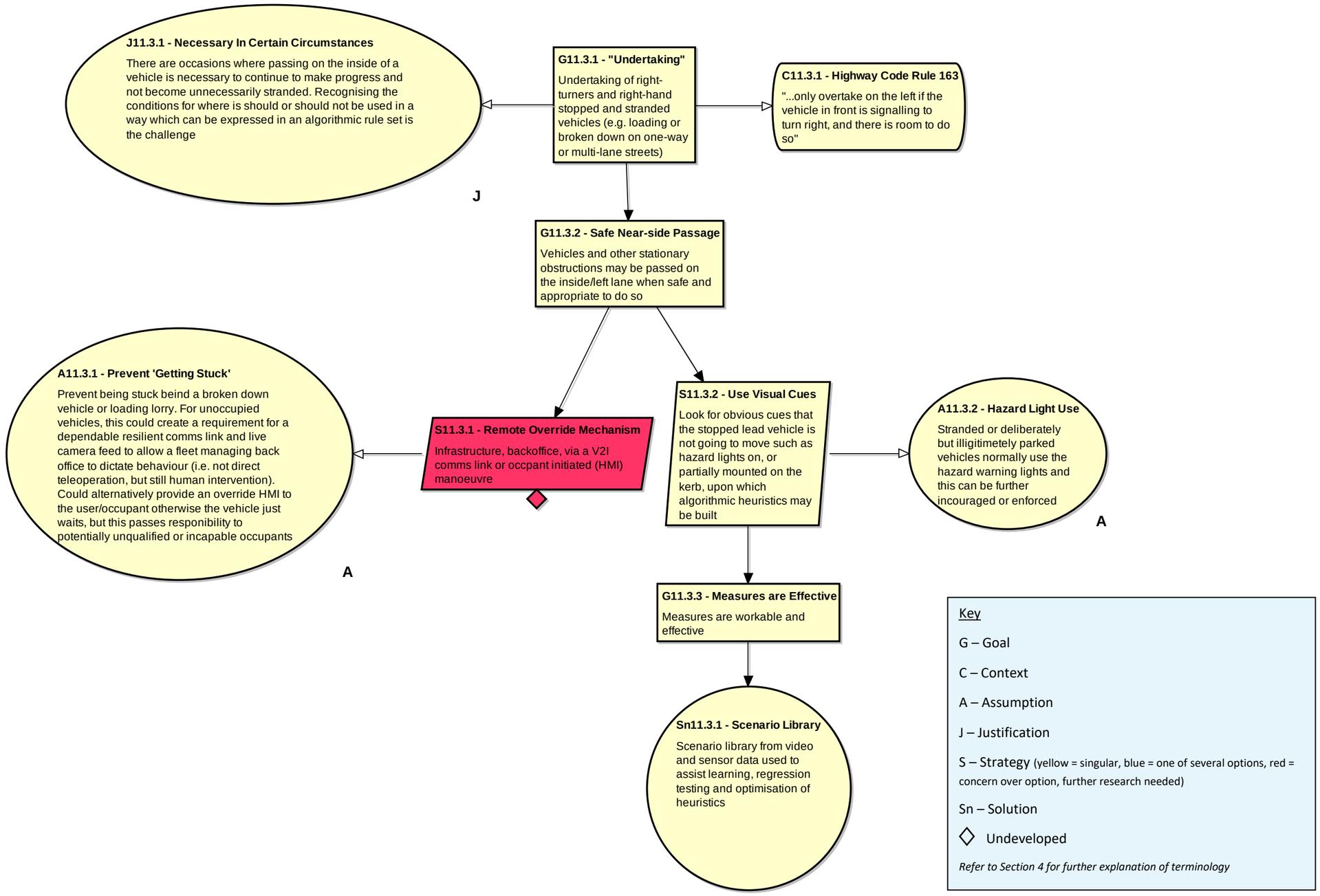


Figure 32 : Urban Pilot GSN – Undertaking sub-module (UP11.3)

6.8 Antisocial Behaviour (UP12)

The Antisocial Behaviour module covers some of the operational issues which should fall outside normal engineering considerations but which may throw up some real-world practicalities that cannot be avoided and may in turn require solutions to be engineered. The change of use from an owner driven vehicle to a possibly fleet owned and operated vehicle which can be operated unoccupied presents many new opportunities for people to cause disruption. There is a process of depersonalisation that may take place when firstly there is no longer a person driving the vehicle to offend and secondly, it may not even have anyone in it to directly observe the treatment of it from outside. From impatient drivers and pedestrians unafraid to force an empty vehicle to a halt, there are more concerning issues of criminal uses of the vehicle which might have been facilitated by its automation. Many of these concerns are related to the unoccupied running of the vehicle, but there are also potentially valid concerns that an automated vehicle may have personal security issues, the main fear being that drivers will be demoted to helpless and hapless passengers vulnerable to attack. This arises from the fact that an urban AV could be stopped and surrounded by any group of people who may have malicious intentions towards its passengers. With existing vehicles, assailants run the risk that the target vehicle may just not stop, or if it does stop, the driver may proceed to force the vehicle past to escape what they perceive to be a dangerous situation. Instructing an automated vehicle to do likewise quickly becomes problematic if not also unlawful. Many of the issues addressed are covered by laws and could be left to existing law enforcement methods, or the enforcement of laws that have been created or updated accordingly. However, laws are often not enough of a deterrent as people usually do not intend to get caught and there is the burden and practicality of proof since identifying the perpetrators of the identified antisocial behaviour after the fact may be a significant challenge such that, to those who may be affected, prevention is better than cure or retribution.

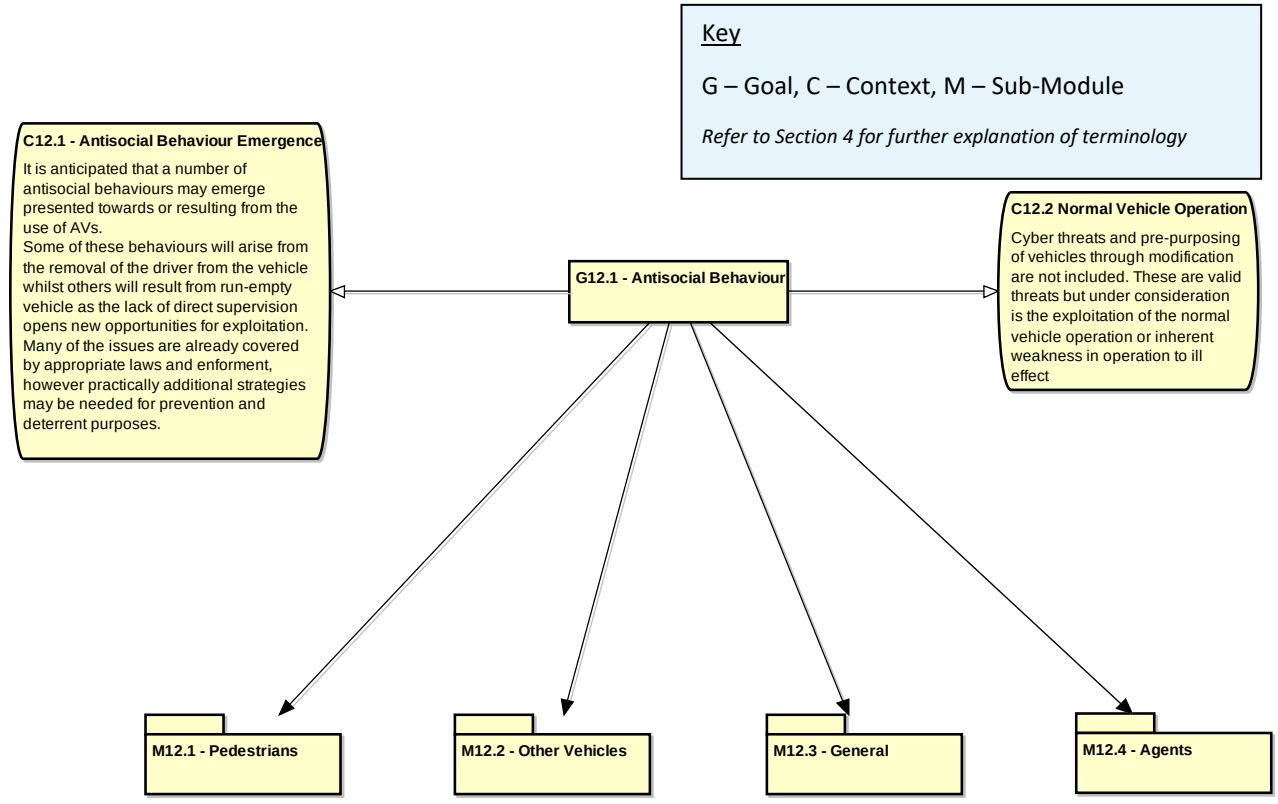


Figure 33 : Urban Pilot GSN – Antisocial Behaviour module (UP12)

The Antisocial Behaviour module structure is shown in *Figure 33*. The structure was derived from common the themes emerging from the issues mind map shown in *Figure 34*.

The Pedestrians sub-module covers the issue of wilful obstruction of the ego vehicle by pedestrians and bystanders or even people with malicious intentions. This can range from people who may feel some level of indignation towards moving out of the way for a machine, impatience to wait to cross, particular from the perspective of a crowd surge where safety in numbers may prevail, though to a deliberate surrounding of the vehicle by people who may wish to try to force entry into it, knowing that it will not run them over. Once fairly benign measures have been exhausted such as horn equivalents, play a reminder jingle or please move audio message there needs to be an approach to proceeding. A human driver would use due diligence to move forward based on the urgency of the situation and possible use of gestures or eye contact. The proposed solutions are to emulate human driving behaviour by providing a nudge forward mode for crowd conditions and a push forward mode for more serious situations which may threaten the personal safety of the occupants. Obviously, the vehicle cannot be programmed to force its way past people if there is a risk of injury to them. The Nudge forward mode would use a creep speed and sensing to detect if people became too close to the vehicle. With the provision of this additional sensing it is conceivable that it could still be allowed when the vehicle is unoccupied as anyone refusing or unable to move would still ultimately prevent the ego vehicle from going any further in that direction. The Push Forward mode is potentially more controversial and it is anticipated that it would be only be appropriate and permitted for use in dangerous or threatening situations for which the vehicle user would be required to take responsibility for. It would therefore have to be initiated and modulated by a vehicle occupant and would not apply to unoccupied vehicles. In effect a limited manual override supported by vehicle systems and monitoring which could possibly provide evidence to support its use after the fact. The level to which the function could be allowed to physically force people out of the way is laid out in the GSN. It could be supported by sensing and trappage prevention/detection measures to prevent it being used to inflict deliberate harm (the other side of the argument addressed in the Agents sub-module under vehicle misuse). The mode is effectively for emergency escape but carries with it equivalent risks that it could be intentionally or unintentionally misused. Measures put in place to help prevent it could cause someone to end up crushed under the vehicle's wheels or underbody which could also limit its effectiveness in the most severe of scenarios which is another ethical and legal trade off to be considered.

The Other vehicles sub-module covers the anticipation of issues arising from interaction with conventionally driven vehicles. As with the rest of the module, much of the problem could be attributed to the depersonalisation of interacting with a machine rather than a person. The machine cannot take offense at being forced to break or the right of way not being respected, particularly if also there are no vehicle occupants to be perturbed. The sub-module breaks up the potential behaviours into aggressive driving, testing for weaknesses (in order to exploit), herding and trapping, as well as induced collisions. Testing for weaknesses involves trying to trigger or induce behaviours that might be used later for ill effect, but might in itself be dangerous as an activity if the limits are pushed too far or the driver(s) involved lose sight of the effects of their own driving whilst concentrating on perturbing the ego vehicle. Motivations may vary from pranks to looking for some future advantage to gain position in traffic flow, "how late can I leave it to jump out on an AV?" for example. Herding or trapping are both similar to testing and may result out of testing, but could have serious consequences if it is found to be able to effectively force a route change of the ego vehicle which could form part of a wider criminal activity. Mitigations and prevention for such activities is difficult beyond trying to detect and report foul play with the ability to do so perhaps acting as a deterrent to try in the first place. Finally, induced collisions used in much the same way are they already are, for compensation claims, but with the added incentive that there may not be a human witness in the ego vehicle and that there is a control system to be falsely accused resulting in the so-called 'Autonomous Scapegoat'. Inducing a collision with an AV might be done to discredit the general use of AVs or again just for financial gain.

The General sub-module covers general interference such as vandalism (of the kind which effects the vehicle's operation, not cosmetic/graffiti), deliberate obstruction and disablement attempts as well as pranks conducted for amusement or the irony of exploiting some aspect of there not being a human driver. Pranks may include 'playing

chicken', by seeing who is brave enough to jump into or out of the way of the ego vehicle at the shortest moment. Also, attempting to attach items or tow another vehicle, hanging on to the outside for a 'free ride' as well as trying to surf or balance on a moving vehicle for peer or social media notoriety. Finally, when the vehicle is unoccupied, particularly if obviously fleet operated (thus depersonalised from an individual owner or potential victim), there may be the temptation by some to stop it in its path and attempt to overturn it or displace it to a new location to see how it copes or to create a photo opportunity by placing it in a place or position of some novelty. Many of these are not entirely new issues, but are perhaps made easier or more tempting when there is nobody inside the vehicle. They could still be left to regular law enforcement, but the lack of witnesses may make it less of a deterrent and harder to pursue after the fact, with prevention being better than punitive measures. Reporting of events via a control centre might be a viable option to deter these kinds of pranks that do not have any serious criminal intentions beyond property damage or risking injury.

The Agents sub-module provides coverage for where the AV is being utilised directly to facilitate some criminal or unwanted behaviour. It further breaks down such activity into terrorism, as a transport mode for general criminality, as a mobile surveillance device, and where the vehicle itself is used either as a weapon or to try to create some form of accident for which it can then be blamed.

Taking terrorism first, car bombs are not new, but there is a new potential to be able to send one or more vehicles to a location to deploy some explosive device or chemical/biological agent, or to just create congestion near a targeted area to amplify the effects of a separate act of terrorism. As far as transport for general crime is concerned, taxis and so-called minicabs already present traceability issue for police as assailants can commit a crime then use a minicab to leave the scene of a crime whilst paying cash so that there is no payment record. The rollout of AVs presents a new opportunity for a convenient escape with potentially even less traceability. In addition, vehicles could be easily equipped with cameras (for example 'dash cams', action cameras and smart phones) then sent to a location to monitor or stalk either persons or places of interest for future crimes. Deliberate misuse of the vehicle by occupants wishing to inflict harm remains a possibility depending upon how the system's overrides (if any) have been implemented. If it is possible for the users to have some form of manual control, then these could potentially be exploited to cause harm. The possibility of user induced accidents (for so-called scapegoating) beyond the use of basic low-speed overrides has not been developed, since by definition of the feature this should not be readily possible without modification to the vehicle which is also out of the scope of consideration. Countermeasures to these types of threats and issues are difficult and will depend on the seriousness and likelihood of occurrence for the intended region of use. Where there is not a strong possibility of frequent occurrence then it could be left to regular law enforcement. If more pre-emptive or preventative measures are required, then active fleet monitoring, spot checks and similar impositions may be required and will have to be considered in balance with the possible infringement of civil liberties and the right to personal privacy.

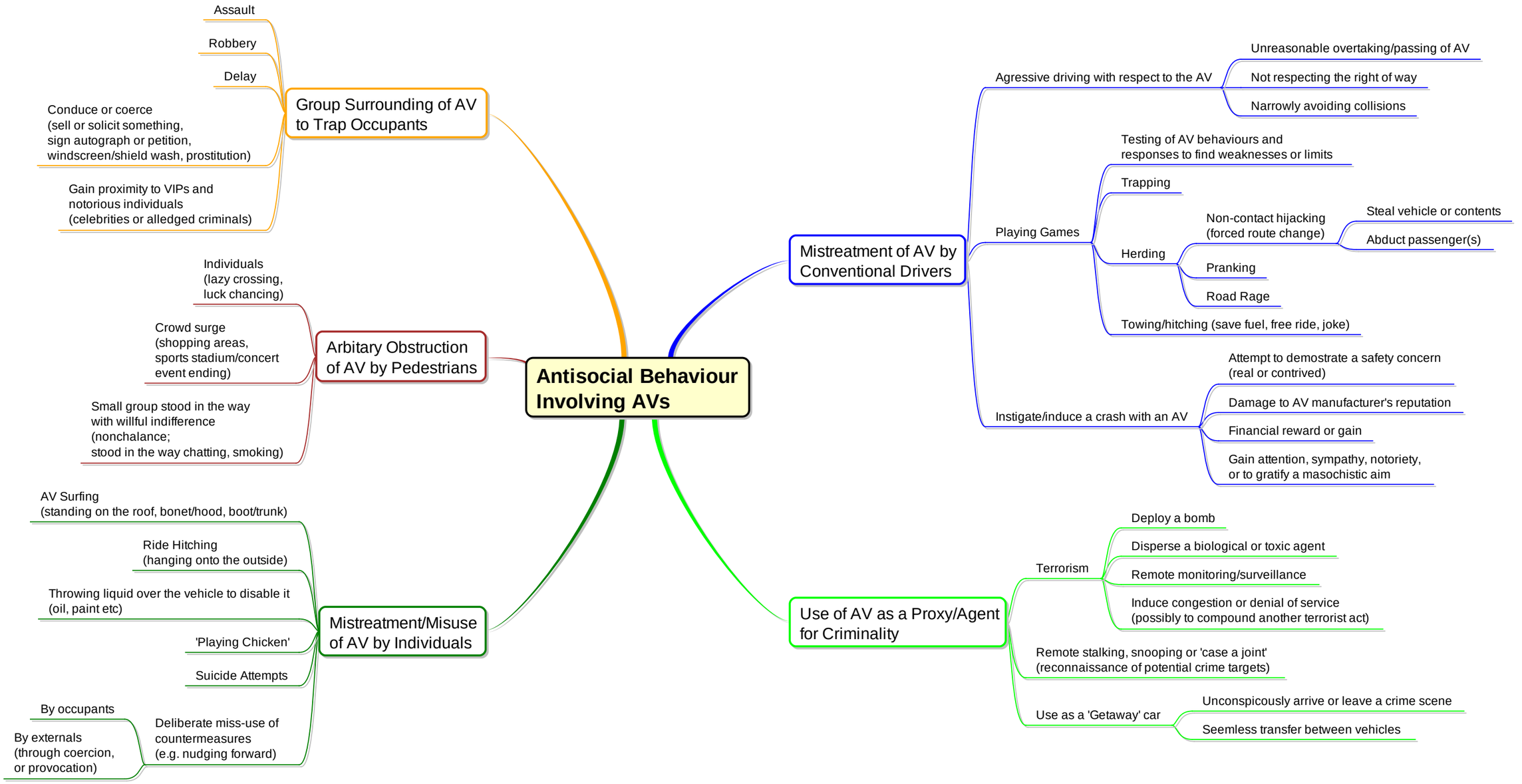
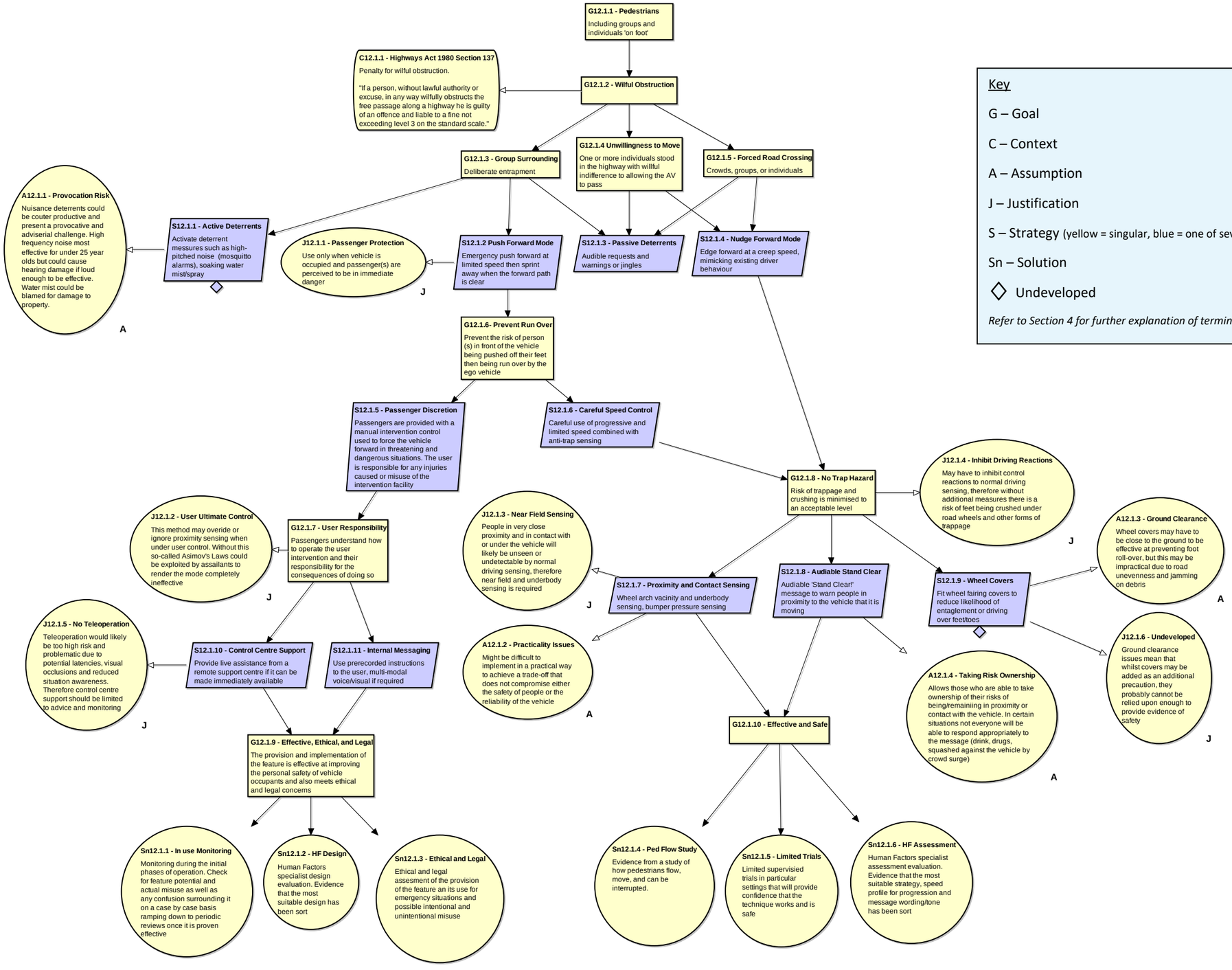


Figure 34 : Mind Map for Antisocial Behaviour Involving AVs



Key

G – Goal

C – Context

A – Assumption

J – Justification

S – Strategy (yellow = singular, blue = one of several options)

Sn – Solution

◇ Undeveloped

Refer to Section 4 for further explanation of terminology

Figure 35 : Urban Pilot GSN – Antisocial Behaviour Pedestrians sub-module (UP12.1)

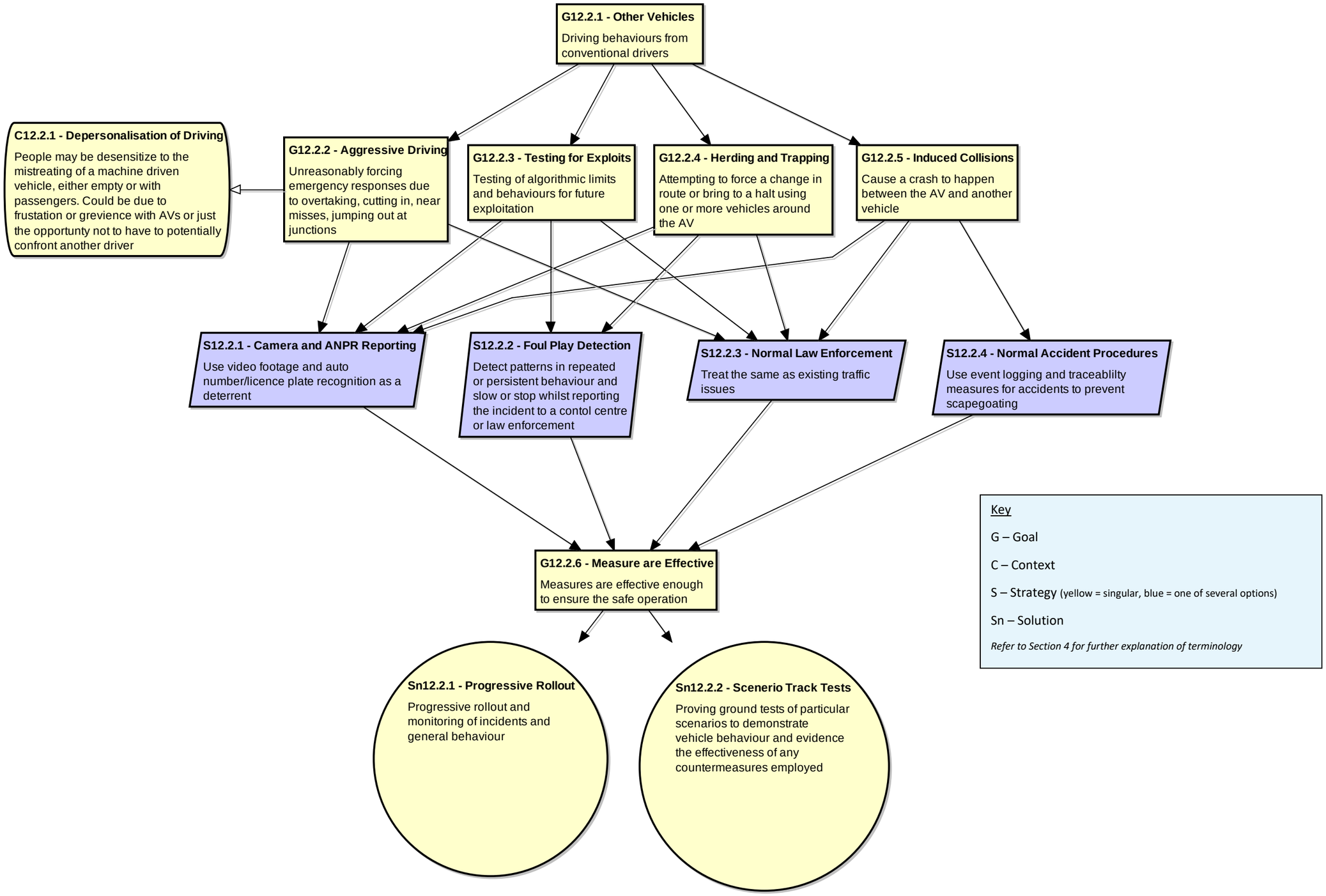


Figure 36 : Urban Pilot GSN – Antisocial Behaviour Other Vehicles sub-module (UP12.2)

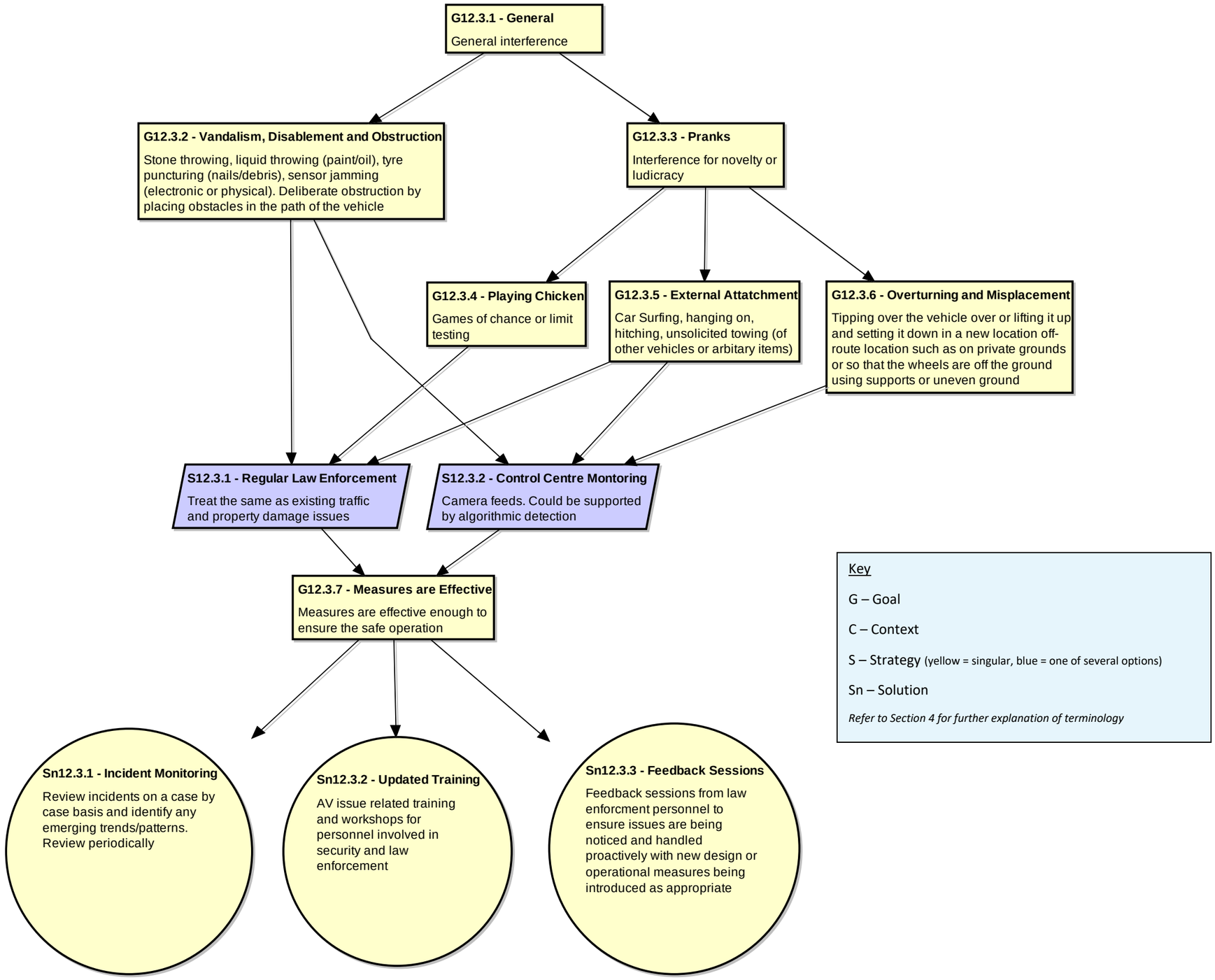


Figure 37 : Urban Pilot GSN – Antisocial Behaviour General Interference sub-module (UP12.3)

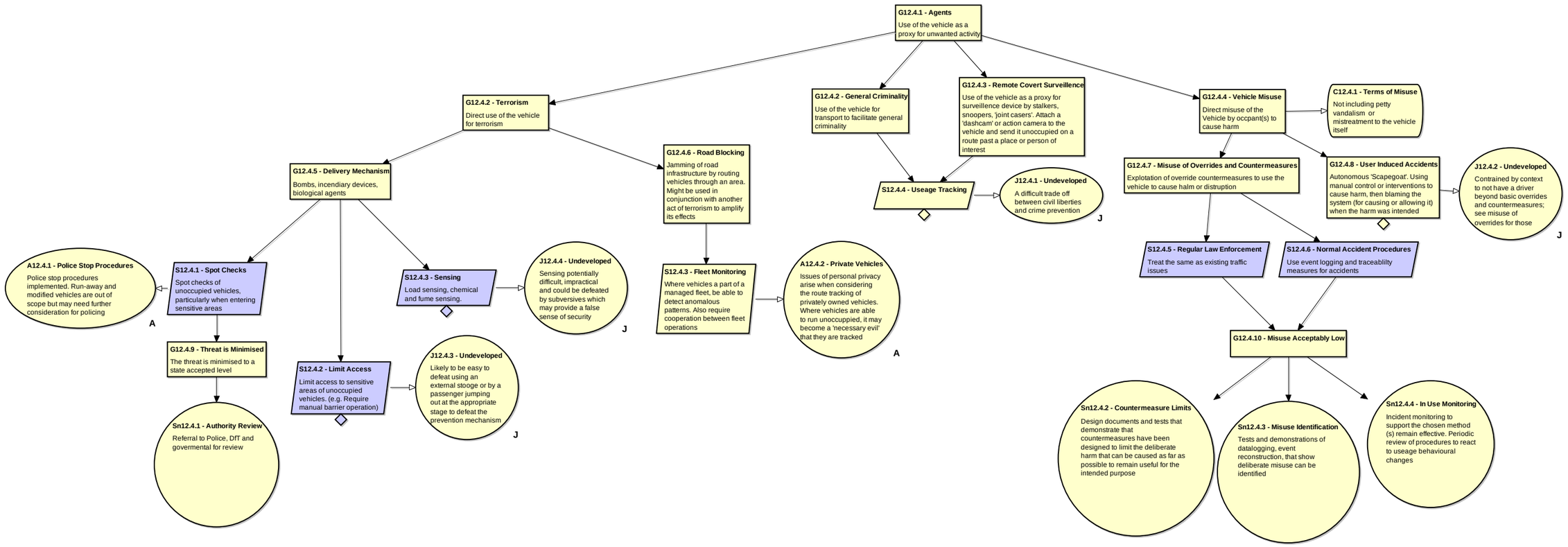


Figure 38 : Urban Pilot GSN – Antisocial Behaviour Agents sub-module (UP12.4)

7 Summary and Concluding Thoughts

7.1 General

It is hoped that this report and the GSN approach will assist with the thought process of understanding the complexities surrounding CAV development and subsequent rollout. There are many challenging situations which arise day to day when driving on the roads which also need to be addressed for vehicles if they are to be operated autonomously. The exercise of strategy selection (pruning) has been deliberately left as an exercise to the reader, to apply ALARP principles. Despite this, in many cases it appears that when the ALARP point has been reached the residual level of risk may still be far too high without resorting to infrastructure support. This is however an issue of what levels of residual risk are acceptable, which is a question for governments and society to answer. A common theme is the need for a robust and resilient digital mapping and vehicle localisation system, without which road casualty rates will be dependent upon image processing and object recognition techniques.

7.2 Roads and Regulation

The following are concluding thoughts and suggestions in relation to roads and regulation:

- It is envisaged that some road types will be challenging for fully automated vehicles for many years to come. Narrow roads, two-way single carriageway high speed rural roads, complex junctions and areas with many road user types could be amongst the most challenging. Restricted use (e.g. geo-fencing) may be required to allow the benefits of automation to be realised more quickly along roads which are more manageable for AVs.
- Type Approval for AVs could be focussed on conformance with particular expected vehicle behaviours (e.g. overtaking and negotiating junctions) that are currently left to the discretion and due diligence of human drivers.
- Consideration should be given to a national mapping infrastructure. This could be licensed to private sector providers in a similar manner to a mobile phone network, but would give a single source of the truth. OEM and technology providers could layer their own proprietary features on top on a national base map to provide additional brand defining functionality.
- Consider developing a road and zone suitability assessment procedure for AVs alongside a classification system for which vehicles could be certified as being capable of automated operation against a particular road or zone class.

7.3 Vehicles and Technology

The following are concluding thoughts and suggestions in relation to the vehicles and associated technologies:

- It is a known human factors trait that people are poor at supervising automation, especially when it infrequently fails. The idea of the handover of a moving vehicle for general drivers is gradually being accepted as a bad idea despite it being a convenient solution from the point of view of system development and evolution as well as confidence building in the technology.
- From this, it seems that a step change from partial to full automation is needed rather than a gradual evolution which relies on a driver to intervene when the system fails.

- Weaknesses in processes associated with automated driving, such as estimation, interpretation and prediction, can have consequences comparable to those of hardware and software failures in traditional safety of automation, and the seriousness of those weaknesses should not be underestimated. This is not currently addressed by ISO 26262.
- Sensor fusion is often only as good as the weakest sensing method and having a mixed bag of sensors does not guarantee sufficient coverage for all situations.
- Localisation systems may fuse several data sources to plug coverage gaps (where the vehicle cannot locate itself - GPS/GNSS denied areas or a lack of local landmark features). These gaps and exactly how they will be managed for all times and locations will need to be understood better. For highly automated driving determinism is preferable to a carousel of weaker solutions that may be switched in with the hope that one will always be available. Heterogeneous redundancy (using multiple location data sources) does not always provide more resilience, and may serve to falsely increase confidence though obfuscation of its failure modes
- The connectivity and automation technical communities need to work closely together and converge their efforts to avoid a connectivity roll-out which only serves to provide driver assistive information and is not of any real use to providing an infrastructure for failsafe automation.
- A system engineering approach to vehicles and infrastructure is needed, rather than stand-alone vehicle or infrastructure centric approaches

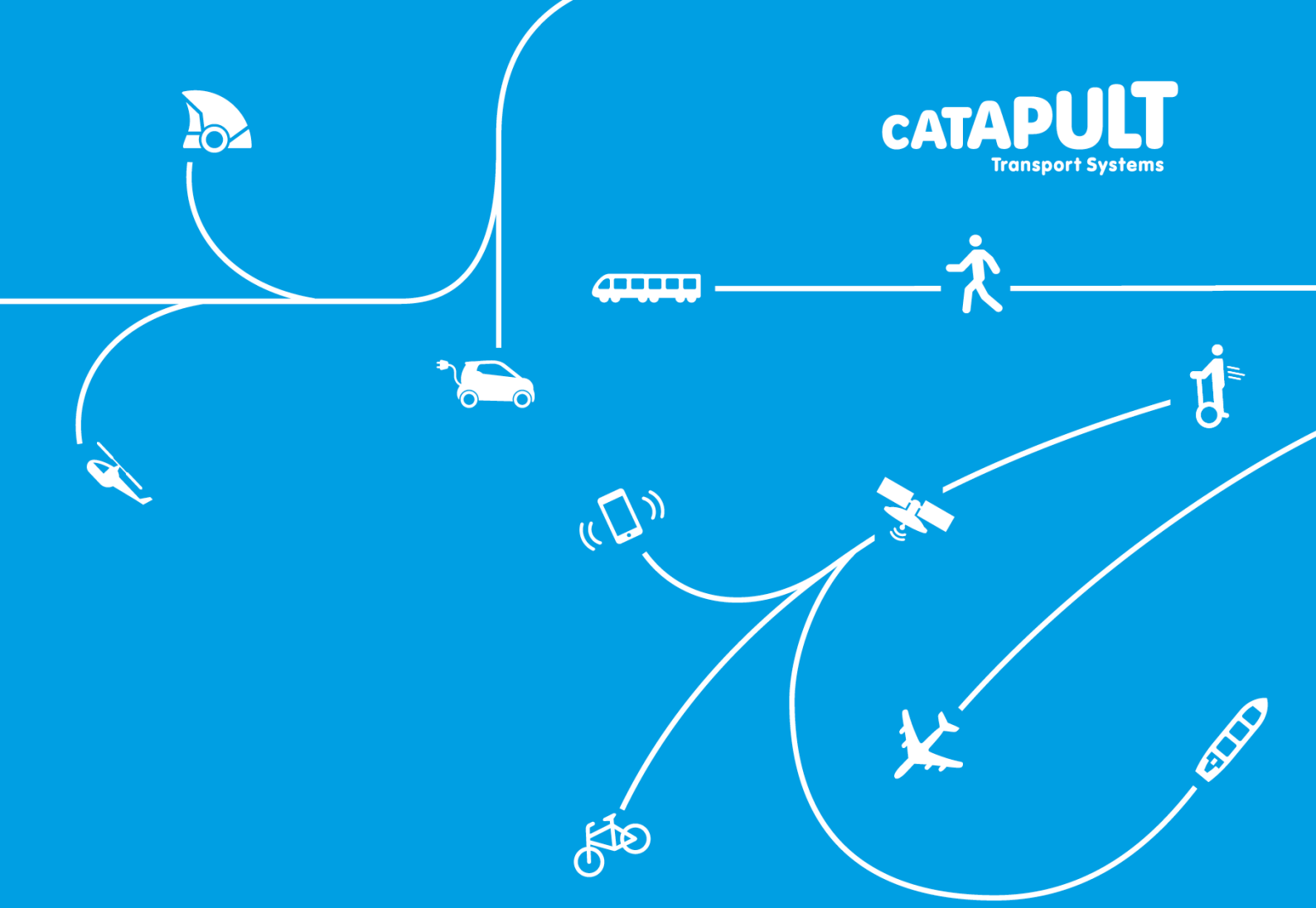
7.4 Standards, Ethics, and Safety Case

The following are concluding thoughts in relation to standards, ethics and safety case:

- Standards normally follow practice, but ISO 26262 was a slight exception in that practice in the industry was to some extent formed from it.
- Current standards cover fault-failures but not system insufficiencies.
- There is a need to address the standards gap related to insufficiencies and safety of the intended functionality.
- There needs to be a conceptual model of the operation and safety case for automated vehicles to form the consensus of practice from which the standards gap can be filled, rather than speculatively forming standards in the hope they will be adopted as the de facto approach. This could greatly assist in the forward progress with the development and roll-out of CAVs by breaking the competitive stalemate that has arisen from the expectation that a fully functional standalone CAV platform will just emerge from one of the technology providers.
- There is a paradox between traditional approaches to the design of hardware and software intended for preventing hazards high risk situations, and the design complexity required to cope with the chaotic random events of real-world driving. This arises from removing the driver as a fall back and arbiter to which responsibility and liability has been previously placed.
- Current solutions do not appear conceptually commensurate with the inherent risks both from a hardware and a software perspective and are not currently implementable against current fault failure standards such as ISO 26262 before even considering algorithmic and sensing insufficiencies.
- An industry consensus or common understanding of the general safety case would help to focus design and vice versa since having some concept of the design is required to base the safety case

upon. Currently this approach is being resisted as it is seen as being anti-competitive despite the fact that the industry has a long history of eventually converging around common designs.

- There is a need for protection of data integrity across domains when relied upon by vehicles for safety.
- Consideration of the residual risks when an ALARP or SFAIRP strategy has been reached is needed and whether or not these risks are acceptable, and what else could be done if not.
- The GSN provided here could be evolved and unused strategies pruned. Preferred approaches could be further developed.



CATAPULT
Transport Systems

Transport Systems Catapult
The Pinnacle, 170 Midsummer Boulevard
Milton Keynes MK9 1BP

Tel: 01908 359 999

www.ts.catapult.org.uk
linkedin.com/company/transport-systems-catapult
Twitter: @TSCatapult