



中华人民共和国国家标准

GB/T XXXXX—XXXX

道路车辆 预期功能安全

Road Vehicles-Safety of the Intended Functionality

（征求意见稿）

（本草案完成时间：2022 年 4 月 29 日）

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX – XX – XX 发布

XXXX – XX – XX 实施

国家市场监督管理总局
国家标准化管理委员会 发布

目 次

前言 IV

引言 V

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 预期功能安全活动概述和组织 9

 4.1 总则 9

 4.2 预期功能安全的原理 9

 4.3 本文件的使用 13

 4.4 SOTIF 活动管理和支持过程 14

5 规范定义和设计 16

 5.1 目的 16

 5.2 功能规范的定义和对设计的考虑 16

 5.3 系统设计和架构的考虑 17

 5.4 性能局限和应对措施的考虑 18

 5.5 工作成果 19

6 危害的识别和评估 19

 6.1 目的 19

 6.2 总则 19

 6.3 危害识别 19

 6.4 风险评估 21

 6.5 残余风险接受准则的定义 22

 6.6 工作成果 23

7 潜在功能不足和潜在触发条件的识别与评估 23

 7.1 目的 23

 7.2 总则 24

 7.3 潜在功能不足与触发条件的分析 24

 7.4 预估系统对触发条件的响应的可接受度 28

 7.5 工作成果 28

8 修改功能以解决 SOTIF 相关风险 28

 8.1 目的 28

 8.2 总则 28

 8.3 改进 SOTIF 的措施 29

 8.4 更新“规范定义和设计”的输入信息 31

 8.5 工作成果 31

9 定义验证和确认策略 31

9.1	目的	31
9.2	总则	32
9.3	集成和测试的定义	32
9.4	工作成果	34
10	已知场景的评估	34
10.1	目的	34
10.2	总则	34
10.3	感知的验证	35
10.4	规划算法验证	35
10.5	执行的验证	36
10.6	系统集成验证	36
10.7	已知危害场景导致的残余风险的评估	37
10.8	工作成果	37
11	未知场景的评估	37
11.1	目的	37
11.2	总则	37
11.3	未知场景残余风险的评估	38
11.4	工作成果	39
12	SOTIF 成果的评估	39
12.1	目的	39
12.2	总则	39
12.3	评估 SOTIF 的方法和准则	40
12.4	SOTIF 发布推荐	40
12.5	工作成果	41
13	运行阶段的活动	41
13.1	目的	41
13.2	一般要求	41
13.3	与观察相关的主题	41
13.4	SOTIF 问题评估和解决流程	42
13.5	工作成果	43
附录 A (资料性)	预期功能安全的通用指南	44
A.1	用 GSN 构建 SOTIF 论证的示例	44
A.2	GB/T 34590 标准功能安全和本文件之间相互作用的说明	66
A.3	简化的预期功能安全 (SOTIF) 应用示例	73
A.4	规范定义和设计的简化示例	75
附录 B (资料性)	场景和系统分析指南	79
B.1	推导 SOTIF 误用场景的方法	79
B.2	SOTIF 安全分析方法的场景因素构建示例	81
B.3	用于识别和评估潜在触发条件和功能不足的安全分析的示例	88
B.4	在 ADAS 和自动驾驶车辆的 SOTIF 研究中应用 STPA 方法	99
附录 C (资料性)	预期功能安全验证和确认指导	104
C.1	验证和确认策略目的	104

C.2	确认目标的导出	104
C.3	预期功能安全适用系统的确认	111
C.4	感知系统的验证和确认	113
C.5	场景参数化及场景抽样指导	118
C.6	减少确认测试的考虑	122
附录 D (资料性)	关于 SOTIF 特定方面的指南	127
D.1	驾驶策略规范指导	127
D.2	对机器学习的建议	134
D.3	地图预期功能安全的考虑	139
D.4	V2X 功能安全的考虑	140
D.5	感知系统性能目标量化与常见传感器性能局限举例	142
D.6	OTA 更新的 SOTIF 考虑	143
附录 E (资料性)	风险接受准则示例	144
E.1	概述	144
E.2	风险接受准则示例	144
参考文献	147

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中华人民共和国工业和信息化部提出。

本文件由全国汽车标准化技术委员会（SAC/TC114）归口。

本文件起草单位：

本文件主要起草人：

引 言

道路车辆的安全是道路车辆行业最为关注的问题。车辆上包含的自动驾驶功能的数量逐渐增加，而这些功能依赖于由电气/电子(E/E)系统实现的感知、复杂算法进行处理和执行。

可接受的道路车辆安全水平指不存在因预期功能及其实现相关的任何危害而导致不合理的风险，包括由失效引起的危害和由规范定义不足或性能局限而引起的危害。

为了实现功能安全，GB/T 34590.1-XXXX将功能安全定义为不存在因电气/电子系统的功能异常表现引起的危害而导致不合理的风险。GB/T 34590.3-XXXX描述了如何开展危害分析和风险评估(HARA)以确定整车层面的危害及其安全目标。GB/T 34590-XXXX的其他部分提供了避免和控制可能违背安全目标的随机硬件失效和系统性失效的要求和建议。

对于一些E/E系统，例如依靠感知车辆外部或内部环境来建立态势感知的系统，尽管没有发生GB/T 34590中所提到的故障，但其预期功能及实现仍可能会导致危害行为，造成此类潜在危害行为的原因，举例来说，包括：

- 功能无法正确感知环境；
- 功能、系统和算法在传感器输入变化、融合策略或不同环境条件等方面缺乏鲁棒性；
- 由于决策算法和/或人的不同期望而导致的非预期的行为。

尤其是这些因素与使用机器学习的功能、系统和算法有关。

不存在因功能不足引起的危害行为而导致不合理的风险被定义为预期功能安全(SOTIF)。功能安全(GB/T 34590所指的)和预期功能安全在安全方面是互补的(见附录A.2，以更好地理解GB/T 34590和本文件的各自范围)。

为解决SOTIF问题，在以下阶段实施消除危害或降低风险的措施：

- 规范定义和设计阶段；

示例1：在开展SOTIF活动的过程中，通过识别出的系统不足或危害场景，对车辆功能或传感器性能要求进行修改。

- 验证和确认阶段；

示例2：技术评审、对相关场景具有高覆盖率的测试用例、潜在触发条件的注入、选定SOTIF相关场景的在环测试[例如：软件在环(SIL)、硬件在环(HIL)、模型在环(MIL)]。

示例3：车辆长期道路测试，车辆场地测试，仿真测试。

- 运行阶段；

示例4：SOTIF事件的现场监控。

这些危害可被场景中的特定条件(即：触发条件)触发，触发条件可包括预期功能的合理可预见的误用。此外，与其他整车层面功能的交互可能会导致危害(例如，在自动驾驶功能激活时，触发驻车制动)。

因此，用户正确理解功能及其行为和局限性(包括人机交互)对于确保安全至关重要。

示例5：驾驶员在使用2级驾驶自动化系统时注意力不集中。

示例6：模式混淆(例如，当某功能未激活时，驾驶员认为该功能已激活)可能直接导致危害。

注1：合理可预见的误用不包括对系统运行的故意更改。

基础设施提供的信息(例如，V2X、地图)，如果可能对SOTIF产生影响，也将作为评估功能不足的一部分。见D.4中关于V2X功能的指南。

示例7：对于自动代客泊车系统，路径规划功能和目标探测功能可由基础设施和车辆共同实现。

注2：根据应用情况，在评估SOTIF时，基于其他技术的要素可能是相关的。

示例8：传感器在车辆上的位置和安装可能与避免因振动而导致的传感器输出噪声有关。

示例9：在评估摄像头传感器的 SOTIF 时，挡风玻璃的光学特性可能是相关的。

针对E/E系统的随机硬件故障和系统性故障（包括硬件和软件故障），在GB/T 34590中给出了降低风险的指导。

本文件中提到的功能不足可能被认为是系统性故障，然而，应对这些功能不足的措施是本文件特有的，同时也是对GB/T 34590中所述措施的补充。GB/T 34590假定预期功能是安全的，且针对的是E/E系统故障，这些故障可导致预期功能的偏离而产生危害。系统及其要素的要求的获取过程可来自此两项标准的相关方面。

表1说明了危害事件的可能原因与现有标准的映射关系。

表1 不同标准应对安全相关主题的概览

危害来源	危害事件的原因	范围
系统	电气/电子系统故障	GB/T 34590
	功能不足	本文件
	不正确和不充分的人机交互（HMI）设计（不恰当的用户态势感知，例如，用户感到困惑、用户负担过重、用户注意力不集中）	本文件 欧洲人机交互原则声明 ^[1]
	基于人工智能算法的功能不足	本文件
	系统技术	特定标准
	示例：激光雷达光束对眼睛的伤害	示例：IEC 60825
外部因素	用户或其他道路参与者的合理可预见的误用	本文件 GB/T 34590
	利用车辆安全漏洞进行攻击	ISO/SAE 21434或SAE J3061
	智能基础设施和/或车辆与车辆通信以及外部系统的影响	本文件 ISO 20077、GB/T 34590、GB/T 20438
	车辆周围环境的影响（例如，其他用户、非智能基础设施、天气、电磁干扰）	本文件 GB/T 34590、ISO 7637-2、ISO 7537-3、ISO 11452-2、ISO 11452-4、ISO 10605及其他标准

道路车辆 预期功能安全

1 范围

本文件提供了用于确保预期功能安全（SOTIF）的措施的通用论证框架和指南。预期功能安全指不存在因预期功能不足引起的危害而导致不合理的风险，功能不足包括：

- a) 整车层面预期功能规范定义的不足；或
- b) 系统中电气/电子要素实现的规范定义不足或性能局限。

本文件为实现和保持SOTIF所需的相关设计、验证和确认措施以及在运行阶段的活动提供指导。

本文件适用于依靠复杂传感器和处理算法进行态势感知且感知的正确性会对安全产生重要影响的预期功能，特别是紧急干预系统的功能和驾驶自动化等级为 L1-L5 的系统的相关功能。

本文件适用于安装在除轻便摩托车外的量产道路车辆上的包含一个或多个电气/电子系统的预期功能。

合理可预见的误用属于本文件的范围。此外，如果远程用户对车辆的操作或协助、或与可影响车辆决策的后台间的通信，可能导致安全危害，也属于本文件的范围。

本文件不适用于：

- GB/T 34590 涵盖的故障；
- 信息安全威胁；
- 因系统技术直接导致的危害（例如，激光雷达光束对眼睛造成的伤害）；
- 与触电、火灾、烟雾、热、辐射、毒性、易燃性、反应性、能量释放等相关的危害和类似的危害，除非危害是直接由电气/电子系统的预期功能引起的；
- 被视为功能滥用的、明显违反系统预期用途的故意行为。

对于已具备成熟且可靠的设计及验证和确认（V&V）措施的现有系统（例如，动态稳定性控制系统、安全气囊）的功能，不在本文件的预期使用范围。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 34590-XXXX（所有部分） 道路车辆 功能安全（ISO 26262:2018，MOD）

3 术语和定义

GB/T 34590.1-XXXX界定的以及下列术语和定义适用于本文件。

3.1

接受准则 acceptance criterion

表征不存在不合理风险(3.23)水平的准则。

注1：接受准则可以是定性的或定量的，例如，当某特定行为被认为是危害行为时所对应的物理参数，每小时的最大事件数，最低合理可行风险水平(ALARP)等。

示例1：从交通统计数据得到合理风险水平是每 X 千米发生一次事故。

示例2：与同等整车层面的影响（在使用中已证明驾驶员对该影响可控）进行比较，可支持接受准则的定义。例如，由非预期的车道保持辅助功能的干扰而引起的轨迹扰动，可类比为横风，以定义该功能的接受准则。

注2：预期功能安全的接受准则包含两个层面：

- a) 第一层：判断车辆行为是否属于危害行为而可能导致危害事件的准则，即危害行为接受准则；
- b) 第二层：判断车辆运行过程中残余安全风险是否处于合理水平的准则，即残余风险接受准则。

示例3：在使用预期功能安全接受准则进行评估过程中，可首先针对车辆某一行为或事件进行评估，判断是否构成有风险的危害行为或危害事件；然后针对一定行驶里程或行驶时间内出现的危害行为可能导致的残余危害风险进行评估，判断是否可以接受。见图 1 所示。

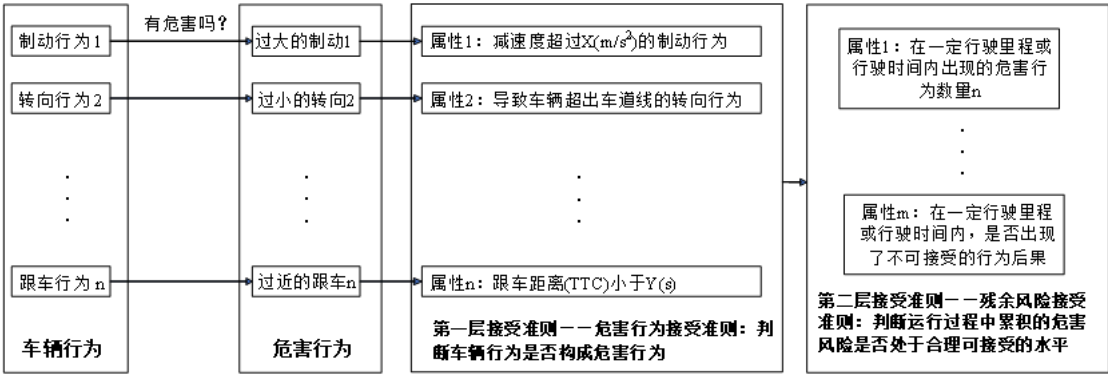


图1 双层接受准则的评估

注3：双层接受准则的指标值可能受多方面因素（如：技术演变、场景变化、交通文明程度提升等）的影响，因而在应用过程中也可能伴随着接受准则及其指标值的变化。

注4：在考虑特定功能场景的风险时，双层接受准则的理念可能存在融合，例如：通过基于一定时间内、某一场景下人类驾驶员操作行为的统计，来定义自动驾驶功能的危害行为接受准则。但这并不影响使用双层接受准则对预期功能安全的评价。

3.2

行为 action

场景快照 (3.27) 中任何参与者所实施的单一行为。

注1：行为/事件 (3.7) 的时序和场景快照是场景 (3.26) 定义的一部分。

示例：自行车 (3.6) 点亮危害报警灯。

注2：在该术语定义中，参与者可是人员、另一目标、另一系统或与所考虑功能相互作用的任何要素。

3.3

驾驶策略 driving policy

定义整车层面可接受的控制行为 (3.2) 的策略和规则。

3.4

动态驾驶任务 DDT dynamic driving task; DDT

在交通中车辆运行所需要的实时操作和决策功能。

注：以下功能属于部分动态驾驶任务：

- 车辆侧向运动控制（操作）；
- 车辆纵向运动控制（操作）；
- 监控驾驶环境（操作和决策），及执行对目标与事件 (3.7) 的响应（操作和决策），见 OEDR (3.20)；
- 运动规划（决策）；
- 通过照明、发信号及打手势等增强可见性（决策）。

3.5

动态驾驶任务后援 DDT fallback

当发生失效、探测到功能不足(3.8)或探测到潜在危害行为后，为执行动态驾驶任务(3.4)或过渡到最小风险状态(3.16)而由驾驶员或驾驶自动化系统做出的响应。

示例：驶离 ODD(3.21)或传感器被冰雪遮挡可导致危害行为，此时要求驾驶员做出响应。

3.6

自车 ego vehicle

SOTIF(3.25)分析时考虑的、装备了相关功能的车辆。

3.7

事件 event

在某一时刻所发生的事情。

注1：行为(3.2)/事件的时序和场景快照(3.27)是场景(3.26)定义的一部分。

注2：虽然每个行为也是一个事件，但不是任何事件都只是一个行为，即行为是事件的子集。

示例1：树倒在车辆前方 50 米的街道上。

示例2：交通灯在给定的时间变绿。

3.8

功能不足 functional insufficiency

规范定义不足(3.12)或性能局限(3.22)。

注1：功能不足包括整车层面或系统中电气/电子要素层面的规范定义不足或性能局限。

注2：SOTIF(3.25)活动包括对功能不足的识别及对其影响的评估。根据定义（见3.12和3.22），功能不足会导致危害行为或无法防止、探测及减轻合理可预见的误用(3.17)。在促成危害行为或无法防止、探测及减轻合理可预见的误用的能力尚未被建立时，可使用术语“潜在功能不足”。

注3：图2-图4描述了SOTIF因果模型，涵盖以下关系：触发条件(3.30)，功能不足，输出不足，危害行为，无法防止、探测及减轻合理可预见的间接误用，危害(3.11)，危害事件(3.7)和伤害。

注4：对于有助于伤害发生的间接误用，通常涉及两类功能不足。一类是导致系统在触发条件下产生危害行为的功能不足，另一个类是导致无法防止、探测及减轻合理可预见的间接误用的功能不足。见图2-4。

示例：一辆配备了 L2 级高速公路驾驶辅助功能的车辆，探测驾驶员注意力不集中的驾驶员监控摄像头是该系统的一部分。为了简化，假设以下描述为真：

——感知要素存在功能不足，若被触发条件 1 触发，该功能不足会导致危害行为，例如，执行错误的车辆运行轨迹；

——驾驶员监控摄像头存在功能不足，若被触发条件 2 触发，该功能不足会导致系统无法探测和减轻合理可预见的间接误用。

为使伤害发生，场景(3.26)需包含：

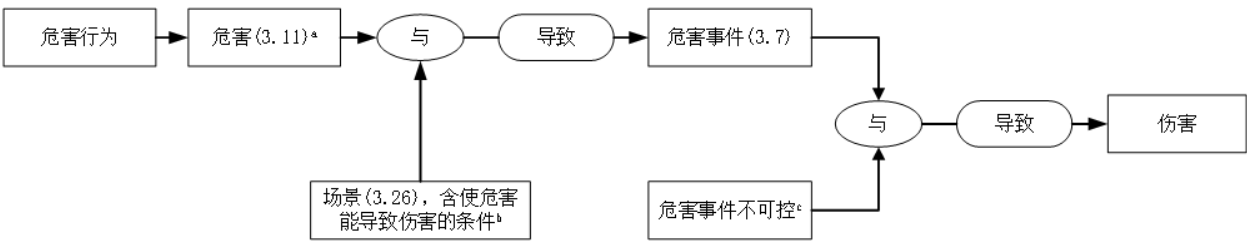
——驾驶员的间接误用：驾驶员注意力不集中且未及时探测到系统的危害行为，从而能够对其进行控制；

——触发条件 2，其导致系统无法及时探测和减轻出现的合理可预见的间接误用；

——触发条件 1，其导致系统危害行为。

注5：若整车层面的功能不足被触发条件触发，这将会导致危害行为，或导致无法防止、探测及减轻合理可预见的间接误用。见图4A。

注6：若要素层面的功能不足被触发条件触发，这将会导致输出不足。见图4B。输出不足本身或与其他要素的一个或多个输出不足相结合，将促成整车层面的危害行为或无法防止、探测及减轻合理可预见的间接误用。见图4B。



说明：

- a——危害是伤害的潜在来源，由整车层面的危害行为导致；
b——包含危害可能导致伤害的条件场景是伤害发生的助推因素，但不是伤害的来源；
c——无法充分控制危害事件是伤害发生的助推因素，但不是伤害的来源。

图2 危害和伤害发生的关联

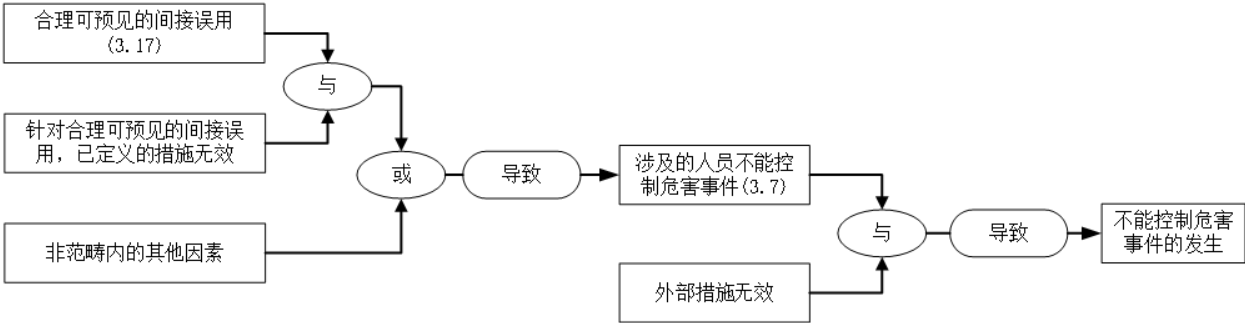
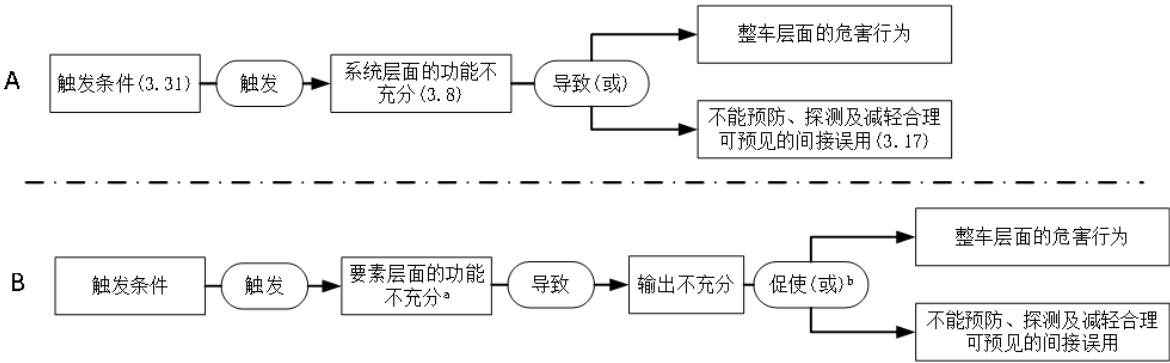


图3 危害事件未得到控制的原因



说明：

- a——取决于系统架构，要素层面的功能不足可被认为是单点功能不足(3.28)或多点功能不足(3.19)；
b——输出不足本身或与其他要素的一个或多个输出不足相结合，促成整车层面的危害行为或无法防止、探测及减轻合理可预见的间接误用。

图4 SOTIF 因果模型

3.9

功能修改 functional modification

功能规范的变更。

注：“功能修改”与GB/T 34590.1-XXXX中定义的术语“修改”不同。本文件中“功能修改”在GB/T 34590术语中称为“变更”。

3.10

后援用户 fallback ready user

有能力操作车辆，并能在一定时间内（对处于已定义的非驾驶期间的用户是合适的）进行干预以按要求执行动态驾驶任务后援（3.5）的用户。

3.11

危害 hazard

由整车层面危害行为导致的伤害的潜在来源。

3.12

规范定义不足 insufficiency of specification

可能不完整的规范定义，当被一个或多个触发条件（3.30）触发时，会导致危害行为或导致无法防止、探测及减轻合理可预见的间接误用（3.17）。

示例1：自适应巡航跟车距离的规范定义不完整，导致自车（3.6）未与前车保持安全距离。

示例2：由于规范定义的偏差，导致系统无法处理不常见的道路标识，即：不常见的道路标识不在规范定义的范围，因而系统不能恰当处理该标识。

注1：在系统生命周期的某个给定时间点上，规范定义不足可能是已知的，也可能是未知的。

注2：SOTIF（3.25）活动包括对规范定义不足的识别及对其影响的评估。在促成危害行为或无法防止、探测及减轻合理可预见的误用的能力尚未被建立时，可使用术语“潜在规范定义不足”。

注3：从规范定义、其他系统或要素的假设、或系统性分析（如第6章包含的分析或引出SOTIF设计及实现所要求的其他分析）中衍生出的要求，可包含在正式的数据库中，以助于确保验证。很多组织可能不认为这些要求是“规范定义”，但对于确保SOTIF是必要的。本文件中使用的术语“规范定义不足”包括此类衍生要求的不足。

3.13

预期行为 intended behavior

预期功能（3.14）的行为。

注1：预期行为是开发者考虑了能力局限的标称功能，这些能力局限来源于所用组件和技术的固有特性。

注2：开发者定义的预期行为，尽管不代表不合理风险（3.31），但可能与驾驶员对系统行为的期望不匹配。

3.14

预期功能 intended functionality

已定义的功能。

注：预期功能是在整车层面定义的。

3.15

驾驶自动化等级 levels of driving automation

基于驾驶自动化系统能够执行动态驾驶任务的程度，根据执行动态驾驶任务（3.4）中的角色分配以及有无设计运行范围（3.21）限制，将驾驶自动化进行的等级划分，分成0级至5级。

注：见表2（参考GB/T 40429-2021附录A，表A.1）。

表2 驾驶自动化等级与划分要素的关系

等级	名称	动态驾驶任务（3.4）		动态驾驶任务 后援（3.5）	设计运行范围 （3.21）
		持续的侧向和 纵向车辆运动 控制	目标和事件探 测与响应 （3.20）		
0	应急辅助	驾驶员	驾驶员及系统	驾驶员	有限制

1	部分驾驶辅助	驾驶员和系统	驾驶员及系统	驾驶员	有限制
2	组合驾驶辅助	系统	驾驶员及系统	驾驶员	有限制
3	有条件自动驾驶	系统	系统	动态驾驶任务 后援用户 (3.10)	有限制
4	高度自动驾驶	系统	系统	系统	有限制
5	完全自动驾驶	系统	系统	系统	无限制*
*排除商业和法规因素等限制					

3.16

最小风险状态 MRC minimal risk condition MRC

当给定的行程无法完成时，为降低风险(3.23)的车辆状态。

注1：这是动态驾驶任务后援(3.5)的一个预期输出。

注2：GB/T 34590中相似的功能安全术语是安全状态。

3.17

误用 misuse

以制造商或服务提供商不期望的方式使用。

注1：误用包括非故意的人员行为，但不包括故意改变系统或以造成伤害为目的而使用系统。

注2：误用可能源于对系统性能的过度信心。

注3：根据与危害行为的因果关系，有两种误用，直接误用和间接误用。

注4：直接误用可能是导致系统发生危害行为的原因，被认为是潜在的触发条件(3.30)。如果其形成了导致危害行为发生的能力，则将其作为触发条件。直接误用也可能是触发条件的一部分，即在直接误用后，需存在场景的其他特定条件，才能发生系统的危害行为。

示例1：直接误用：在城市环境中激活用于高速公路的功能，导致车辆未探测到停车标志并做出响应的场景(3.26)。

示例2：直接误用：在用户手册定义的 ODD(3.21) 范围以外，驾驶员激活了自动化系统。该直接误用与系统是否包含自行车(3.6)定位组件无关，此定位组件用于防止功能在定义的 ODD 范围外被激活。

注5：间接误用会导致对危害行为的可控性降低，或导致事故的严重度增加，或二者兼有。间接误用不被视为潜在的触发条件，因为其不能导致系统自身的危害行为。

示例3：间接误用：一个脱手 L2 级高速公路辅助系统存在已知的感知问题，要求驾驶员持续监控系统对 DDT(3.4) 的正确执行，并在必要时做出干预。间接误用是驾驶员睡着，而没有监控。该间接误用与驾驶员监控系统是否探测到此情况并做出纠正无关。

示例4：间接误用：当自行车处于自动驾驶状态并在运动过程中，乘员解开安全带。该间接误用潜在增加了事故的严重度，但不是触发条件。

注6：见图2-4。

3.18

误用场景 misuse scenario

发生误用(3.17)的场景(3.26)。

3.19

多点功能不足 multiple-point functional insufficiency

在一个或多个触发条件(3.30)触发下，且仅当与其他要素的功能不足结合出现时，才会导致危害行为或无法防止、探测及减轻合理预见的间接误用(3.17)的要素的功能不足(3.8)。

3.20

目标和事件探测与响应 OEDR object and event detection and response OEDR

动态驾驶任务(3.4)中的任务,其包含监控驾驶环境并对目标和事件(3.7)执行恰当的反应,以完成动态驾驶任务(3.4)和/或动态驾驶任务后援(3.5)。

3.21

设计运行范围 ODD operational design domain ODD

对于给定的驾驶自动化系统,在设计时确定的功能运行的特定条件。

注1:条件可以是空间的、时间的、法律的或环境的。

注2:驾驶自动化系统自身的条件,例如:车速、计算能力及感知能力,也属于ODD的范畴。

3.22

性能局限 performance insufficiency

技术能力局限,其在一个或多个触发条件(3.30)触发下,促成危害行为或无法防止、探测及减轻合理预见的间接误用(3.17)。

注1:在系统生命周期的某个给定时间点上,性能局限可以是已知的,也可以是未知的。

注2:性能局限考虑系统中的电气/电子要素,及与实现SOTIF(3.25)(见3.8注1)相关的其他技术要素。

注3:SOTIF(3.25)活动包括对性能局限进行识别并评估其影响。在促成危害行为或无法防止、探测及减轻合理可预见的误用的能力尚未建立时,可使用术语“潜在性能局限”。

示例:技术能力局限包括有限的计算性能、有限的传感器感知范围、有限的执行等。

3.23

风险 risk

伤害发生的概率及其严重度的组合。

3.24

反应 reaction

场景快照(3.27)中任何参与者对行为(3.2)的反馈。

3.25

预期功能安全 SOTIF safety of the intended functionality SOTIF

不存在因预期功能(3.14)或其实现的功能不足(3.8)引起的危害(3.11)而导致不合理的风险(3.31)。

注1:可导致危害的系统危害行为,是由场景(3.26)中的触发条件(3.30)引发的(见图2)。合理可预见的直接误用(3.17)被认为是潜在的触发条件。

注2:在识别危害事件(3.7)时,对“预期使用和合理可预见的间接误用(3.17)”与“因规范定义不足(3.12)或性能局限(3.22)导致的危害行为”进行结合考虑。

3.26

场景 scenario

按场景快照(3.27)的先后顺序,对几个场景快照的时间关系进行的描述,包括特定情况下受行为(3.2)和事件(3.7)影响的目标和取值。

注1:每个场景都开始于一个初始的场景快照。可定义行为、事件、目标及取值,以对场景中的时间关系进行特征化。与场景快照不同,场景持续了一定的时长。

注2:提及的“目标和取值”是预期功能的条件参数。目标可以是“保持在车道线以内”,取值可以是“行人安全的优先级高于避免财产损失”。

3.27

场景快照 scene

环境快照,包括风景、动态要素、所有参与者和观察者的自我表征以及这些实体之间的关系。

注1:场景快照可包含环境要素(状态、时间、天气、照明和其他周边条件)、道路设施或内部要素(道路或内部几

何构造、拓扑结构、质量、交通标识、道路边界等）及对象/参与者（如果适用，静态的、动态的、运动的、交互、动作）。

注2：一个包含所有实体（如风景、动态要素、参与者）的包罗万象的场景快照（即客观场景或基本事实）只能在模拟中建模。在现实世界中，场景快照是由传感器感知的。自车（3.6）或人类驾驶员感知到的场景快照是对真实情况的不完整、不准确、不确定和潜在错误的投影。

注3：场景快照也可包含自车及实施预期功能（3.14）的系统的情况，如：胎压、用户使用情况及系统部件存在的失效。

3.28

单点功能不足 single-point functional insufficiency

在一个或多个触发条件（3.30）触发下，直接导致危害行为或无法防止、探测及减轻合理可预见的误用（3.17）的要素的功能不足（3.8）。

3.29

态势感知 situational awareness

对态势的理解。

3.30

触发条件 triggering condition

场景（3.26）中的特定条件，这些条件引发了系统的后续反应，这些反应促成了危害行为或无法防止、探测及减轻合理可预见的间接误用。

注1：“触发”的概念包含有多个条件逐步发生而导致危害行为或无法预防、探测及减轻合理可预见的误用的可能性。

注2：场景（3.26）中的触发条件触发了功能不足（3.8），导致系统的后续反应。见图2-4。

示例：在高速公路上运行时，车辆自动紧急制动（AEB）系统误将道路标志识别为前车，导致了Xg的制动持续了Y秒。在此示例中，触发条件是在高速公路上运行时导致误识别路标的环境条件，而AEB具有相关的性能局限（3.22）（例如，感知精度低或算法分类错误）。

注3：SOTIF（3.25）活动包括对触发条件进行识别并评估系统的响应。在引发相应反应的能力尚未被建立时，可使用术语“潜在的触发条件”。

注4：合理可预见的直接误用可直接引发系统的危害行为，被视为潜在的触发条件。

注5：见图2-4。

3.31

不合理的风险 unreasonable risk

按照现行的安全观念，被判断为在某种环境下不可接受的风险（3.23）。

3.32

用例 use case

一组相关场景（3.26）的描述。

注1：用例可包括以下的系统相关信息：

- 一个或几个场景；
- 功能范围（如最大允许车速、最大允许减速度）；
- 预期行为；
- 系统边界；
- 关于环境和人员操作的假设。

注2：典型的用例描述不包括该用例的全部相关场景详单。相反，用这些场景的更概要的描述。

3.33

确认目标 validation target

用于论证满足接受准则（3.1）的值。

注1：确认目标的定义取决于目标市场和运行场景 (3. 26)。

注2：在SOTIF (3. 25) 范围内，确认是一种保证，基于检查和测试，确保接受准则（对于已识别的危害）会得到实现且具有充分的置信度水平。

示例：功能在 Y 小时的耐久运行中无危害行为，或在 X 次泊车中发生一次某种严重度的危害行为。

注3：为完全满足一个给定的接受准则，满足多个确认目标可能是必要的。

3. 34

整车层面安全策略 VLSS vehicle level safety strategy VLSS

针对预期功能 (3. 14) 的一组整车层面需求，用于支持设计、验证和确认活动以实现SOTIF (3. 25)。

注：可为每个SOTIF相关的系统定义一套整车层面安全策略。

3. 35

优先度子集 Priority Subset

按照一定的规则对要素的属性进行排序，所挑选出的部分要素的集合。

示例：根据场景要素中与预期功能不足的相关性（触发条件），及这些要素的发生概率，可以给出预期功能的一组场景优先度子集或用例优先度子集。

注1：由于场景和用例的完整描述难以实现，优先度子集有助于支持SOTIF分析、验证和确认、评估活动中，应对大量的场景和用例。

注2：对要素属性的排序规则可以是定性的，如：与预期功能是否相关，也可以是定量的，如：场景要素出现的概率大小。

注3：场景优先度子集是给定规则的场景概要描述，可支持对用例和场景的分析和建立，以及对场景覆盖率的论证。

4 预期功能安全活动概述和组织

4. 1 总则

第4章提供了：

- a) 预期功能安全原理的概述；
- b) 关于预期功能安全活动的工作流程和使用本文件的指南；
- c) 关于预期功能安全活动管理和支持过程管理的指南。

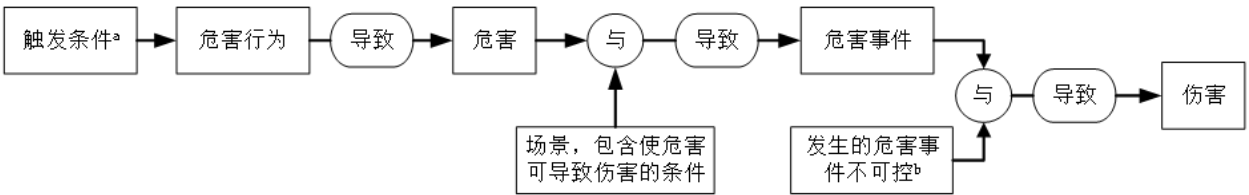
本文件中规定的活动适用于整车、系统和组件层面。

4. 2 预期功能安全的原理

4. 2. 1 SOTIF 相关的危害事件模型

本文件的主要目的是，对用于确保SOTIF相关危害事件达到足够低的风险水平的活动和理由进行描述。

功能和系统规范及设计包含了相关用例，而这些用例又由多个场景组成。这些场景可能包含导致伤害（简化的示意图见图5，详细的示意图见图2-4）的触发条件。为了避免伤害，适当的态势感知是必要的。



说明:

- a ——触发条件包括合理可预见的直接误用。
- b ——无法控制危害事件也可能是由于合理可预见的间接误用导致的，例如，驾驶员没有按照规定的方式监控系统。

图5 SOTIF 相关危害事件模型的示意图

示例1：当一个仅适用于高速公路的功能在城市环境中被激活时，该功能将难以识别和解释弱势道路使用者的运动。

示例2：对系统运行模式的错误理解，例如，系统未激活，但驾驶员误认为系统处于激活状态。在这种情况下，用于防止这种混淆的系统人机交互的潜在不足或（如果驾驶员行为可以被监控的情况下）系统缺乏恰当的反应，可能也被视为系统的危害行为。

注1：正确的态势感知依赖于：

——对相关环境条件具有足够全面、准确的感知，对场景快照的正确理解（例如，探测停车标志），以及关于每个道路参与者的状态的预测模型（例如，行驶方向，速度）。定位、自车运动、与其他车辆或环境的通信等信息可进一步用于态势感知。

——驾驶时，适当的行为或反应（例如，遵守与停车标志相关的规则）。

在车辆的整个运行生命周期中，可能会产生以下变化：

——环境（例如，新型交通标志、道路标记、车辆）；

——适当的响应（例如，新交通标志要求的新驾驶行为，驾驶场景的变化，驾驶法规的变化）。

注2：本文件第13章描述了对此类变化的监控。

注3：该问题可以通过源于驾驶策略的要求来解决。见附录D.1中的示例。

在定义设计运行范围（ODD）和系统开发过程中（例如，风险识别，定义适当的措施），考虑上述这些变化，以确保运行阶段的预期功能安全。

4.2.2 场景的四个区域

在本文件中，危害场景指的是导致危害行为的场景。作为相关用例的一部分，车辆运行场景可以分为四个区域（见图6和图7）。

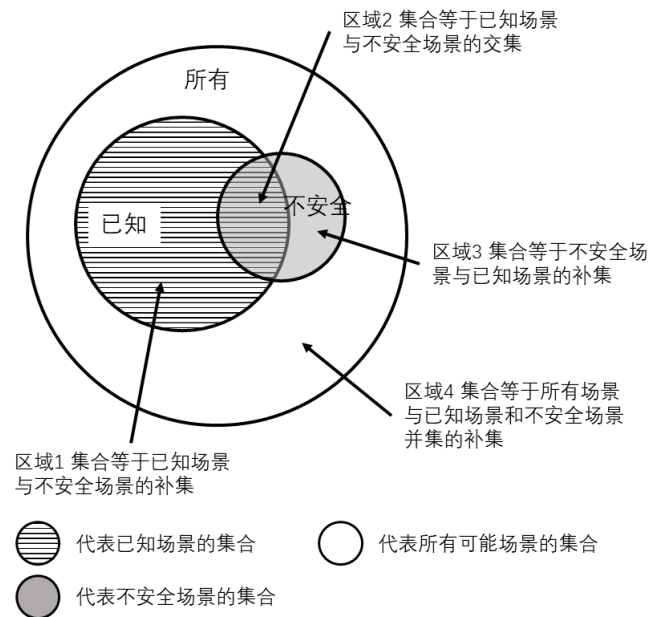


图6 场景的可视化分类

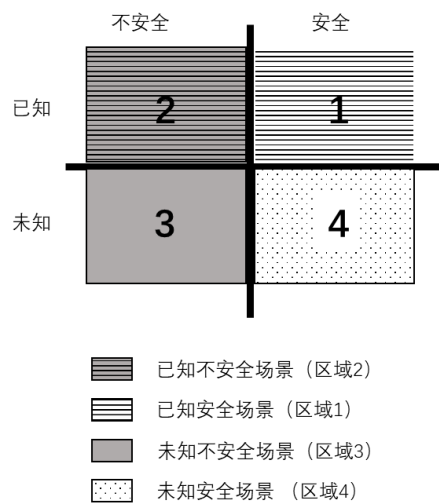


图7 场景可视化分类的替代方案

定义区域1、2、3和4，旨在构建和指导对本文件的理解：

- 已知安全场景（区域 1）；
- 已知不安全场景（区域 2）；
- 未知不安全场景（区域 3）；
- 未知安全场景（区域 4）。

示例：未知区域与以下场景相关：

- 定义了潜在的触发条件（例如，极端低温、不同驾驶场景的特殊组合），但系统的行为是未知的；
- 存在未知的触发条件（例如，概率极低的偶发事件）；或
- 场景的已知参数可组合成未知的潜在触发条件（例如，天气和交通条件的组合）。

注1：区域4中的场景是未知但安全的，不会导致伤害风险。一旦区域4中的场景被发现（即成为已知的），它就会被移至区域1。

该模型是一个抽象概念，其代表了SOTIF活动的目标，即：

- 基于对预期功能的分析，评估区域 2 的风险是否可接受；
- 通过功能修改（见第 8 章），将区域 2 中引起危害行为的已知不安全场景的概率，降低到可接受的标准；
- 通过充分的验证和确认策略（见第 9 和 11 章），将区域 3 中引起潜在危害行为的未知不安全场景的概率，降低到可接受的标准。

注2：各区域的大小代表场景的数量大小，而不是这些场景导致的风险大小。然而，这只是一种概念性描述方法，因为这些区域的大小并不是真正可以测量的。SOTIF的任务是为预期功能的风险足够低提供论据，其中，场景的数量是其中的一个方面，但不是唯一的方面。造成伤害的严重度和危害场景出现的可能性会影响预期功能的风险，但这些并未在区域中体现。

注3：即使在所应用的系统开发方法中，某些SOTIF相关活动未计划使用场景的方法，也不会改变SOTIF的目标，即，避免不合理风险。

一个给定的用例可以包含已知和未知的场景。探索每个用例的场景可以识别出以前未知的场景。

SOTIF活动的最终目标是评估区域2和区域3中存在的潜在危害行为，并提供论据以证明这些场景导致的残余风险足够低，即达到或低于接受准则。虽然明确评估了区域2中已知场景产生的风险，也需要通过基于统计数据的测试，以论证区域3中未知场景产生的风险足够低。

期望降低区域2和区域3的残余风险。通过不断增加区域1（见图8和9）中的场景集合，将提高实现SOTIF的信心。

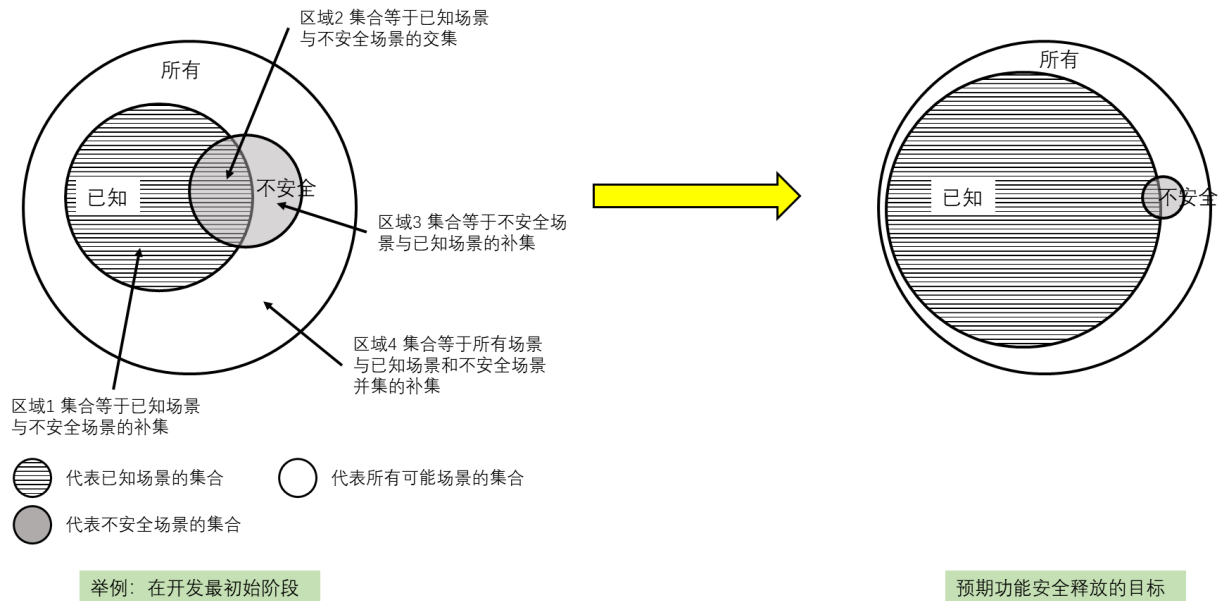


图8 预期功能安全活动带来的场景区域演变

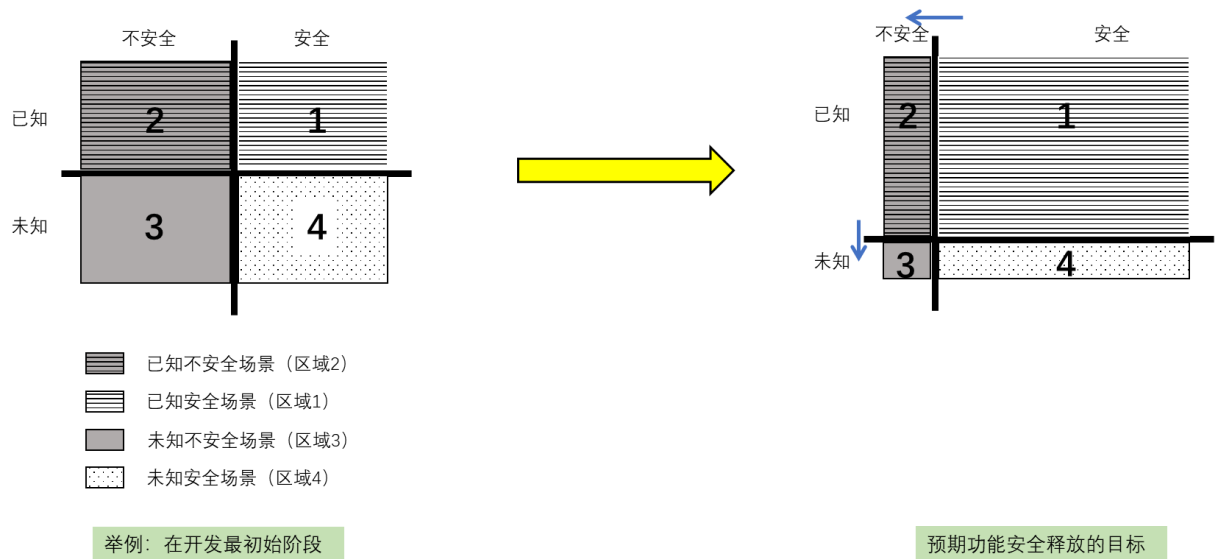


图9 预期功能安全活动带来的场景区域之间的演变替代表现形式

4.2.3 “感知-规划-执行”模型

本文件中，导致危害行为的可能原因，与系统能力密切相关，这些能力包括创建足够准确的环境模型，基于环境模型做出正确决策并推导出正确的控制动作，以及执行该控制动作。

关键系统要素及其交互由“感知-规划-执行”模型表示（见图10）。“感知”要素执行感知部分（包括定位），即依据从车辆外部和内部环境以及车辆和系统状态所感知并接收到的信息，创建环境模型。“规划”要素将其目标和策略应用于由“感知”要素提供的环境模型，以得出控制行为。最后，“执行”要素执行控制行为。

注：决策算法被包含在“感知-规划-执行”模型的所有要素中（例如，分类、传感器数据、融合、态势分析、行为

决策)。

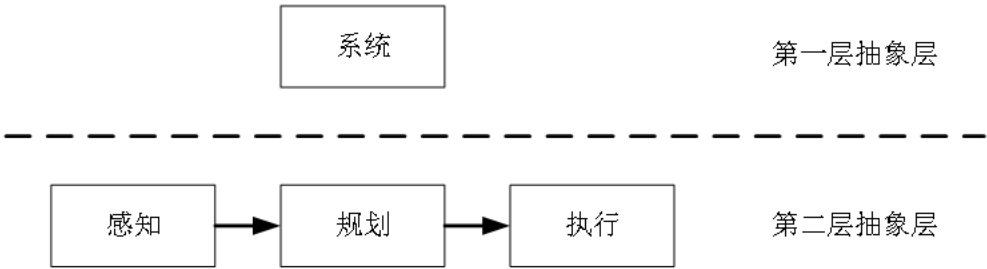


图10 “感知-规划-执行”模型

基于“感知-规划-执行”模型，对有能力的整体系统架构的选择，是实现高效SOTIF过程的一个重要考虑因素，在高效的SOTIF过程中，与整体能力相关的活动可在早期阶段启动，并贯穿整个功能开发生命周期。选择一个有能力的系统架构对于确保SOTIF至关重要。因此，可在系统开发的早期阶段启动系统架构的定义工作。此外，在系统整个生命周期中，定期评审系统架构，并在必要时进行更新。

4.3 本文件的使用

4.3.1 本文件的流程图和结构

SOTIF流程(见图11)开始于规范定义和设计(见第5章)。规范定义和设计中包含了那些在后续SOTIF活动和周期开始前就已知的性能局限和功能不足。SOTIF活动的迭代可能会导致规范定义和设计的更新，以及新的先前未发现的性能局限和功能不足。开始于规范定义和设计的每次迭代，也致力于将规范定义和设计更新到最新的状态。

对预期功能的潜在危害行为进行危害识别和风险评估(见第6章)。对已识别出的危害事件进行风险评估，并定义相应的风险接受准则。如果证明危害事件不会导致不合理的风险，则不需应用额外的设计措施。第6章不考虑预期功能的危害行为的原因，而仅考虑其对安全的影响。因此，重点是评估可能由危害行为引起的危害事件，并定义所需满足的接受准则。

第7章旨在识别可能导致预期功能危害行为的根本原因(见图2)，并评估由所识别出的潜在功能不足和触发条件引起的风险是否合理。

根据第6、7、9、10、11、12、13章的活动，如果认为有必要，则对功能进行修改(例如，改进传感器的能力，对ODD进一步的限制)，以改进SOTIF，见第8章。

制定验证和确认策略，以提供证据来证明：与SOTIF相关的整车层面残余风险符合可接受的水平，要素满足其功能要求(见第9章)，对设计运行范围(ODD)的覆盖是充分的。为了能够收集所需的证据，可以从该策略中导出相应的验证和确认测试用例，且ODD上的测试用例具有足够高的覆盖率(见第10和11章)。

评估SOTIF活动的结果是否充分，足以论证实现了SOTIF，见第12章。

在运行阶段定义了识别和解决可能出现的SOTIF现场运行问题的流程(见第13章)。

图11描述了本文件中为确保预期功能安全所需的活动流程。带圆圈的数字表示本文件中的相应章节。

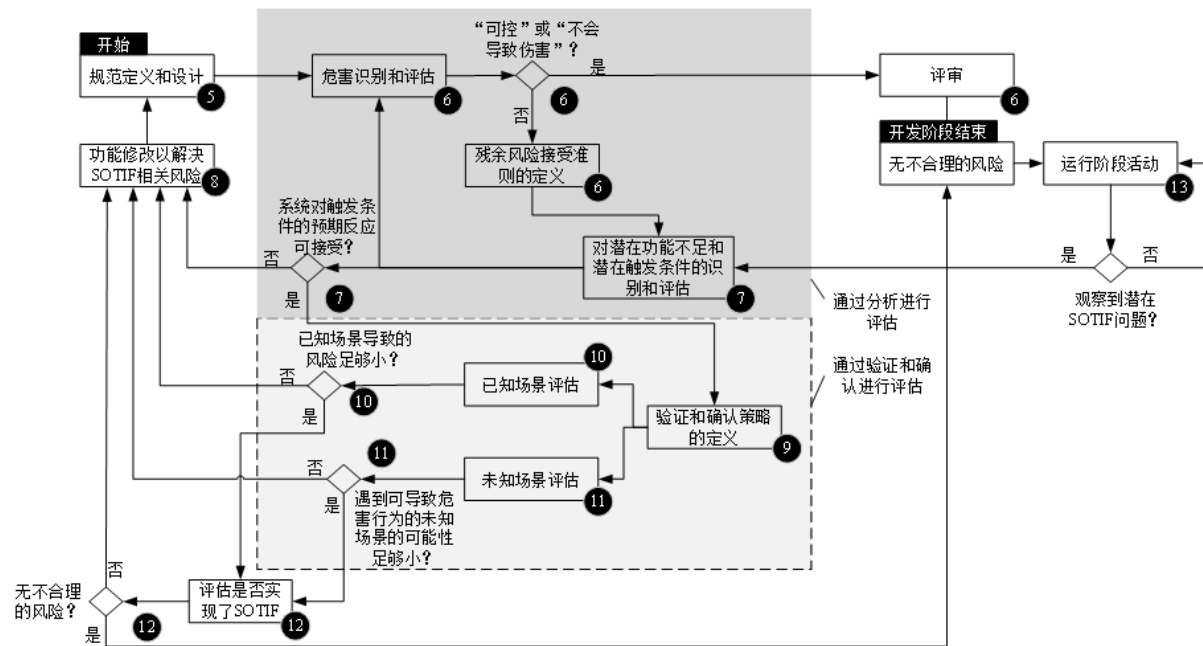


图11 SOTIF 活动的相关性

附录A 提供了关于SOTIF的一般指南。

附录B 提供了关于场景和系统分析的指南。

附录C 提供了关于SOTIF验证和确认的指南。

附录D 提供了关于SOTIF特定方面的指南，例如，驾驶策略的定义，对机器学习的影响以及对地图和V2X的考虑。

附录E 提供了风险接受准则的示例。

4.3.2 规范的条款

通过在各章节的开始列出目标，提供实现目标的论据，并记录相应的工作成果，可以声明符合本文件。目标的规范性特征是通过使用关键词“应”来表达的，表示一种要求。

注：附录A.1给出了基于目标结构符号（GSN）的论证示例。

4.3.3 表的解释

本文件中的一些表列出了为实现某一开发目标的一系列方法和措施。这些条目旨在说明可能的方法和措施，而表的条目可能并不详尽。可以采用其他等效的方法和措施。这些表的目的是支持开发团队选择一个或多个适当的措施和方法。

注：选择一组适当的方法可能取决于各种因素，例如，危害事件的复杂性或暴露概率。

4.4 SOTIF 活动管理和支持过程

4.4.1 质量管理、系统工程和功能安全

为了开发安全的产品，严格的工程和质量管理体系至关重要。这些已经在其他标准中予以说明，如IATF 16949、GB/T 34590和ISO 15288。本文件仅关注这些过程中的SOTIF特定方面。

注1：在产品开发过程中，本文件和GB/T 34590中规定的活动是并行开展的。一般而言，实施的措施可能对SOTIF和功能安全产生影响，并按照这两种方法论进行评估。附录A.2为并行应用GB/T 34590和SOTIF提供了实践指南。

对于管理活动和支持过程，可以将GB/T 34590第2、7和8部分扩展到SOTIF活动中。5.3章条和10.2章条进一步描述了需求传递和追溯性要求。

对于SOTIF相关的活动，按下文选择一组方法和措施：

——SOTIF 过程（见图 11）始于规范定义和系统及其架构的设计（见第 5 章）；

——对预期功能的潜在危害行为进行危害识别和风险评估（见第 6 章），以识别危害及其相应的危害事件。如果证明危害事件不会导致不合理的风险，则不需应用额外的设计措施。

注2：第6章不考虑预期功能的危害行为的原因，而仅考虑其对安全的影响。因此，重点是评估可能由危害行为引起的危害事件，并定义所需满足的接受准则。

——第 7 章旨在识别可能导致预期功能危害行为的根本原因（见图 2），并评估由所识别出的潜在功能不足和触发条件引起的风险是否合理。

——根据第 6、7、10、11、12、13 章的活动，如果认为有必要，则对功能进行修改（例如，改进传感器的能力，对 ODD 进一步的限制），以改进 SOTIF。见第 8 章。

——制定验证和确认策略，以提供证据来证明：与 SOTIF 相关的整车层面残余风险达到可接受的水平，要素满足其功能要求（见第 9 章），对设计运行范围(ODD)的覆盖是充分的。为了能够收集所需的证据，可以从该策略中导出相应的验证和确认测试用例，且 ODD 上的测试用例具有足够高的覆盖率（见第 10 和 11 章）。

——根据先前活动的结果，评估残余风险（见第 12 章）。

——在运行阶段定义了识别和解决可能出现的 SOTIF 现场运行问题的流程（见第 13 章）。

注3：关于GB/T 34590功能安全和本文件之间的交互，附录A.2给出了进一步解释。

4.4.2 分布式 SOTIF 开发活动

在分布式产品开发的情况下，所有相关方之间定义了开发接口协议（DIA），DIA的目的是在项目的早期阶段确认SOTIF活动的所有职责，并在开发方之间交换充分的技术信息。

IATF 16949提供了一个基本流程框架，也可在SOTIF过程中予以考虑。本章条侧重于如何将DIA扩展到分布式SOTIF开发和实施中。GB/T 34590提供了关于功能安全方面的DIA和供应协议框架。为了将此框架应用于SOTIF，可以进行剪裁，增加与SOTIF开发和实施相关的各方职责。考虑并商定各方的责任，以计划和执行第5至第13章的整个SOTIF活动。对将要共享的信息和工作成果进行定义。可根据GB/T 34590.8中5.4.1、5.4.2、5.4.3、5.4.4和5.4.6所描述的过程，并为SOTIF进行剪裁，来完成这些活动。在项目开发初期就文档格式达成一致。

4.4.3 独立于环境的 SOTIF 相关要素

为了实现SOTIF，对不同系统（硬件和软件）之间的接口的描述是必要的。为了确保集成的系统在所定义的ODD内是安全的，每个子系统（例如，独立的感知系统）的边界都要经过仔细的评估。因为环境因素（例如，ODD、场景）是SOTIF开发的基本问题，所以系统及其要素依层级的不同具有不同的关注点。就这些系统和要素的开发而言，它们可以分为以下三种类型之一：

- a) 在环境中的开发：整个系统的所有 SOTIF 活动遵循 V 模型进行开发。对于从事系统及其要素的分布式开发的各方，根据任务角色，来确定要求，包括规范定义和设计（见第 5 章）以及其他活动（见第 6、7、8、9、10、11，12 和 13 章）。在 GB/T 34590 中，这种开发被视为“在环境中”的开发。
- b) 独立于环境的 SOTIF 相关要素：对于这些要素，可以对它们在整个系统中的使用及其对预期功能的贡献做出假设。因此，可以对与 SOTIF 相关的输出不足及其容许的发生概率目标做出假设。记录这些假设并将其作为这些要素后续开发的输入。SOTIF 活动提供了实现相应概率目标的证据。对于独立于环境的 SOTIF 相关要素，记录已识别出的触发条件及其引起的输出不足。

通过整车层面功能环境下的 SOTIF 活动，确定假设的有效性（见 GB/T 34590.10 第 9 章的 SEooC）。

- c) 非特定的 SOTIF 相关开发：这些要素的功能可通过多种方式帮助实现预期功能，因此，如果使用这些要素的环境未知，提前预估 SOTIF 相关要求在实践上是不可行的。

5 规范定义和设计

5.1 目的

本章的目的是实现以下目标：

- a) 规范定义和设计应包含充分的信息以开展 SOTIF 相关活动；
- b) 在 SOTIF 相关活动的每次迭代后，应根据要求对规范定义和设计进行更新（见图 11）。

5.2 功能规范的定义和对设计的考虑

规范定义和设计可包括本条中所列的各个方面。某些方面仅与特定的自动化等级或特定的应用相关。此外，某些方面与整车层面或要素层面的功能规范定义相关。

考虑的方面（如适用）包括但不限于以下方面：

- 对预期的功能、支持子系统和组件的功能的描述，包括：
- 运行设计范围 ODD；
- 自动驾驶功能控制车辆动态任务的权限级别和细节；
- 车辆层面的 SOTIF 策略；
- 功能可被激活或关闭的用例，以及激活和关闭用例间的转换；
- 决策逻辑的描述（例如，路径规划、驾驶策略，见 D.1）；
- 实现预期功能的相关系统及其要素的设计；
- 为实现预期功能而安装的传感器、控制器、执行器或其他输入及组件（例如，地图，见 D.3）的性能目标；

注1：自动驾驶系统的性能目标（见D.5），例如，包括在ODD范围内对关键目标和事件（例如，行人、车辆、自行车、摩托车和交通标志）的探测与响应。

- 预期功能与如下条目间的依赖关系、交互或接口：
 - 驾驶员；
 - 驾驶员交互（例如，人机交互HMI），及如何使用交互以减轻已知的合理可预见的误用；
 - 乘客、行人、骑自行车的人和其他道路使用者；
 - 相关环境条件；
 - 道路基础设施和道路设备；
 - 云端、车辆间或其他通信基础设施（例如，V2X/X2V，见D.4）以及涉及诊断和参数更新的在用远程信息处理之间的数据交换。
 - 软件更新的远程刷写；
 - 车辆上可能干扰预期功能的其他功能，包括信息交换和相应的使用假设；
- 合理可预见的误用（直接和间接）；
- 系统及其要素的潜在的性能局限、已识别出的触发条件、应对措施；

注2：在SOTIF活动中，识别出的一些潜在的性能局限和风险可能是可以接受的，无应对措施。在此情况下，可在规范定义和设计中予以记录。

- 实现预期功能的系统和整车架构；

——报警和降级概念：

- 报警策略；
- DDT后援：接管/后援的条件和方案，用于在各自的用例中，将控制权从自动驾驶系统转移到驾驶员或另一个系统；
- 最小风险状态方案（例如，自动离开车道并停车、在路径中停车、后援用户）；
- 驾驶员监控系统及其对后援策略的影响。

——在预期功能的开发过程中和开发后，支持数据收集和监控的程序：

- 数据收集的目的和要求；
- 在SOTIF发布前，支持开展所需数据收集的架构、实现和机制；
- 运行阶段，支持数据收集的要求、设计和机制，以用于SOTIF分析（见13.5），包括：可能基于云、OTA或射频通信的技术。

——运行阶段，形成风险缓解能力的机制、设计和要求（见D.6）。

注3：附录A.4给出了一个简化后的规范定义和设计框架性示例。

5.3 系统设计和架构的考虑

规范定义和设计提供了对系统及其要素、功能和性能目标的充分理解，以便执行后续阶段的活动。这包括一份已知的功能不足、相关的触发条件及其应对措施（如果适用）的详尽列表。在开展SOTIF相关活动前，一些潜在的功能不足、触发条件、应对策略是已知的并记录下来，而其他的则在开展SOTIF活动过程中被揭示出来。通过实施应对措施对系统进行设计，以缓解已知功能不足对整个系统的影响。

SOTIF相关活动（见图11）的每次迭代，可引发在任何相关层面上更新规范定义和设计的工程活动。每次迭代依赖于在任意相关层面上对规范定义和设计的更新，这也体现出此前迭代中发现的所有信息。

各开发方【主机厂（OEM）、1级供应商（Tier1）、N级供应商（TierN）】之间的合作，对于发现集成系统、组件或要素的潜在功能不足，并在开发阶段（见4.4）针对这些不足制定应对措施是必要的。将设计和规范的相关部分传达给较低层面的系统及组件开发人员。在每个开发周期/迭代之后，所使用的假设、可预见的误用和潜在的性能局限从一个层面传递到相邻的层面，包括OEM。

随着SOTIF活动识别出新的功能不足和触发条件（见第7章），并定义改进SOTIF的措施（见第8章），规范定义和设计将作为每个开发周期的一部分进行更新，见图11。

SOTIF工作成果如果影响规范定义和设计，包括预先存在的相关内容，则其与规范定义和设计是相关联的（如5.2中的定义）。这确保了从之前的迭代中获取到所有信息，同时确保规范定义为下一个迭代周期做好了准备。

注：规范定义和设计（工作成果5.5）的可追溯性和完整性可通过与SOTIF措施（工作成果8.5）的关联性予以证明，这些措施可以进一步关联到：

——相关设计文档；

——工作成果，来自于：

- 第6章：对已识别出的危害事件的评估（例如，需要达到S=0、C=0或需要满足更宽泛的接受准则）；
- 第7章：系统对潜在触发条件响应的评估（例如，与分析关联，该分析表面触发条件的风险不可接受）；
- 第9章和第10章：已知不安全场景的验证报告（例如，与验证测试报告关联，该报告表明与要求相关的性能不可接受）；
- 第9章和第11章：未知不安全场景的确认报告（例如，与确认测试报告关联，该报告表明与不安全场景或确认目标相关的性能不可接受）；
- 第12章：SOTIF发布的论证（例如，与记录拒绝发布请求理由的报告关联）；
- 第13章：现场监控（例如，与记录现场监控期间发现的新的不安全场景的报告关联）。

第6章（6.4）和第7章（7.4）中与风险评估相关的SOTIF技术假设不一定与第8章（8.3）中的SOTIF措施相关，但仍可将这些技术假设追溯至规范定义和设计。那些可提供基于模型的设计和不同模型工件（要求、组件、接口、分析、测试用例和结果）之间的自动追溯的设计工具，可以支持这一过程。

5.4 性能局限和应对措施的考虑

设计包括对可能因要素输出值引起潜在性能局限的考虑，这些性能局限可能潜在导致整车层面的危害行为。潜在性能局限的非穷尽示例包括：

- 分类不充分；
- 测量不足；
- 跟踪不足；
- 目标选择不足；
- 运动学估算不足；
- 假阳性探测（误报，例如，鬼影、幻影物体）；
- 假阴性探测（漏报）；
- 驾驶策略层面的局限，例如，对堵塞区域的考虑。

B.3提供了关于识别功能不足和相应整车层面危害行为的可能方法的指南。功能不足与系统在已定义的ODD范围内运行时最为相关。但系统对离开ODD的探测方式，以及在离开ODD时控制权限转换期间系统如何运行，与支持完整的分析也相关。

系统开发基于对设计中的性能局限的假设。实施措施以应对这些性能局限，从而确保SOTIF。集成到规范定义和设计中的设计和措施，降低了残余风险，提高了整体鲁棒性（见图6和7）。

注1：第7章详述了识别潜在功能不足及其触发条件的方法和措施。

注2：第8章描述了解决功能不足的方法和措施，例如，冗余、多样性和互补要素。

注3：规范定义和设计中的SOTIF内容，按照第10章的详细说明进行验证。

以下是性能局限和可能的应对措施的示例，相关内容包含在规范定义和设计文档中：

示例1：用于车道保持等功能的高速公路车道边界探测算法，可能由于道路上的废弃物而错误地确定车道。然而，导致碰撞的车道偏离可通过其他自动驾驶功能来缓解，例如：使用高清地图和定位来确认车道，基于前面车辆的轨迹将自车轨迹合理化，保持与其他车辆距离的防碰撞算法，即使这意味着离开已识别出的车道等。

示例2：目标探测算法将滑板上的人视为行人，但由于其速度不可信而拒绝该目标。在这种情况下，通过对目标探测算法和感知及处理算法之间进行抽象，并使用其他不同的合理性检查的系统，可以缓解与滑板者的碰撞。

示例3：在某些区域，用具有三维视错觉的人行横道（见图 12）来提醒驾驶员。在道路上绘制图像的目的是为了欺骗人类的感知，但也可能欺骗视觉系统，使其探测到不存在的物体。在这种情况下，基于光流的分析机制可防止错误制动。光流分析和基于雷达的环境识别作为相互替代的应对措施，以应对由视觉分类局限而导致的此类情况。



图12 可能欺骗视觉系统的视觉幻象图的示例

示例4：使用自动泊车系统时，有物体从打开的后备箱中伸出，可能会导致危害事件。系统设计中的应对措施是只允许在后备箱关闭时进行自动泊车。

5.5 工作成果

工作成果是满足5.1 a)和5.1 b)的规范定义和设计。

注1：规范定义和设计可以拆分或链接到多个文档，例如，SOTIF相关系统的需求规范、功能规范定义和设计规范。

注2：缓解措施的SOTIF规范可集成到现有功能安全设计文档中，例如，功能安全概念和/或技术安全概念。

6 危害的识别和评估

6.1 目的

本章的目的是实现以下目标：

- a) 应系统地识别整车层面定义的预期功能所导致的危害。
- b) 应系统地识别和评估由预期功能的危害行为导致的风险，及危害行为可能导致危害的相应场景。应定义预期功能的行为被视为不安全的情况下的参数。
- c) 应定义残余风险的接受准则。

6.2 总则

为实现本章的目标，可以考虑以下信息：

- 规范定义和设计，按照 5.5；
- 用于推导接受准则的可用数据。

6.3 危害识别

系统地确定整车层面由功能不足引起的危害。这种系统地识别主要是基于对功能的认识及可能因功能不足而引起偏差的认识。这可以通过应用GB/T 34590.3中定义的方法来实现。GB/T 34590和本章要

求的危害分析的通用要素的说明，如图13。图14以AEB系统为例，展示了如何使用图13中的术语。该示例显示了由相同的危害行为导致的两种危害。在附录A. 2. 5中以AEB为例进一步阐述了危害分析的应用。

示例1：AEB 系统可能因预期功能导致危害行为和功能异常导致危害行为而引发危害。对于功能约束范围内和范围外的非预期制动导致的危害，在危害分析和风险评估中，可从功能安全的角度进行分析。对于功能约束范围内非预期制动相关的危害，也需要进行 SOTIF 分析。

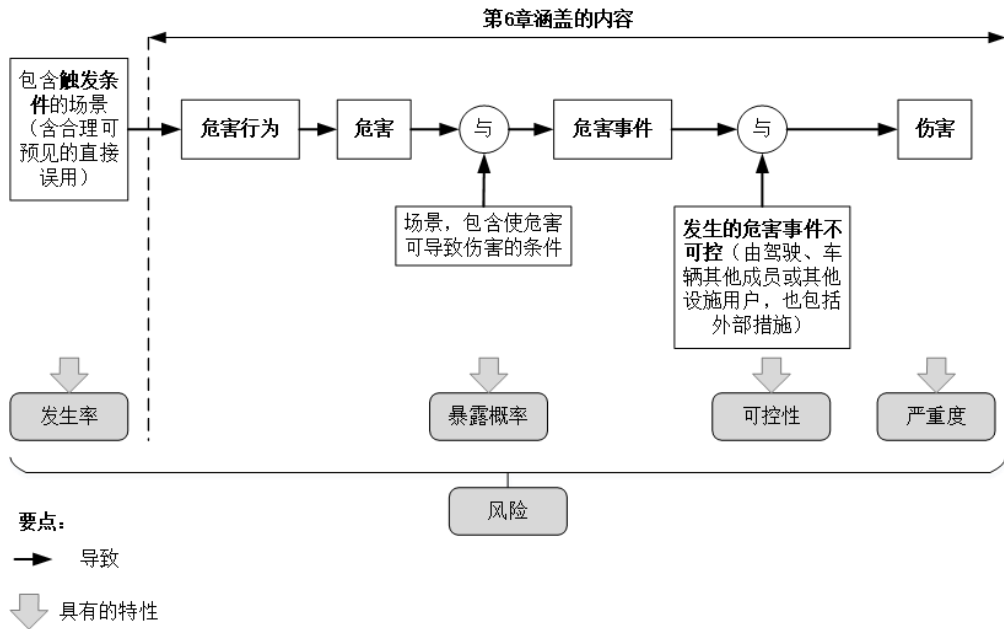


图13 GB/T 34590 和本文件中危害分析的通用要素示意图

注1：与GB/T 34590. 3不同，在分析SOTIF相关的危害时，不会为危害事件确定汽车安全完整性等级（ASIL）。然而，严重度(S)、暴露概率(E)和可控性（C）参数可用于调整确认工作所需的工作量。

注2：发生率反映了功能在运行阶段遇到触发条件的概率。

触发条件的发生率和危害导致伤害所处的场景的暴露概率，两者有重要的区别。一般来说，触发条件并不独立于场景。因此，为支持在SOTIF风险降低的论证中使用场景暴露概率，评估时需要考虑处于场景的暴露概率和遇到触发条件的概率之间的统计相关性。

示例2：对于行驶在高速公路上的场景和针对高速领航驾驶功能的触发条件之间，不能假设存在统计学上的独立性。在某些特定情况下，可对统计学上的独立性进行假设，如图14所示。

可控性参数C可用于评估 SOTIF 相关危害是否可控（见10.6和表10）。关于道路参与者反应的研究或假设可用于支持可控性评级。

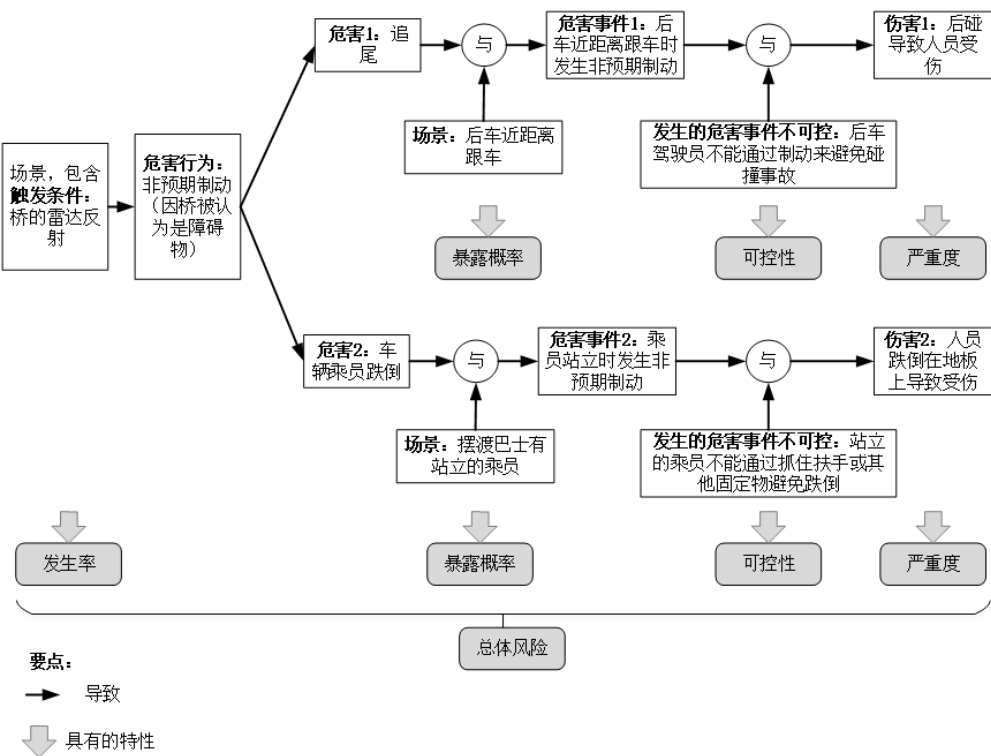


图14 使用图 13 中术语的 AEB 示例

注3：图14中的示例说明可从两个层面评估所导致的风险：首先，是给定场景中特定危害相关的风险；其次，是与全部危害行为相关的总体风险，这包括对多个危害和相应场景的评估。

除针对可能的功能偏差导致的危害进行系统性识别外，还可以考虑驾驶员或用户与系统的交互（包括合理可预见的误用），以识别出更多的危害。根据合理可预见的误用与导致危害的关系，对其进行区分。对预期功能的直接误用会形成触发条件，而对预期功能的间接误用会降低可控性，或增加因危害行为造成的危害事件的严重度（例如，驾驶员注意力不集中或驾驶员对功能局限的误解）。

第6章涵盖了对合理可预见的间接误用的识别及其影响的分析。

注4：7. 3. 4和附录B. 1提供了分析合理可预见误用的通用指南。

6. 4 风险评估

风险评估旨在评估给定场景中危害行为的危险；这有助于定义 SOTIF相关风险的接受准则。

注1：由整车层面功能不足导致的危害行为（如果有）是本评估的一部分。

可以使用GB/T 34590. 3-XXXX第6章中描述的方法来评估伤害的严重度和危害事件的可控度。尽管共享分析方法，但SOTIF分析中观察到的结果和特定危害的评估参数可以是不同的。

注2：GB/T 34590. 3-XXXX中给出了可控性、严重度和暴露概率等级。在本章中，仅考虑危害事件在总体上是否可控，或者是否会导致伤害。暴露概率不是本章风险评估的决定参数。这是因为，对场景中需要评估的风险的选择已经意味着它们的暴露与SOTIF 相关，否则分析中将不作考虑。

注3：为定义确认目标，可考虑暴露于特定场景的概率（见第9章）。

示例1：对于因自动紧急制动引发的与主车辆的追尾，可通过限制制动干预的幅度来降低严重度。幅度限制可以被视为增加可控性的安全措施，或视为对预期行为的功能修改。在分析危害时，该限制被认为是预期行为的一部分；与此不同的是，对于该限制在实施过程中的失效，则属于其他安全标准的范畴，例如 GB/T 34590。

考虑危害事件的严重度和可控性，以确定危害行为在给定场景下产生的风险是否合理。严重度和可控性的评估考虑功能规范（根据第 5 章规范定义和设计）。可控性评估包括相关人员为控制危害而做出的“无反应”或“延迟反应”，例如这些反应可能是由合理可预见的间接误用导致的。评估也可考虑外部措施。如果可控性被评为“可控”（即 $C=0$ ）或严重度被评为“无伤害”（即 $S=0$ ），则认为不存在不合理的风险。其他情况的危害事件被认为与 SOTIF 相关。对这些危害行为，使用可测量的参数（如速度的偏差和与其他物体的最小距离）来进行描述，这些参数是危害行为接受准则（第一层接受准则）的来源，对于不符合这些接受准则的危害行为，需要进行合理的应对，并在接下来的SOTIF残余风险评估中予以考虑。

注4：对于危害行为的风险评估和接受准则的定义，应考虑目标市场人员和场景的合理水平，包括可能的驾驶习惯、文化、人员能力、交通情况及其随时间的演变等。

注5：支持危害行为接受准则定义的可测量参数，也可基于所考虑场景下的交通数据统计结果，例如：基于交通数据统计分析，给出的高速上有风险的跟车距离的定义。

注6：危害行为接受准则可用于支持整车安全策略及安全要求的制定，附录E提供了一个基于交通数据分析来定义危害行为接受准则及整车安全策略的示例。

示例2：在危害事件归类中，可考虑高级驾驶员辅助系统（ADAS）无法以安全方式处理某一环境条件，并因此需要驾驶员恢复控制。

驾驶员延迟的反应或不恰当的反应，包括驾驶员为获得足够的态势感知和恢复控制所需的时间，可能会影响可控性的评估，是SOTIF相关分析时需要考虑的。

注7：SOTIF分析中考虑人员可能由于未及时、正确的反应，而比功能安全分析时考虑人员及时反应而定义的可控性更低，或者因为人员对车辆预期行为存在误解而导致在特定场景中的误操作，使严重度升高，例如：驾驶员由于紧张或缺乏信心，可能对自动驾驶车辆的一些行为偏差矫枉过正。

如果在功能修改后（见图11），危害事件被判断为 $S=0$ 或 $C=0$ ，则该危害已得到充分解决。

示例3：表 3 给出了针对 AEB 系统 SOTIF 相关危害事件的潜在后果进行评估的示例。

表3 危害事件的示例

危害行为	潜在后果	严重度		可控度		不合理的风险？
		评级	注	评级	注	
高速公路行驶时，非预期的AEB激活，达到 $x \text{ m/s}^2$ 的减速度并持续 y 秒	与后车发生追尾	$S>0$	有效碰撞速度： $v \geq x \text{ km/h}$	$C>0$	后方车辆可能无法通过制动来避免碰撞。	是

6.5 残余风险接受准则的定义

针对不满足危害行为接受准则的那些行为，可导致 $S>0$ 且 $C>0$ 的危害事件，需要为这些危害行为定义残余风险接受准则，即第二层接受准则，并从第7章继续开展相关活动，最终实现对残余风险接受准则是否得到满足的评估。

作为SOTIF流程的一部分，对 $S=0$ 或 $C=0$ 分级的论证进行评审，包括对分级证据（例如，测试或分析结果）的评审。

接受准则考虑：

- 适用的政府和行业法规；
- 一项功能是新功能还是市场上已存在的功能；

- 对于可能暴露于风险的人（例如，自动公共交通系统中的车主、操作员、行人或乘客），风险是否合理；
- 已存在的功能的接受准则；
- 以模范方式行事的驾驶员的表现。

示例1：接受准则可以是每小时事故的最大数量，在第9章中的验证和确认策略是基于特定接受准则定义的。

在制定接受准则时，可以考虑的方法包括：

- 目标市场的可用交通数据（例如，事故统计、交通分析）（见C.2.2.4）；
- 现场运行中，类似功能的现有准则。

示例2：已量产的类似碰撞报警系统，每x公里产生的误报事件数量（采用类似的测试分布）。

如果给出有效的理由，就可以选择适当的定量接受准则。总体理由可以基于以下单个理由中的一个或多个的组合：

- 风险容忍原则，例如GAMAB或GAME，这两个法语术语的意思是“至少在全球范围内一样好”。遵循这一原则，任何新系统的残余风险（关于安全方面）不应高于那些具有类似功能或危害的现有系统。
- 正向风险平衡原则。这种风险容忍原则应用于考虑了新系统所有危害的整体残余风险，可以进行相关风险权衡。对于某个危害，即使残余风险增加，系统也可发布，前提是通过降低一个或多个其他残余风险来达到平衡补偿。
- ALARP原则。ALARP风险管理框架可以提供一个有用的风险降低原则，尤其是在目前不存在“良好实践”的情况下，开发和引入新技术。通过承认零风险状态是不可能的，ALARP原则旨在通过权衡风险与进一步降低风险所需的努力，将风险降低到“合理可行”的水平。
- MEM（最小内源性死亡率）原则。MEM原则基于的理念是技术系统的引入不应显著增加社会死亡率。由技术系统导致的死亡概率的定量接受准则源自自然原因导致的最小死亡概率。

注1：本文件中的理由仅包括与SOTIF相关的风险，不包括来自其他安全领域（例如，电气安全）的风险。

注2：附录C.2和C.6给出了定义和评估接受准则和确认目标的示例。

注3：关于GAMAB、ALARP和MEM的描述，见EN 50126-2:2017, A.1 (RAMS) [4]。

注4：有效的理由可基于整个车队的风险或与单个车辆相关的风险。即使车队进入包含触发条件的某一场景的概率非常低，但如果给定的单个车辆面临这种场景的概率很高，系统的响应也可能是不可接受的。

注5：附录E提供了一个基于交通数据分析定义风险接受准则的示例。

6.6 工作成果

- 6.6.1 整车层面的危害，实现目标6.1 a)。
- 6.6.2 危害行为的风险评估，实现目标6.1 b)。
- 6.6.3 接受准则，实现目标6.1 c)。

7 潜在功能不足和潜在触发条件的识别与评估

7.1 目的

本章的目的是为了实现以下目标：

- a) 应识别出潜在规范定义不足、潜在性能局限和包括合理可预见的直接误用在内的潜在触发条件，并确定导致危害行为的来源。
- b) 应对系统的响应进行SOTIF可接受性评估。

注1：包括在合理可预见的直接和间接误用情况下，识别功能不足和相关触发条件。

注2：本活动考虑整车层面预期功能的潜在规范不足以及系统中E/E要素的潜在规范不足或潜在性能局限。

注3：根据危害行为接受准则，判断规范不足、性能局限及潜在触发条件是否导致了不可接受的危害行为。对于新识别出的、不可接受的危害行为，对危害行为接受准则进行补充完善。

7.2 总则

为了实现本章的目的，可考虑以下信息：

- 规范定义与设计，按照 5.5；
- 整车层面的危害，按照 6.6.1；
- 危害行为的风险评估，包括已识别出的合理可预见的间接误用，按照 6.6.2；
- 接受准则，按照 6.6.3；
- 基于外部信息或经验教训（例如，13.5）的，可导致危害行为的系统及其要素的已知潜在功能不足和已知潜在触发条件（包括合理可预见的直接误用）。

7.3 潜在功能不足与触发条件的分析

7.3.1 总则

对潜在的功能不足和触发条件进行系统性分析。该分析可考虑从相似项目或专家中获得的现场经验和知识。可从以下两方面同时进行该分析：

- 基于已知的潜在规范定义不足和性能局限，确定导致已识别出的危害行为的场景（包含触发条件）；
- 基于已识别出的环境条件和合理可预见误用，确定潜在的规范定义不足和性能局限。

注1：关于SOTIF分析技术的更多细节见附录B。此外，也可参考ISO 34502。

注2：归纳、演绎或探索性的方法可用于支持分析。

注3：可进行定性或/和定量分析。

注4：定量目标可向下定义至要素层面，其源自整车层面的接受准则或确认目标。

注5：对所有相关用例参数的适当抽象（例如，生成和使用等价类或优先度子集），有助于处理大量用例组合。

注6：交通统计数据可用于那些可导致潜在危害行为的合理用例。

注7：可通过仿真来支持分析，例如，使用蒙特-卡罗方法。

可使用表4中所列举的方法，进行适当的方法组合，来识别和评估潜在规范定义不足、性能局限、输出不足和触发条件。

表4 潜在功能不足和触发条件的分析方法

方法	
A	需求分析
B	设计运行区域ODD、用例和场景分析 ^a
C	事故统计分析 ^b
D	边界值分析
E	等价类分析
F	功能相关性分析
G	常见触发条件分析 ^c
H	来自现场经验和经验教训的潜在触发条件分析 ^d
I	系统架构分析（包括冗余）
J	传感器设计与潜在技术局限分析 ^e

方法	
K	算法及其输出或决策分析
L	系统老化分析 ^f
M	车辆运行周期内可能的环境变化分析（例如，干扰）
N	内部和外部接口分析 ^g
O	执行器设计和潜在局限分析
P	事故场景分析 ^h
Q	合理可预见误用分析 ⁱ
<p>^a 包含 ODD 边界分析。</p> <p>^b 例如 STATS19 (UK) [6]、GIDAS (Germany) [7]、GES (US) [8]、CARE[9]、IGLAD[10]。</p> <p>^c 单一触发条件可以激活多个性能局限或规范定义不足（例如大雨能够影响如雷达、摄像头等不同传感器的性能）。</p> <p>^d 考虑市场上相似系统、先前系统和项目以及客户索赔分析。</p> <p>^e 考虑技术上的局限（例如由于相机成像器、雷达天线设计局限或缺少如密封和振动等环境隔离而导致的较低分辨率）和由于安装位置所导致的技术局限（例如由于传感器未能覆盖车辆周边全部 360 °视野所导致的盲区）。</p> <p>^f 例如在规定范围内由于老化效应而变得暗淡的相机镜头。</p> <p>^g 例如车辆与车辆之间、车辆与基础设施之间、OTA 地图。</p> <p>^h 例如基于自动驾驶数据存储系统/事件数据记录器（DSSAD/EDR）中记录进行分析。</p> <p>ⁱ 表 5 列出了分析方法。</p>	

注8：安全分析方法可用于识别和评估潜在功能不足和触发条件及其对危害的影响（例如，原因树分析、事件树分析（ETA）、归纳性的SOTIF分析或危害与可操作性分析（HAZOP））。附录B. 3提供了对安全分析方法进行调整的示例。

基于系统架构，要素的潜在功能不足可分类为：

——单点功能不足；或

——多点功能不足。

该分类有助于确定充分的功能修改，以实现SOTIF（见第8章）。其可用于导出要素层面所需的要求，以实现整车层面的SOTIF（见第5章）。

示例1：如果整车层面实现了 SOTIF 特定的接受准则，则可将性能目标分配给不同的相关要素，如图 15 所示。与单个传感器系统相比，每个传感器可以分配到较少的约束性性能目标（例如，误触发探测率）。



图15 具两个不同传感器融合的系统架构示例

在定义确认策略过程中，也可以使用这种分类，其中，根据独立性的考虑，可以降低多点功能不足的确认目标（见第9章和附录C. 6. 3）。

对于给定的性能局限或规范定义不足，可能存在多个触发条件导致危害行为。此外，已知的环境条件和合理可预见的误用可激活多个整车层面或要素层面的性能局限或规范定义不足。需建立并维护危害行为、触发条件与整车层面或要素层面潜在性能局限或规范定义不足之间的追溯性。

图16给出了危害、触发条件与整车层面或要素层面潜在的性能局限或规范定义不足之间关联的说明和示例。

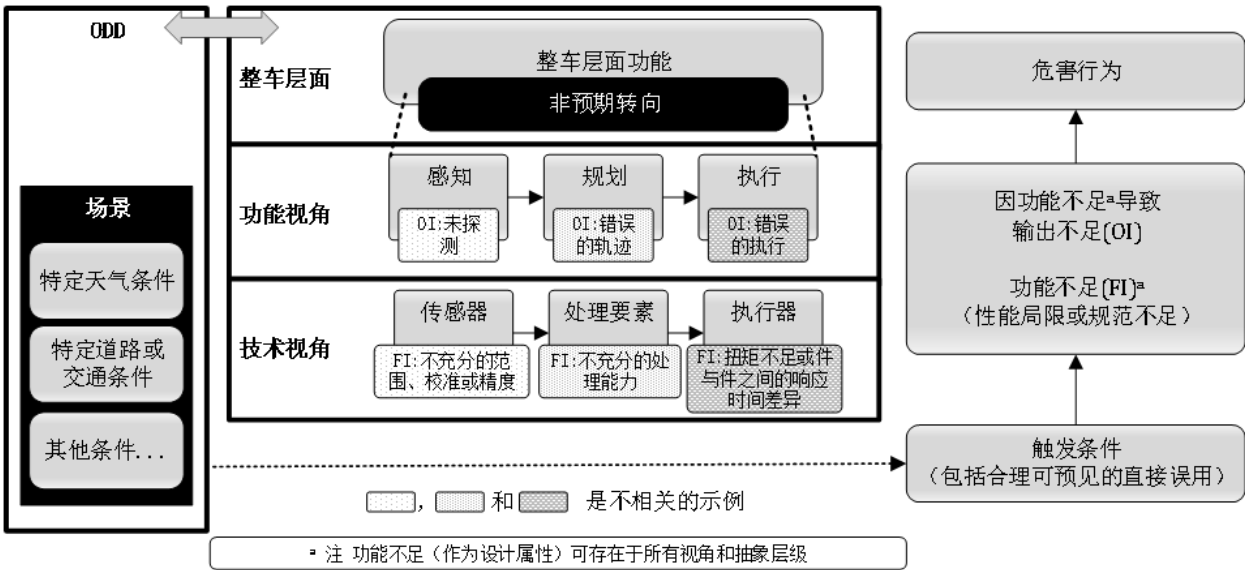


图16 潜在功能不足与触发条件之间关联的说明

为了方便展示，在下面子章节分别对规划算法、传感器和执行器进行说明。如果有益，传感器和执行器中的潜在功能不足和触发条件也可用于规划算法分析，反之亦然。

7.3.2 与规划算法相关的潜在功能不足和触发条件

分析可考虑以下类别：

- 环境与位置；
- 道路基础设施；
- 城市或乡村基础设施；
- 高速基础设施；
- 驾驶员或用户行为（包括合理可预见的误用）；
- 其他驾驶员或道路使用者的潜在行为；
- 驾驶场景（例如，施工现场、事故、紧急通道上交通堵塞、车辆行驶方向错误）；
- 已知的规划算法局限（例如，无法应对可能的场景或不确定的行为）；
- 已知的机器学习的规范定义不足；
- 已知的机器学习的测量数据不足；
- 已知的功能不足和功能改进。

7.3.3 与传感器和执行器相关的潜在功能不足和触发条件

分析可考虑以下类别：

- ODD；
- 天气条件；
- 机械干扰（例如，因传感器在车上的位置导致的震动而引起的传感器噪声输出）；
- 传感器上的污垢；
- 电磁干扰（EMI）；
- 来自其他车辆或其他来源的干扰（例如，雷达或激光雷达）；

- 声音干扰；
- 眩光；
- 低质量的反射；
- 精度；
- 范围；
- 响应时间；
- 由于耐久性、磨损、老化导致的性能影响；
- 权限能力（适用于执行器，例如，液压制动系统预期的最大可用制动压力）；
- 多传感器数据融合；
- 传感器的校准与安装。

示例1：雨和雪可影响雷达性能。

示例2：车辆前方升起的太阳可影响视频摄像头的性能。

示例3：行人身上厚重的羊毛外套可影响超声波传感器的性能。

示例4：不正确的校准可影响许多个传感器类型。

注1：已知的潜在功能不足和触发条件的某种组合可能导致某个潜在的危害行为。

注2：详细分析类别见附录B。对于每个分类，详细的干扰列表可基于知识和经验（包括来自相似项目和现场经验的知识积累）来制定。

注3：如果由基础设施要素所提供的传感器输入与自动驾驶（AD）或ADAS功能相关，则本子章节也适用于这种情况，以分析功能不足。

注4：附录D. 5给出了驾驶自动化系统常见传感器的性能局限示例。

此外，在可能的数值范围内（包括潜在的和观察到的场景），对每个环境的输入进行系统性分析。

7.3.4 合理可预见的直接或间接误用分析

预期功能的合理可预见的直接和间接误用可导致不合理的风险水平。

一方面，作为潜在触发条件分析的一部分，第7章覆盖了针对直接误用的分析。另一方面，可能导致针对间接误用的措施无效的潜在功能不足也在第7章的范围内。

合理可预见的直接或间接误用的原因可以是：

- 用户缺乏对系统的理解，例如，驾驶员被市场中某个具有不同操作规程的相似系统误导；
- 用户对系统的错误期望，例如，呈现给驾驶员的不充分、不恰当或不正确的信息；
- 注意力不足；
- 对系统的过度信赖；
- 系统设计中对用户交互的错误假设。

可使用表5中描述的方法来支持对合理可预见的误用的分析。

此外，附录B. 1描述了一种导出SOTIF误用场景的方法。

表5 识别合理可预见的误用的方法

方法	
A	来自现场经验和其他来源的经验教训的已知误用场景分析 ^a
B	测试项目相关的研究
C	用例和场景分析
D	用户与系统交互的分析 ^b
E	HMI分析
F	已知的人员行为模式分析，例如，缺乏使用、误用和对自动化水平过度自信

方法	
G	针对执行任务或在任务间切换的人员能力分析 ^e
H	相关标准、法规和指南的应用 ^d
<p>^a 例如，一些用户视频，演示了系统或其他相似系统如何以一种合理可预见的方式被误用。</p> <p>^b 例如，驾驶员提醒，系统理解，运行模式的混淆。</p> <p>^c 例如，人员重新获得态势感知的能力的分析。</p> <p>^d 例如，ADAS 设计和评估的代码实践[11]、欧洲人机界面原则声明[1]。</p>	

注1：详细方法见附录B.1。

注2：在有需要时，驾驶员无法保证驾驶任务的情况下使用车辆，被认为是一种滥用，不属于本文件的范畴。

示例1：驾驶员受到管制药物的影响。

示例2：在雪地上，以超出车辆动态控制能力的不合理高速行驶。

为防止或减轻合理可预见误用（间接或直接的），所需的额外措施及其有效性，可在预估系统对潜在触发条件的响应的可接受度时进行评估。可在验证和确认阶段对这些措施的有效性进行论证。

7.4 预估系统对触发条件的响应的可接受度

对包含已识别出的触发条件的场景进行评估，以确定是否实现SOTIF。

注1：这些已知场景被第10章中的验证活动所覆盖，用于提供其可接受度的最终评估。

注2：特别地，本评估中使用的或由此产生的与实现SOTIF相关的假设，见第10章。

注3：在评估过程中所考虑的假设可能包括系统及其要素的预期行为，或假设的用户行为。

在以下情况下，无需进一步功能修改（见第8章），SOTIF被视为可以实现的：

——导致危害事件的系统残余风险低于 6.5 中定义的接受准则；

注4：注风险评估中所使用的证据将会在验证与确认活动中生产（第9、10和11章）。

——没有可导致特定道路使用者的不合理风险的已知场景。

注5：即使车队遇到包含某个触发条件的场景的概率很低，但如果特定单个车辆遇到此类场景的概率很高，则系统的响应也可能是不可接受的。

示例：一种集成在环形交叉口或桥墩中的特殊结构，系统性地引起 AEB 系统制动，从而导致与跟随车辆发生追尾碰撞的概率达到不合理的水平。

根据上述条件，如果系统对触发条件的响应是不可接受的，将启动进一步的功能修改（见第8章）。

7.5 工作成果

7.5.1 已识别出的潜在规范定义不足、性能局限和触发条件（包括合理可预见的直接误用），满足 7.1 a）。

注：为满足7.1 a而进行的分析的报告包含在7.5.1中。

7.5.2 评估系统对已识别出的触发条件的响应对于实现 SOTIF 的可接受度，满足 7.1 b）。

8 修改功能以解决 SOTIF 相关风险

8.1 目的

本章的目的是实现以下目标：

- a) 应确定并实施解决 SOTIF 相关风险的措施；
- b) 应更新“规范定义和设计”的输入信息。

8.2 总则

为了实现本章的目的，可考虑下列信息：

- 规范定义和设计（按照 5.5 章）；
- 危害行为的风险评估（按照 6.6.2 章）；
- 识别潜在的规范定义不足、性能局限和触发条件（按照 7.5.1 章）；
- 已知场景的验证和确认结果（按照 10.8 章），如果有；
- 未知危害场景的确认结果（按照 11.4.1），如果有；
- SOTIF 发布的论证（按照 12.5 章），如果有。

8.3 改进 SOTIF 的措施

8.3.1 介绍

在满足下列条件的情况下，可开展第8章中关于制定解决SOTIF相关风险的措施的活动（以下简称“SOTIF措施”）：

- 当前规范定义和设计（第 5 章）中的预期功能，被识别出存在危害场景，需要深入分析（如同在危害事件的风险评估中所评估的，将造成伤害）（第 6 章）；
- 系统对已识别出的可导致危害行为的触发条件的响应被评估为是不可接受的（存在已知场景，其中，引起危害事件的残余风险不满足接受准则并导致不合理的风险）（见第 7 章）。
- 通过反复应用第 8 章中的 SOTIF 措施，更新规范定义和设计（第 5 章），并使用更新的规范定义和设计对预期功能进行风险评估（第 6 章和第 7 章），从而完善系统。

接下来，在V&V阶段对完善的系统（包括SOTIF措施的有效性）进行评估，且如果满足以下任一条件，则可通过第8章在设计阶段进行系统改进的迭代活动：

- 已知危害场景的残余风险被确定是不可接受的（第 10 章）；
- 未知危害场景的残余风险被确定是不可接受的（第 11 章）；或
- 残余风险被认为是不可接受的（第 12 章）。

在上述情况下，重复第5章和第8章对系统进行改进。

对SOTIF “避免”或“缓解”措施的恰当组合进行选择，以实现SOTIF相关风险的降低。

注：“避免措施”代表本质安全的设计措施，其首要任务是消除风险（旨在实现第6章风险评估中的S=0或C=0），修改功能（特别是新增功能）是一种典型的方法。但是，这并不一定意味着将实现S=0或C=0。

当在避免风险方面存在已知的困难时或认为能将风险降低到可接受的水平时，可以考虑应用“缓解措施”来尽可能多的降低风险，并期望通过“缓解措施”与避免措施或其他缓解措施的结合，提升风险降低的效果。

在实施SOTIF措施时，可考虑以下几点：

- 对其他要素没有不利的影响；
- 与其他危害场景没有相互作用。

此外，即使精心设计并实施SOTIF措施，也可能不会产生预期的结果，并可能产生非预期的后果。因此，实施第13章所述的监控和评审是实现SOTIF措施的重要组成部分，以确保SOTIF措施的有效性。

第8.3.2条~8.3.5条描述了可能的SOTIF措施。

8.3.2 系统修改

系统修改措施的目的是尽可能地维护预期功能。这些措施包括但不限于：

- a) 通过下列方法提高传感器性能和/或精度：

- 改进的传感器技术；

示例1：提高传感器测量的精度；

示例2：更新为新的和改进后的传感器，以解决已知的局限。

——改进传感器扰动探测机制，以触发适当的报警和降级策略；

——多样化传感器类型；

示例3：增加额外的感知装置，以适当的方式提高覆盖率。

——改进的传感器标定和安装；或

示例4：针对具有潜在的性能局限的临界情况，将传感器安装在合适的位置提升覆盖率。

示例5：封装传感器以避免或减少干扰，从而达到可接受的水平。

示例6：进行传感器覆盖率分析，并优化传感器的选择(类型、技术、数量)及其在车内的相对位置。

——传感器遮挡探测及清洗方法；

示例7：使用边缘探测方法探测摄像头上的污垢，并用液体和雨刷器清洗。

b) 通过改进执行器技术，提高执行器的性能和/或精度(例如，提高精度，延长或限制输出范围，减少响应时间，可重复性，对能力进行仲裁，利用其他功能来辅助或增加新的执行器来辅助)；

c) 通过算法改进，提高识别和决策算法的性能和/或精度；

示例8：改进的传感器识别算法(例如，改进相机图像中探测对象的特征描述，如 HOG(定向梯度直方图))。

示例9：考虑模型中额外的输入信息。

示例10：改进算法以提供更好的鲁棒性、精度(例如，从线性模型切换到非线性模型或使用机器学习)(见附录 D.2)。

示例11：随着计算能力的提升，加快图像处理速度。(例如，使用机器学习加速器或运行高效的硬件)。

示例12：退出 ODD 的识别。(例如，识别高速公路出口坡道的方法)。

示例13：识别已知的不支持的环境条件(例如，根据地理位置、一天中的时间、季节等，预测遇到太阳眩光的情况)。

注：在执行高级算法时，可考虑提升硬件性能。

d) 突出自车的可见性，以在自车发生危害行为时，提高其他交通参与者的可控性。

示例14：在当地法规允许的情况下，安装反光板、开雾灯、开转向灯、主动发出声响等。

8.3.3 功能限制

功能限制措施的目的是通过预期功能的降级(或限制)来维持部分功能。这些措施包括但不限于：

a) 特定用例的预期功能限制；

示例1：当车道探测装置不能清晰地探测到车道时，车道保持辅助功能限制了转向辅助扭矩，以避免出现不期望的转向干预。

示例2：ODD 的限制，包括环境的、地理的或时段的限制。

示例3：限制或约束驾驶策略(见附录 D.1)以确保制定决策的安全性。

示例4：摄像头被午后阳光引起的周围光线的反射导致失明；通过雷达和其他传感器进行有限的操作(例如降低允许的最大车速，限制车道保持功能施加的最大转向扭矩)。

b) 解除特定用例的预期功能权限；

示例5：所有的感知传感器都被暴风雪导致失明，驾驶员被要求接管控制。

示例6：自动驾驶车辆不能处理收费站收费或无标识的施工区域，驾驶员被要求接管控制。

8.3.4 权限移交

将权限从系统移交给驾驶员的措施旨在提高较低自动化水平的可控性。这些措施包括但不限于：

a) 修改人机交互(HMI)；

示例1：HMI 清晰地向驾驶员传递移交请求，为驾驶员提供必要的信息，以支持驾驶员实现适当的态势感知并执行该任务。

b) 修改用户通知和 DDT 后援策略；

示例2：当系统探测到视野受限(例如，因泥浆导致距离传感器感知范围缩小)时，车速会降低，相应的 HMI 要求驾驶员接管驾驶任务。如果驾驶员在规定的时间内没有执行接管，系统将把车速降为零。

注1：针对不同的自动驾驶等级，权限移交可能不同。

注2：只有在过渡本身可控且不会给驾驶员带来额外风险的情况下，才能提高可控性。

注3：可考虑对HMI研究的指南。

示例3：ADAS 设计和评估的实施规范^[11]。

8.3.5 解决合理可预见的误用

解决合理可预见的误用的措施包括但不限于：

a) 客户教育（信息和培训）；

示例1：用户手册、培训课程、市场营销、销售演示。

b) 改进 HMI；

示例2：通过提供正确的操作信息来帮助驾驶员。

c) 驾驶员监控和报警系统的实施。

注：探测和警告驾驶员注意力不集中的系统，是自动驾驶车辆系统中防止合理可预见的驾驶员误用的有效方法。有效的驾驶员监控系统的选择和实现，取决于目标的误用情况。

示例3：当方向盘被松开时，警告驾驶员。

示例4：忽略可能导致危害行为的输入/命令，并将原因告知驾驶员。

d) 实施防止误用的措施。

示例5：尽管发出警告信息，如果驾驶员监控系统探测到仍持续误用，则可采取措施阻止危害行为；例如，在多次脱手警告后，车道保持辅助功能可在后续的行程中解除或降级，并提供适当的警告信息，直到下一个驾驶循环。

示例6：合理可预见的误激活功能，例如，在车速过高时激活驻车辅助功能，可通过在功能的激活条件中增加速度限制来防止误用。

8.3.6 支持 SOTIF 措施实施的考虑

在实施SOTIF措施后，根据自动驾驶等级，监控和评审对于确保SOTIF措施的有效性非常重要的，为了支持这一点，在设计系统时可考虑一些方面。这些考虑包括但不限于：

——对 SOTIF 相关系统行为的可测性；

——对 SOTIF 相关系统行为的诊断能力；

——对 SOTIF 相关系统行为的数据监控能力。

8.4 更新“规范定义和设计”的输入信息

“规范定义和设计”的输入信息，是基于按照8.3章中已识别出的和实施的SOTIF措施的规范来更新的。

8.5 工作成果

SOTIF 措施的定义，满足8.1 a) 和 b) 。

9 定义验证和确认策略

9.1 目的

本章旨在实现以下目标：

a) 应定义 SOTIF 的验证和确认策略，包括确认目标，并应考虑：

1) 对潜在危害场景的必要的评估；

- 2) 对相关场景空间的充分覆盖;
 - 3) 必要的证据(例如, 分析结果、测试报告、专门调查);
 - 4) 生成证据的过程。
- b) 应提供所选验证和确认方法和确认目标的适用性理由。

9.2 总则

为实现本章的目标, 可考虑以下信息:

- 传感器或外部数据源(例如, 来自基础设施)提供足够准确的环境信息以满足性能要求的能力;
- 假设的外部数据源的可依赖性的充分性(例如, 通信网络突然中断或暂时没有更新的可能性);
- 传感器处理算法精确模拟环境的能力;
- 决策算法的能力:
 - 安全地处理潜在的功能不足;
 - 根据环境模型、驾驶策略和当前目标(例如, 目的地)做出适当的决策;
- 系统或功能的鲁棒性, 例如:
 - 系统对不利环境条件的鲁棒性;
 - 自动化系统对已知触发条件的反应的适当性;
 - 预期功能及其对不同场景条件的监控的敏感性;
- 不存在因预期功能的危害行为引起的不合理风险;
- 系统(例如, HMI)防止合理可预见的误用的能力;
- 系统安全处理 ODD 之外用例的能力(例如, ODD 之外的系统激活、向 ODD 外的过渡等);
- OEDR 的适用性, 以及跨 ODD 执行驾驶策略(或行为)的鲁棒性;
- DDT 后援的适用性; MRC 的适用性;
- 在运行阶段, 在整车层面上对接受准则的符合性具有足够的信心。

为实现本章的目标, 可考虑以下信息:

- 规范定义和设计, 按照 5.4;
- 危害行为的风险评估, 按照 6.6.2;
- 接受准则, 按照 6.6.3;
- 已识别出的潜在规范定义不足、性能局限和触发条件(包括合理可预见的直接误用), 根据 7.5.1;
- SOTIF 措施的定义, 根据 8.5;
- 系统集成和测试计划(来自外部来源);
- 现场监测过程中的经验教训, 根据 13.5;
- 在传感器的历史中观察到的经验教训, 可能来自其他领域(例如, 大气风暴事件导致 GNSS 信号延迟, 会潜在导致危害事件)。

验证和确认策略不仅关注ODD内部的性能评估和风险识别, 还关注ODD的边界和外部。该策略的一个方面包括验证系统不能在ODD之外的任何地方激活。

另一个方面是验证从ODD内部到ODD外部的过渡是否伴随着上升给驾驶员或后援系统, 以实现最小风险状态。

注: 这些方面对于论证相关场景空间的充分覆盖非常重要。

9.3 集成和测试的定义

定义了验证和确认策略，以提供实现目标的论据，以及如何实现确认目标。验证和确认策略涵盖了车辆的全部预期功能，包括E/E要素和其他被认为与实现SOTIF相关的技术要素。验证和确认策略也支持对SOTIF相关的外部来源的数据进行监控。

定义确认目标是为了提供证据，证明符合接受准则。根据选择的确认方法，可以通过多种方式确定确认目标。

对于从表6、7、8、9、10、11或其他来源中选择的每个方法，定义了适当的开发工作（例如累积测试长度、分析深度）。提供了定义每项工作的理由。这可能包括场景的数量或分布、试验数量或模拟持续时间。

注1：接受准则解决已知和未知危害场景产生的风险。在推导确认目标时考虑了这一点，区域2和区域3的确认目标可能不同。

注2：C. 2和C. 6给出了定义和评估接受准则和确认目标的示例。

示例1：考虑对与功能相关的先前未知的触发条件进行搜索。确定确认目标是为了支持剩余未知触发条件不会带来不合理风险的假设。

示例2：确认目标可以使用被测功能预定义的假阳率和假阴率来设置。

如果只是场景的子集与特定危害相关，那么在确定目标值和确认持续时间时，可以考虑对该子集的暴露度。

注3：表B. 5和表B. 6分别提供了基于定性规则和定量规则的场景优先度子集的示例。

注4：在评估触发条件违反定量目标的可能性时，可考虑所产生行为的暴露度、可控性和严重度。这可以减少证明暴露在触发条件下所需的工作量。有关通过考虑暴露度、可控性和严重度来减少确认工作的方法，请见C. 2. 1。

示例3：考虑来自 6. 3 的示例 3 中，只有存在跟随车辆，意外制动才会导致追尾。在制定确认目标时，可考虑跟随车辆的暴露度。

注5：在验证和确认策略的定义和阐述中，考虑了触发条件参数的可变性。

注6：由于通过预期功能安全（SOTIF）活动迭代（图8）进行了功能修改，因此对系统进行分析，以确定现有功能是否受到功能迭代修改的影响，并通过回归测试重新测试这些功能。这将确保功能修改不会在现有功能中导致潜在的危害行为。如果有适当的理由，回归测试的范围可以做裁剪。

注7：为确保保持正确的功能行为，对于任何释放用于生产的版本都要文档记录完整的验证和确认（V&V）活动。这包括记录未修改要素的文档和受变更影响的重新测试要素的文档。

注8：附录D. 2. 4讨论了离线训练（如用于机器学习）的验证和确认活动。考虑到集成级别，验证和确认策略（例如，集成测试用例、分析）的规范可以使用适当的方法组合导出，如表6所示。

表6 验证和确认活动的导出方法

方法	
A	需求分析
B	外部和内部接口分析 ^a
C	等价类的生成与分析
D	边界值分析
E	基于知识或经验的错误猜测法
F	功能的相关性分析
G	常见限制条件和次序的分析
H	环境条件和操作用例分析 ^b
I	现场经验和教训分析 ^c
J	系统架构（包括冗余）分析

方法	
K	传感器设计及其已知潜在局限性分析
L	算法及其决策路径以及各自的已知局限性分析
M	系统和部件老化分析 ^d
N	触发条件分析
O	性能目标分析 ^e
P	危害分析中可测量参数的分析
Q	边界值中角临界情况和边临界情况的分析 ^f
R	现有系统的SOTIF相关更新的分析
S	使用包含收集到的测试用例和场景的数据库
T	场景和用例的优先度子集分析和使用
U	接受准则的分析
V	事故场景数据分析
W	执行器中已知潜在限制的分析
<p>^a 如果可以的话，还包括 V2x，地图。</p> <p>^b 包括系统或其要素潜在危害行为的已知来源。</p> <p>^c 这考虑了各种驾驶条件、驾驶风格、驾驶环境和终端用户要求。</p> <p>^d GB/T 34590 通常会考虑导致失效的半导体老化效应。与 SOTIF 相关的半导体老化效应，即影响标称性能的老化效应，都在本文件的范围内。</p> <p>^e 性能目标可以在不同的抽象级别上指定，例如传感器级别（雷达的探测范围、摄像头的角分辨率）以及系统级别（例如目标探测的误报率）。</p> <p>^f “角临界情况是指两个或多个参数值均在系统能力范围内，但合起来共同构成挑战系统能力的罕见情况。边临界情况是指极端值或者甚至一个或多个参数的存在都会导致挑战系统能力的场景条件。” [12]</p>	

注9：有关汽车感知系统验证和确认的进一步实践，请见C.4。

9.4 工作成果

实现目标9.1 a) 和b) 的验证和确认策略的定义。

10 已知场景的评估

10.1 目的

本章的目的是实现以下目标：

- a) 对已识别出的潜在危害场景应评估其是否具有危害性；
- b) 对于已知危害场景和合理可预见的误用，系统及其要素的功能表现应符合其定义的方式；
- c) 对于由已定义的整车层面行为导致的潜在危害行为，应评估其可接受度；
- d) 根据验证和确认策略，已知场景应被充分覆盖；
- e) 验证结果应证明确认目标得到了满足。

注：这包括对动态驾驶任务后援和最小风险状态（MRC）的适当性的评估。

10.2 总则

为实现本章的目标，可以考虑以下信息：

——规范定义和设计，按照 5.4；

- 识别出的潜在规范定义不足、性能局限和触发条件（包括合理可预见的直接误用），按照 7.5.1；
- 解决 SOTIF 相关风险的措施，按照 8.5；
- 验证和确认策略的定义，按照 9.4。

注：对于“规范定义和设计”中已识别的预先存在的 SOTIF 相关内容，以及由SOTIF活动迭代导致功能修改的可追溯性，5.3给出了指导。

10.3至10.5条的结构遵循了4.2.3中介绍的感知(10.3)、规划（10.4）和执行(10.5)的模式。10.6涉及集成方面。

10.3 感知的验证

用于证明感知部分的预期用途，以及合理可预见的误用所表现的正确的功能性能、时间、准确性和鲁棒性的方法，如表7所示。

注1：有些问题可以分配给不同的验证活动，例如，目标分类可以被视为规划算法的一部分（见10.4）。在这种情况下，可以应用不同章条规定的验证方法。

表7 感知的验证

方法	
A	验证传感器定义的充分性（例如，范围、精度、分辨率、时序约束、带宽、信噪比、信干比的充分性） ^a
B	基于需求的测试（例如：分类、传感器数据融合）
C	注入触发功能不足的输入 ^b
D	对选定的SOTIF相关用例和场景，结合已识别的触发条件进行在环测试（例如，SIL、HIL、MIL） ^c
E	对选定的 SOTIF 相关用例和场景结合已识别的触发条件进行实车测试 ^c
F	在ODD范围内，不同环境条件下的传感器测试（例如：低温、潮湿、光照、能见度条件、干扰条件）
G	验证传感器老化影响（例如加速寿命测试等） ^d
H	评估来自该传感器，或此类传感器的现场经验（包括现场监控）
I	通过对已知危害场景进行回注仿真，以验证已实施的风险缓解机制的效果
H	验证架构属性，包括触发条件之间的独立性（如果适用）
^a 这还包括传感器组装期间的终检线测试（EOL）（例如：雷达天线和雷达天线罩之间的安装校准、摄像头成像器与摄像头镜片的安装校准）。	
^b 在一些测试用例中，可以通过在仿真层面进行错误注入的方式，来模拟传感器某个潜在的功能不足。同时需要提供错误模型可以代表所测试现象的理由。仿真的结果可以与触发条件的分析结果相结合。	
^c 使用已经识别出的传感器模型的局限性来选择测试环境（HIL/SIL/MIL 或实车）。	
^d 在某特定传感器有着行业共识的老化故障模型的情况下，传感器老化效应的验证可部分在模拟仿真中完成。	

注2：对于测试用例的生成，可以采取组合测试的原则^[13]。

注3：附录C.4提供了感知传感器验证的示例。

10.4 规划算法验证

根据4.2.3，规划算法基于感知部分提供的环境模型推导出控制动作。用于验证规划算法在需要时做出反应的能力，及其避免不期望动作的能力的方法，如表8所示。

表8 规划算法验证

方法	
A	对于输入数据不受其他来源干扰的鲁棒性验证，例如：白噪声、音频、信噪比降级（如通过噪声注入测试）
B	基于需求的测试（例如：场景分析、功能、传感器数据的可变性） ^a
C	验证架构属性，包括触发条件的独立性（如果适用）
D	对选定的SOTIF相关用例和场景，结合已确定的触发条件进行在环测试（例如：SIL/HIL/MIL）
E	对选定的SOTIF相关用例和场景，结合已确定的触发条件进行实车测试
F	注入触发潜在危害行为的输入
G	验证是否正确遵守驾驶策略【例如，实现最小风险状态（MRC）和退出ODD时的操作】 ^{a, b}
H	通过对已知危害场景进行回注仿真，以验证已实施的风险缓解机制的效果
^a 包括车辆选择并实现了适当的最小风险条件（MRC）的验证。 ^b 附录 D.1 中介绍了驾驶策略指南。	

注：对于测试用例的推导，可以采用组合测试的原则^[13]。

10.5 执行的验证

验证执行器的预期用途和合理可预见的误用的方法，如表9所示。

表9 执行验证

方法	
A	基于需求的测试（例如准确性、分辨率、时序约束、带宽）
B	验证执行器被集成在整车环境中或系统测试台架中时的特性
C	不同环境条件下的执行器测试（例如：低温条件、潮湿环境）
D	不同载荷条件下的执行器测试（例如：从中等载荷变化到最大载荷）
E	验证执行器老化的效应（例如加速寿命测试） ^a
F	对选定的SOTIF相关用例和场景，结合已确定的触发条件进行在环测试（例如：SIL、HIL、MIL）
G	对选定的SOTIF相关用例和场景结合已确定的触发条件进行实车测试。
H	验证架构属性，包括触发条件的独立性（如果适用）
I	通过对已知危害场景进行回注仿真，来验证已实施的风险缓解机制的效果
^a 对于某特定执行器具有行业共识的老化故障模型时，可以使用仿真来完成一部分执行器老化效应的验证。	

注：如能表明执行系统没有任何功能不足或触发条件，仅根据GB/T 34590或其他相关领域特定标准进行测试是充分的。

10.6 系统集成验证

验证集成到整车中的系统的鲁棒性和可控性，以及系统组件的正确交互，可应用表10所示的方法。

表10 集成系统验证

方法	
A	验证系统鲁棒性（例如通过噪声注入测试） ^a
B	在整车集成环境或系统测试台架上进行的基于需求的测试（例如：性能目标和行为特征，可测量参数、范围、精度、分辨率、时序约束、带宽）

方法	
C	对选定的SOTIF相关用例和场景，结合已确定的触发条件进行在环测试（例如：SIL、HIL、MIL）
D	不同环境条件下的系统测试（例如：低温、潮湿、光照、能见度条件、干扰条件）
E	验证系统老化影响（例如：加速寿命测试）
F	定向随机输入测试 ^b
G	对选定的SOTIF相关用例和场景，结合已识别的触发条件进行实车测试。
H	可控性测试（包括合理可预见的误用）
I	验证内部和外部接口
J	车载传感系统特性验证 ^c
K	验证架构属性，包括触发条件的独立性（如果适用）
L	通过对已知危害场景进行回注仿真，来验证已实施的风险缓解机制的效果。
<p>^a 这还包括整个 ODD 和 OEDR 的鲁棒性的验证，以及包括退出 ODD 在内的最小风险条件策略的稳健执行的验证。</p> <p>^b 预期的现实世界场景通常很难重现，因此随机输入测试可以作为替代，例如如下情况：</p> <ul style="list-style-type: none">——图像传感器添加翻转图像或更改的图像块；——雷达传感器添加虚假目标以模拟多路径返回；——雷达传感器因多车雷达干扰增加虚假目标或丢失检测目标。 <p>^c 这包括不同传感器在不同运行条件下的工作（例如：当一种传感器技术的能力不足时，如雾或挡风玻璃反射率影响摄像头、保险杠/标志的形状和油漆类型影响雷达）和传感器位置的误差。</p>	

注1：对于非确定性系统的验证，可以使用统计方法或风险管理技术进行已知危害场景的评估。

示例：驾驶策略行为依赖于对道路参与者的假设，特别是在存在遮挡时，在某些情况下遵循已知的非危害行为可能会导致碰撞。

注2：附录C.4提供了系统集成验证的示例。

10.7 已知危害场景导致的残余风险的评估

第9章中定义的确认目标，提供了在运行阶段以足够的置信度满足接受准则的论据。因此，验证结果表明，已知危害场景的验证目标已经实现，已知危害场景的残余风险并非不合理。

如满足下列条件，已知的危害场景是可以接受的：

- 已知场景导致危害行为的概率符合验证目标；
- 不存在可能导致特定道路使用者，面临不合理的风险的已知场景。

示例：当地地理属性（例如，某条隧道或桥梁）不会导致风险的不合理增加。

10.8 工作成果

验证和确认结果，以表明预期功能在已知场景中按照期望运行，实现10.1章的目的。

11 未知场景的评估

11.1 目的

本章的目的是确认结果应证明来自未知危害场景的残余风险满足接受准则并具有足够的置信度。

注：一个方面是整个V&V活动集合对可能场景空间的代表性覆盖。

11.2 总则

为了实现目的，可以考虑以下信息：

- 规范定义和设计，按照 5.4；
- 已识别出的潜在规范定义不足、性能局限和触发条件（包括合理可预见的直接误用），按照 7.5.1；
- 解决 SOTIF 相关风险的措施，按照 8.5；
- 验证和确认策略的定义，按照 9.4；
- 表明预期功能在已知场景中按照期望运行的验证和确认结果，按照 10.8。

11.3 未知场景残余风险的评估

现实生活中可遇到未知场景。为评估可触发集成到整车上的系统产生危害行为的现实生活场景的残余风险，可应用表11所示的方法。

表11 残余风险的评估

方法	
A	对信噪比降级的鲁棒性的确认（例如：通过噪声注入测试）
B	确认架构设计的效果和特性，包括触发条件的独立性（如果适用）
C	采用随机测试用例（源自技术分析和错误猜测）的在环测试
D	随机输入测试 ^a
E	考虑已识别的触发条件，对选定的测试用例（源自技术分析和错误猜测）进行整车层面测试
F	长期车辆试验
G	车队道路测试
H	基于现场经验的测试
I	极端场景和边缘场景的测试 ^b
J	与现有系统的比较
K	随机场景集合的仿真
L	对随机使用和新手驾驶员潜在误用的测试
M	考虑场景特定条件的功能敏感度分析 ^c
N	相关参数的分析/仿真 ^d
O	现实世界里的场景发掘 ^e
P	功能分解和概率建模（即考虑一个要素的不足条件由其子要素的多个输出不足组成；见C.6.3.3）
Q	相较于真值的确认

^a 预期的现实世界场景通常很难重现，因此随机输入测试可以作为替代，例如如下情况：

- 图像传感器添加翻转图像或更改的图像块；或者
- 雷达传感器添加虚假目标以模拟多路径返回；或者
- 雷达传感器因多车雷达干扰增加虚假目标或丢失检测目标。

^b 极端场景是指两个或两个以上的参数值都在系统的能力范围内，但共同构成了挑战其能力的罕见情况。边缘场景指由于系统处于极值状态，或者系统的一个或多个参数导致挑战系统能力的情况。[12]

^c 如果该情况的微小变化可能导致车辆整车上显著不同的行为，则功能被视为对场景的特定情况很敏感。

^d 见第 7.3 条的第 5、6 和 7 条。第 7.3.1 条所述的触发条件列表可用于识别相关用例参数。

^e 发掘的方法是通过现实世界场景覆盖一个多样性集合，来搜索未知的场景。这可以包括系统地或随机地改变场景的相关参数。参数选择是通过敏感性分析、统计分析等方法来论证的，以证明所选参数是相关的。

对于在公共区域进行的测试，可能需要采取额外的安全措施（例如：急停装置），以防止或减轻试验车辆对公众造成的潜在风险。

注1：每次引入变化，如算法变化、ODD变化、OEDR变化、在原环境中导入新车型和驾驶策略变化等变化时，都会出现新的未知危害场景。一旦引入了这些变化，表11中的方法也可以用于重新评估残余风险。

所选择的一组方法足以在区域3中识别潜在的危害场景，例如，通过使用典型用例的输入，以及通过关注具有挑战或罕见的运行环境、特定的用例、情景或场景。同时，需要提供所选方法充分性的理由。

车辆测试里程的确定（例如，长期测试、车队道路测试）可以考虑以往整车项目的经验、驾驶员可控性或选定测试路线的关键度。当使用带有错误注入的随机输入测试时，可以选择一定数量的仿真场景，来匹配代表目标地理市场所需的测试里程和内容。

当考虑不同的测试方法，如场地测试、仿真测试或开放道路测试，为每种测试方法分配适当的公里数或运行时长，并为这种分配提供理由。

注2：由于仿真是对不完整真实世界的模拟，即使决策算法的连续随机模拟循环可以模拟数百万公里的操作，但可能与真实世界暴露概率的权重不同。

注3：C.4提供了SOTIF相关系统确认的示例。

根据第9章，选择确认目标的方式是，通过确认目标的实现，可确保满足可接受准则。在这些条件下，由于未知的危害场景而产生的残余风险是可以接受的。

示例：确认目标可以是在一组测试场景中，遇到的以前发生的未知危害场景的最大数量。如果在执行完这些测试场景之后，遇到的以前发生的未知危害场景的数量小于定义的目标值，可认为确认目标被满足了。

11.4 工作成果

11.4.1 未知危害场景的确认结果，以满足 11.1 规定的目的。

11.4.2 残余风险的评估，以实现 11.1 规定的目的。

12 SOTIF 成果的评估

12.1 目的

本章的目的是实现以下目标：

- a) 应对 SOTIF 活动产生的工作成果的完整性、正确性和一致性进行评审。
- b) 应根据本文件各章的目的及相应工作成果的完成情况，为 SOTIF 的实现情况提供论证；及
- c) 应对 SOTIF 的成果的论证进行评估，并提出批准或拒绝 SOTIF 发布的建议。

12.2 总则

为了实现本章的目的，可考虑下列信息：

- 规范定义和设计（依照 5.5）；
- 整车层面的危害（依照 6.6.1）；
- 危害行为风险评估（依照 6.6.2）；
- 接受准则（依照 6.6.3）；
- 识别的规范定义不足、性能局限和触发条件（依照 7.5.1）；
- 系统对触发条件响应的评估（依照 7.5.2）；
- SOTIF 措施的定义（依照 8.5）；
- 验证和确认策略的定义（依照 9.4）；
- 验证和确认的结果，以表明在已知危险场景中预期功能的行为符合预期（依照 10.8）；

- 未知危害场景的确认结果（依照 11.5.1）；
- 残余风险的评估（依照 11.5.2）；
- 现场监控流程（依照 13.5）。

12.3 评估 SOTIF 的方法和准则

每项工作成果应就其完整性、正确性和一致性进行检查。

基于第5到11章的目的以及第13章中定义的现场监控措施（例如流程和必要的硬件资源）的实现情况进行论证，以表明SOTIF得到了实现。

注1：使用GSN方法论证框架示例，见A.1。

对该论证的评估包括但不限于以下几个方面：

- a) 是否对危害、潜在的功能不足和触发条件进行了分析，以及是否实施和评估了为实现 SOTIF 而做的所有必要的设计更改，以确保这些设计更改已根据所有指定用例中的接受准则充分降低了风险？
- b) 预期功能是否达到了最小风险状态，在必要时为驾乘人员或其他道路使用者提供了一种没有不合理风险的状态，考虑以下方面：
 - 1) 已定义的驾驶员的干预；
 - 2) 合理可预见的误用；
 - 3) 已定义的对驾乘人员和/或其他道路使用者的警告；
 - 4) 已定义的功能降级；
 - 5) DDT 后援（以达到最小风险状态）？
- c) 验证和确认策略是否覆盖所有已知的危害场景，并且是否提供具备足够置信度的论据以证明未知危害场景中的残余风险满足接受准则？
 - 1) 测试结果是否覆盖已识别的触发条件，包括环境条件以及直接和间接的误用？
 - 2) 在验证和确认策略中，是否包含充分的确认活动来限制由已知和未知场景带来的风险？
- d) 是否完成了充分的验证和确认，并满足确认目标，以确信不存在不合理的残余风险？
 - 1) 是否充分执行预期功能以评估常见行为和潜在危害行为？
 - 2) 对危害行为而言，是否有证据证明不存在不合理的风险？
 - 3) 测试是否提供了足够的覆盖度作为论据来证明驾驶策略在所有用例和/或 ODD、OEDR 中具备鲁棒性？
- e) 是否具有实现运行阶段活动所需的方法（根据第 13 章）？

注2：如果第13章中描述的运行阶段活动导出了SOTIF措施，则应在第12章的活动中评审这些措施。

示例：见 C.2.2。

注3：对SOTIF活动结果的检查可与GB/T 34590.2-XXXX的功能安全评估一并考虑。

12.4 SOTIF 发布推荐

根据12.3所述方法论的证据，可以确定“接受”、“有条件接受”或“拒绝”SOTIF发布的建议。在“有条件接受”的情况下，条件将被记录并在最终发布前验证其是否被满足。

注：“有条件接受”是一种过渡状态。在这种情况下，条件将被记录并在最终发布前验证其是否被满足。即：当满足条件时可接受最终发布。

示例：作为长期耐久测试的一部分，可以根据 6.5 中定义的接受准则设置行驶里程的中间目标值。如果满足所有条件，则判定为可接受。如果解决 SOTIF 问题所做的设计改进而产生的回归测试未完成，但其余条件均满足，则适用于有条件接受。回归测试成功完成后，可以发布。

需记录对SOTIF实现的评估。

12.5 工作成果

SOTIF发布论证，以实现12.1规定的目的。

13 运行阶段的活动

13.1 目的

- a) 应在发布之前定义现场监控流程，以确保运行期间的 SOTIF；及
- b) 应执行现场监控流程，以保持运行阶段的预期功能安全的实现。

13.2 一般要求

第5至12章中描述的SOTIF活动，目的是在SOTIF发布时将风险降低到可接受的水平。然而，在一些情况下风险评估可能需要被重新考虑，例如：

- 在功能运行期间，现场发现了一个以前从未识别的危害；
- 在功能运行期间，现场发现了以前从未识别的功能不足和/或触发条件；
- 与功能开发期间定义的假设相比，环境条件或交通规则等假设发生了变化。
- 为实现本章的目的，可考虑以下信息：
- 第 5 章中定义的规范定义和设计；
- 第 6 章中定义的接受准则；
- 第 7 章中定义的已识别的潜在规范定义不充分，潜在的性能局限和触发条件（包括合理可预见的误用）；
- 第 10 章中定义的验证活动的结果；
- 第 11 章中定义的确认活动和残余风险评估的结果。

运行阶段的活动范围见图17。

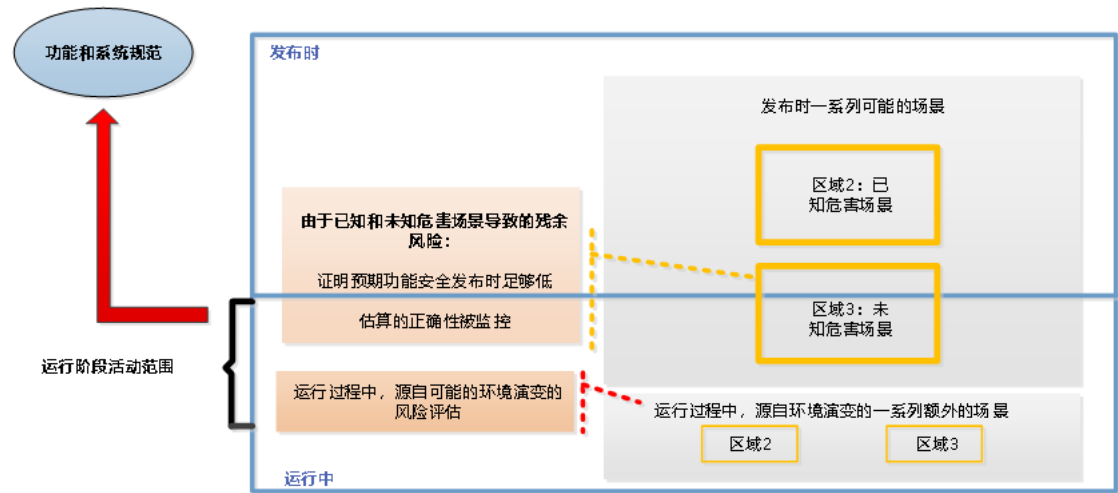


图17 运行阶段的活动范围

注：为全生命周期实现SOTIF所必需的、保持与规范定义和设计相一致的有关活动，包括GB/T 34590.7-XXXX涵盖的生产、运行和服务活动，不在本章范围内。

13.3 与观察相关的主题

现场监控流程的预期取决于驾驶自动化等级、预期功能的复杂性以及危害的严重性。对于较低的驾驶自动化等级,普通的市场监管可能就足够了。对于较高的驾驶自动化等级,额外的手段可能是必须的,例如,自动驾驶数据存储系统(DSSAD)/汽车事件数据记录系统(EDR)、车载/远程安全监控系统。

观察的内容包括但不限于:

- a) 功能已造成或潜在会造成伤害的事件,或功能已超过定义值并在不同情况下可能导致伤害的事件,

示例1: 这些事件可能包括:

- 事故或事件报告;
- 驾驶员报告里声称的问题;
- 车载措施信号传递的潜在弱点;
- 合理可预见的误用报告;
- 违背了与障碍物的最小距离;
- 接近触发特定系统响应的场景。

注: 对于高等级的驾驶自动化系统,可能需要实施监控机制,例如车载监控。这些机制能在事故发生之前探测到潜在的功能不足(如:导致险肇事故的功能不足、引起要素层面输出不足的条件)。在这种情况下,需要在开发阶段中定义SOTIF车载监控机制的要求。

示例2: 车载监控机制可以:

- 捕捉触发紧急系统响应的场景;
- 捕捉驾驶员非预期接管的场景;
- 捕捉导致最小风险状态的场景。

- b) 知识体系,

示例3: 知识体系可能包括:

- 来自公共安全机构(包括其他车辆制造商)的市场上公开可用的、可能与该功能相关的事件;
- 来自其他相似系统设计或相似功能的经验教训。

- c) 可能影响 SOTIF 并可能导致重新考虑 SOTIF 评估的环境演变。

注: 环境演变代表了场景中的变化,这些变化包括但不限于运行范围的变化和用户的系统交互的变化。

示例4: 该演变可能包括:

- 道路和交通的演变;
- 法规修改;
- 基础设施改变;
- 新的用法及其误用;
- 道路使用者特征的演变;
- 通常的用户习惯的改变,或因使用系统而导致的改变。

13.4 SOTIF 问题评估和解决流程

在SOTIF的问题评估和解决流程中,定义了以下角色和职责:

- 将相关数据提供给涉及的团队,如:开发团队;
- 评估收集的数据以确定风险是否仍然合理;
- 如有必要,定义和推出确保 SOTIF 的措施。

运行阶段活动包括但不限于以下:

- a) 监控和分析

监控的步骤是根据13.3定义的观察相关主题进行连续监控。监控可能是被动的[见13.3的列项a)]和主动的[见13.3的列项b)、c)]。此外,在开发阶段,监控可能发现未识别的潜在危害场景。

如果观察到了任何SOTIF相关的事件，需要分析对SOTIF论证的影响，并对SOTIF论证的有效性再次评估。

注1：监控目标可以在开发阶段定义。

注2：SOTIF相关观察可被用于升级或丰富数据库，以用于支持进一步开发所需的SOTIF分析（经验教训）。

注3：关于SOTIF论证的示例见附录A。

注4：如有必要，SOTIF论证可以被更新。

b) 风险评估和危害缓解

如果预期功能安全论证不再有效，就需要评估风险。基于SOTIF观察的风险，来决定风险缓解的方法。为缓解不合理风险，快速响应可能是必须的。在执行相应SOTIF活动来最终修复以前，可以采取不需要任何额外SOTIF活动的措施（如通过空中升级来部分或完全抑制功能）。增加新的SOTIF措施并升级系统作为长期行动项是必要的，这需要执行额外的SOTIF活动，并进行一次新的SOTIF发布。SOTIF发布之后，必要的系统和功能更改，需要考虑5至12章进行开发。

注5：空中升级（OTA）能提供一种灵活方便的方法来实施修改，以便在运行阶段及时解决识别出的功能不足。

注6：对于量产后，通过空中升级（OTA）方式新增或更新驾驶自动化功能时，不能以具备现场安全监控系统为由，代替或部分代替SOTIF验证和确认活动。

13.5 工作成果

13.5.1 现场监控的流程，以实现 13.1 的目的 1）。

13.5.2 现场监控到的不合理风险事件和由此触发的 SOTIF 更新结果，以实现 13.1 的目的 2）。

附录 A
(资料性)

预期功能安全的通用指南

A.1 用 GSN 构建 SOTIF 论证的示例

A.1.1 总则

A.1提供了如何使用目标结构表示法 (GSN)^[14]来呈现SOTIF论据的两个例子。表A.1和A.2描述了GSN示例中使用的元素。论据可以使用不同的方式来组织。在A.1.2和A.1.3中给出了两种可能的,但不是唯一的结构示例。




GSN是在安全领域广泛应用的一种方法。GSN的目的是将支撑最高目标“不存在不合理风险”达成的合理性记录下来。即展示目标是如何被分解成子目标,并最终得到证据(解决方案)的支持,同时明确用于达成目标所采取的策略和背景信息。

注: GSN也可用于应对其他标准中的目标和目的,例如, GB/T 34590标准。

表A.1 使用的 GSN 元素的描述

符号	名称	描述
	目标	目标, 用矩形表示, 表示主张(构成论据的一部分)。
	策略	策略, 用平行四边形表示, 描述了目标与其支持目标之间存在推理的性质。
	解决方案或证据	解决方案或证据, 用圆形呈现, 表示对某一证据项的参考引用。
	背景信息	背景信息, 用左图表示, 表示一个上下文环境。这可以是对上下文信息的引用, 也可以是一个陈述。有时用于定义目标或策略中的术语。
	假设	假设, 用右下角带有字母A的椭圆形表示, 表示一个未经证实的陈述。
	证据支持连接线	证据支持连接线, 用一条带有实心箭头的线表示, 可以记录推断或证据关系。
	环境条件连接线	环境条件连接线, 用一条带有空心箭头的线表示, 声明环境条件关系。
	多重关系连接线	一种表示在实例上也许有多个相应关系的表达方式。实心球象征着(零或更多)。球旁边的标签表示数量关系。

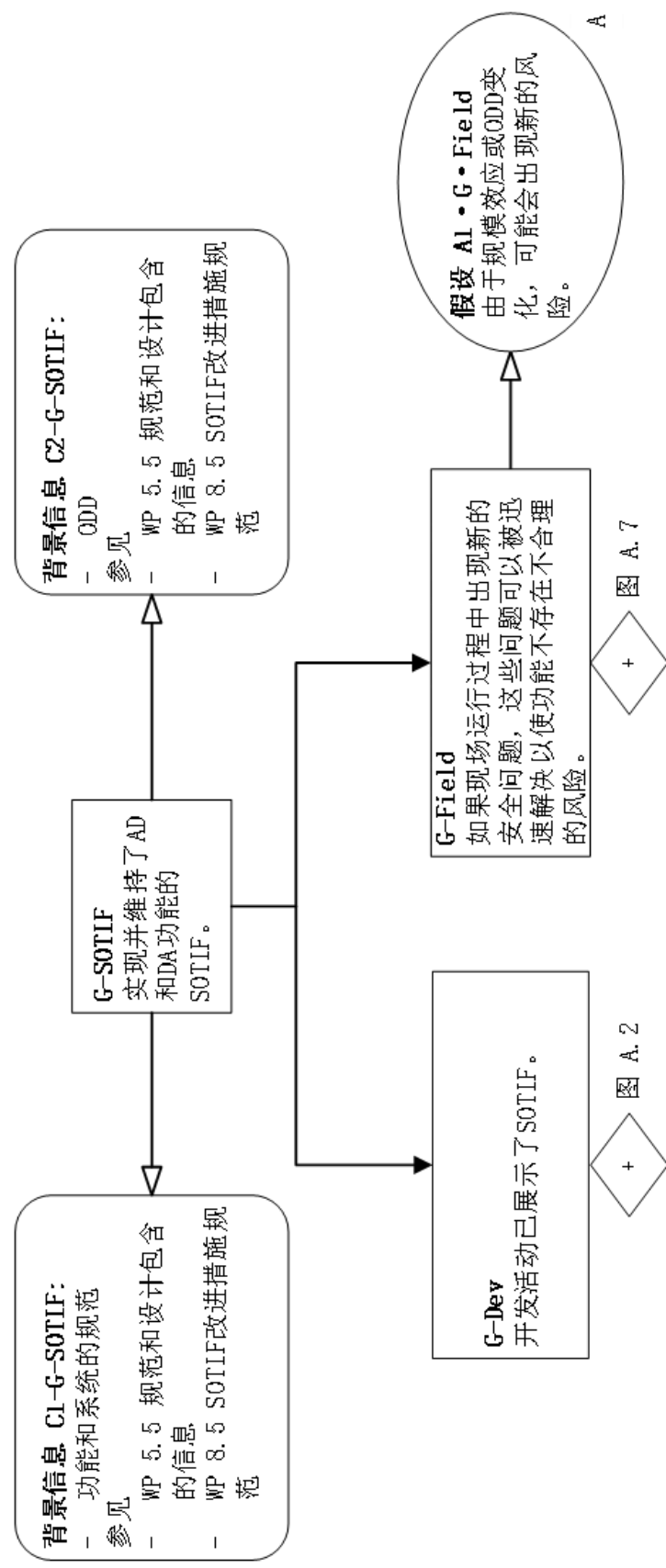
表A. 2 使用的 GSN 官方标准之外的标记元素的描述

符号	名称	描述
	保证主张点	引用有关于两个元素之间关系的论证的一种方法。 注：安全论证包含对以下信息的参考： 提供背景信息； 说明假设；及 展示证据。 这些参考资料的充分性和适当性可能被质疑。对这类质疑的解答将是对所参考信息是充分且适当的这一主张的论证。保证主张点（ACP）的使用是一种方便的语法形式，表明存在或需要一个支持的可信论证，同时不会使主论证图混乱。ACP背后的论据将由单独的图提供。
 图 A. X	图形参考	图A. X的参考指引，在其中继续进行论证。
	表格参考	表A. X的参考指引。

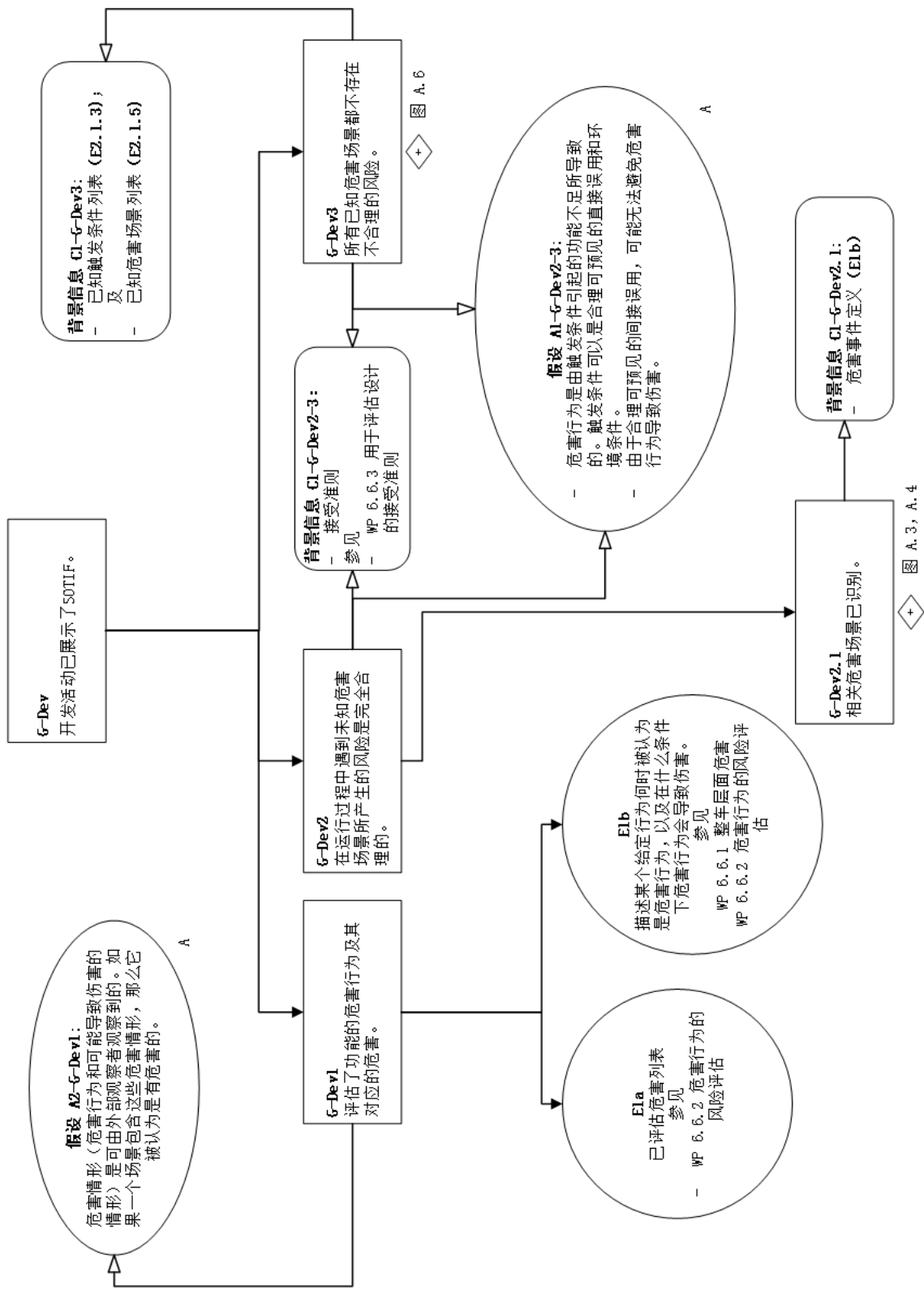
A. 1. 2 GSN示例1

示例1 GSN论证（图A. 1-A. 7）是基于不存在由于已知（即区域2）和未知（即区域3）的潜在危害场景所导致的不合理风险。

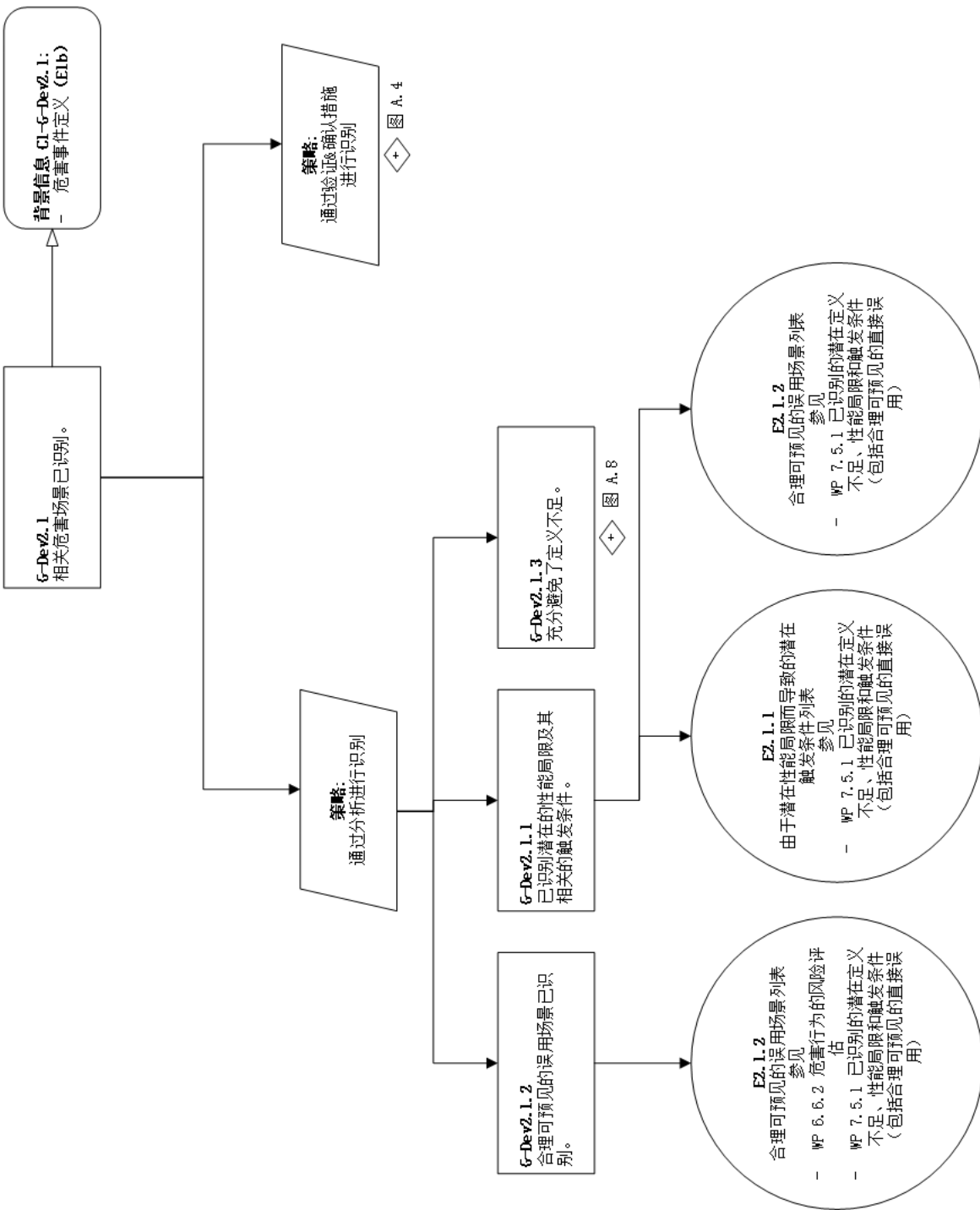
注：在GSN示例中，AD用作“自动驾驶”的首字母缩写，DA用作“驾驶辅助”的首字母缩写。



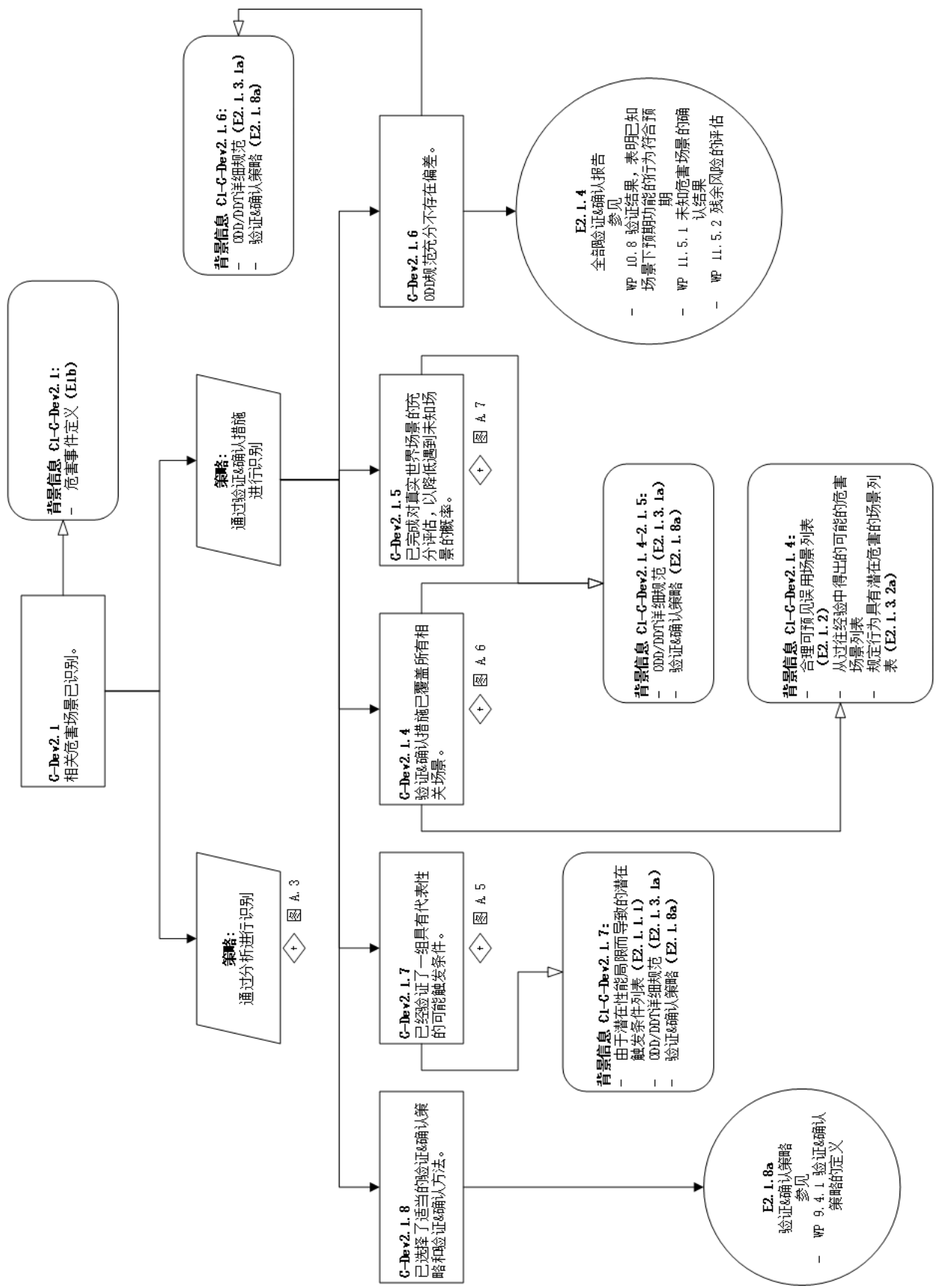
图A.1 G-SOTIF：实现并维持了AD和DA功能的SOTIF



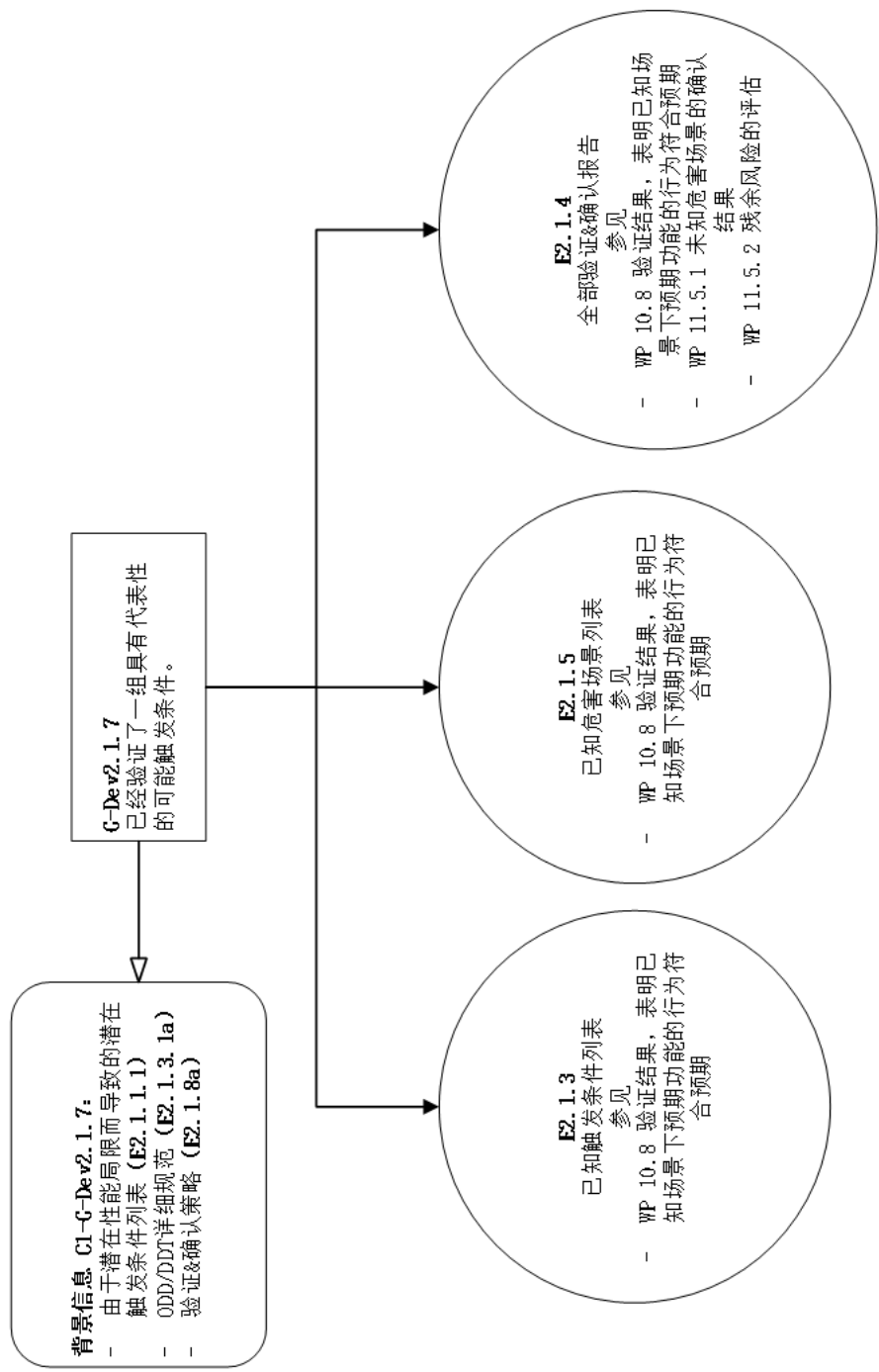
图A. 2 G-Dev: 开发活动已展示了 SOTIF



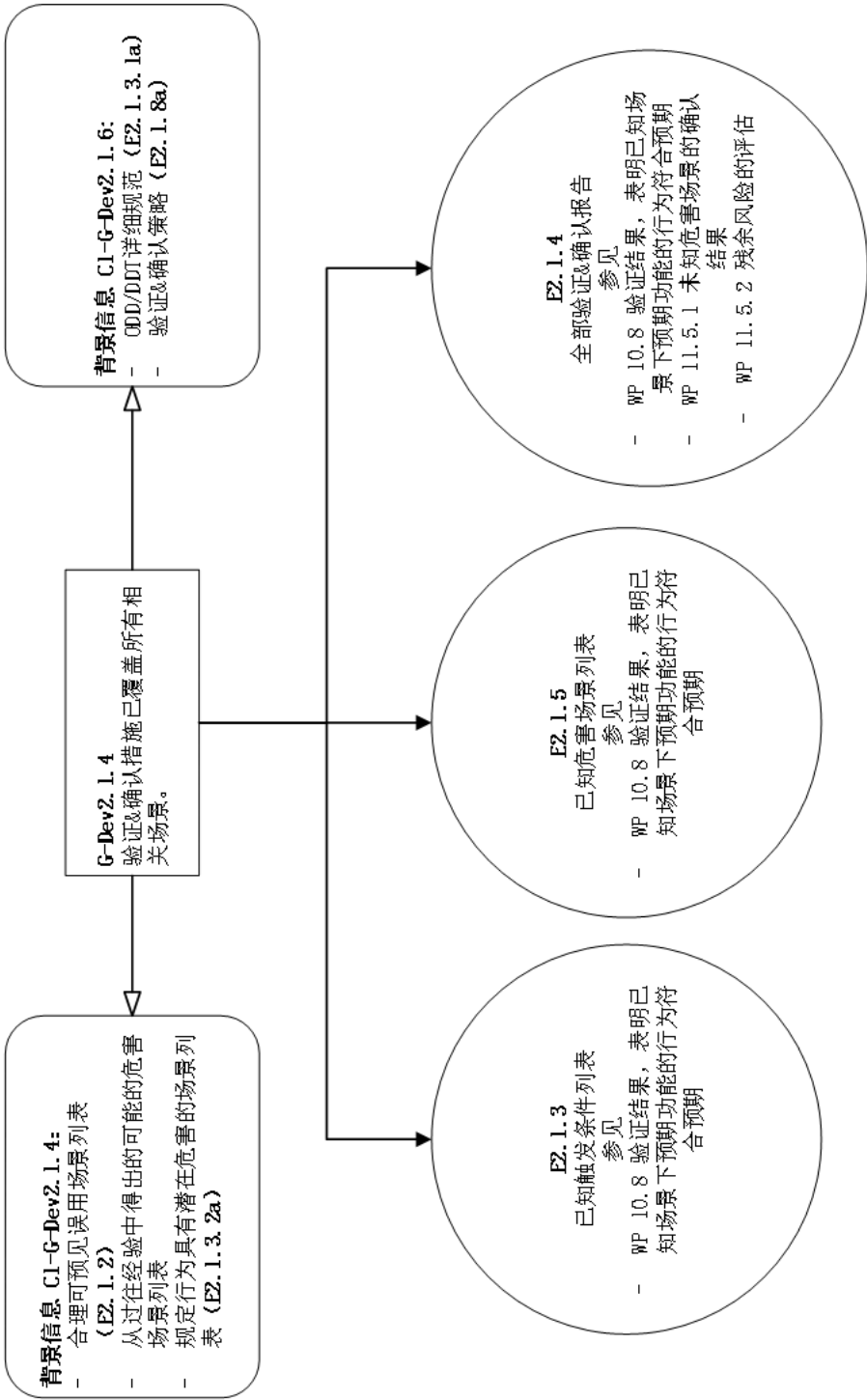
图A. 3 G-Dev2. 1： 相关潜在危害场景已识别-第一部分



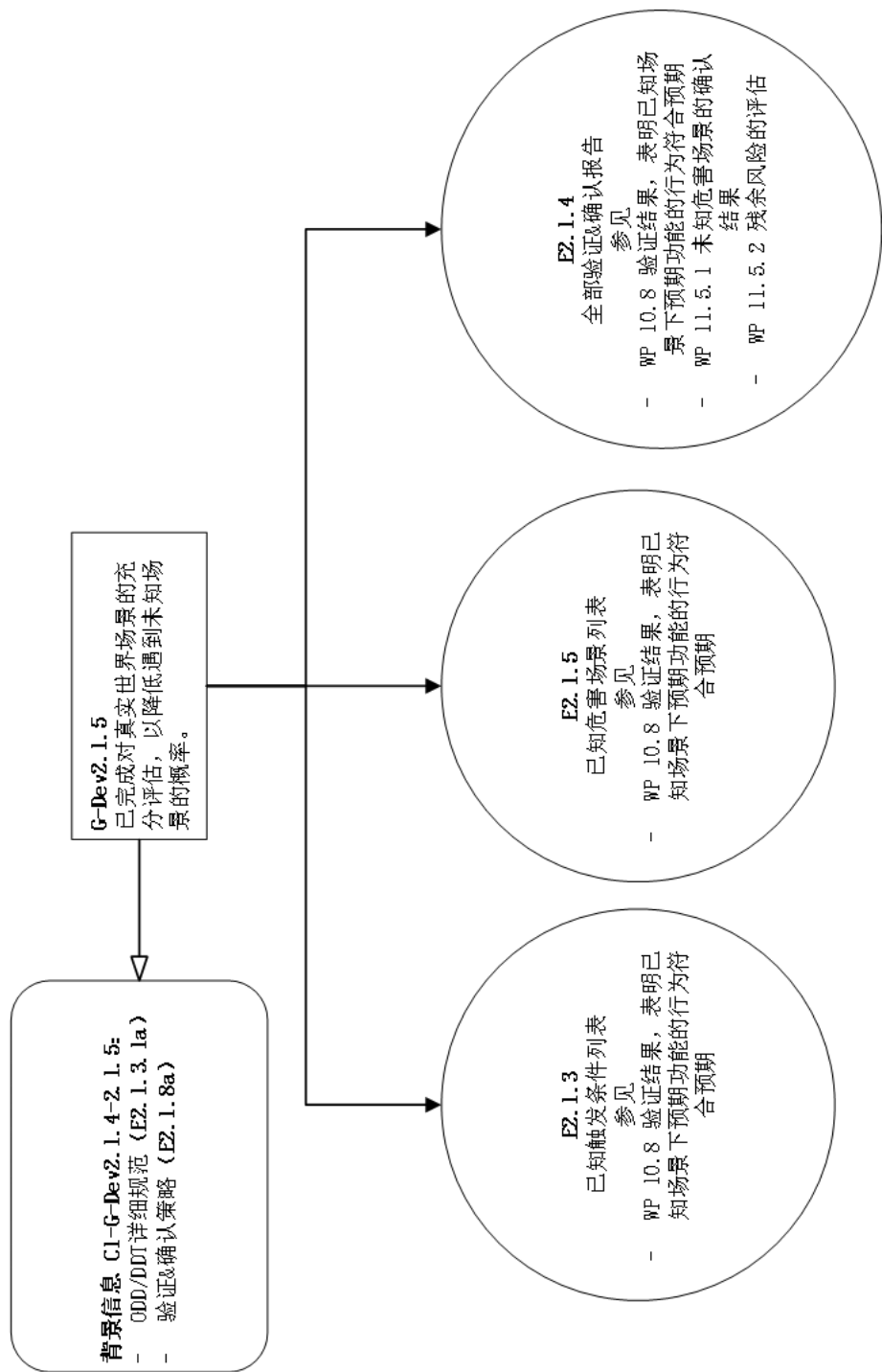
图A.4 G-Dev2.1：相关潜在危险场景已识别-第二部分



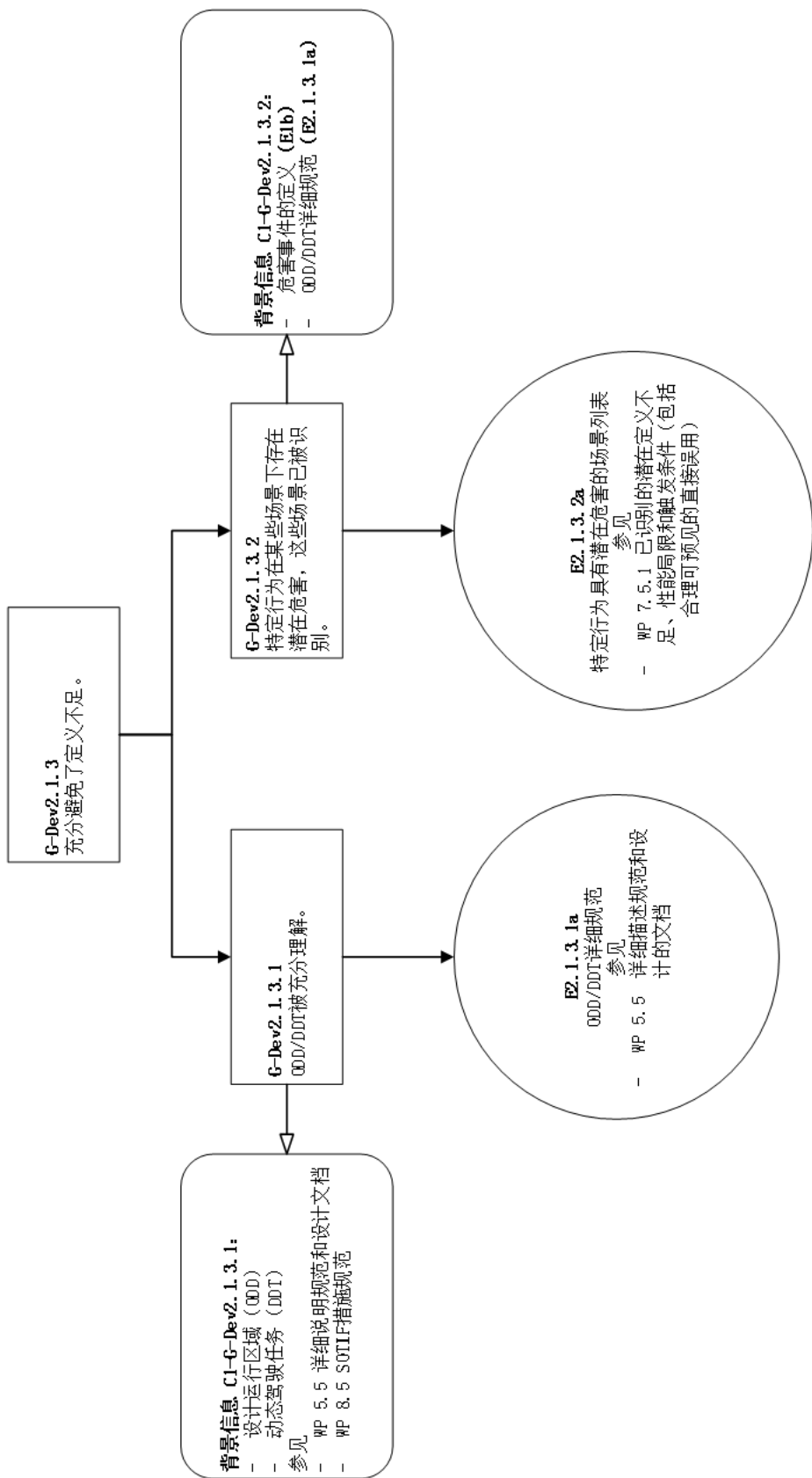
图A. 5 G-Dev2.1.7



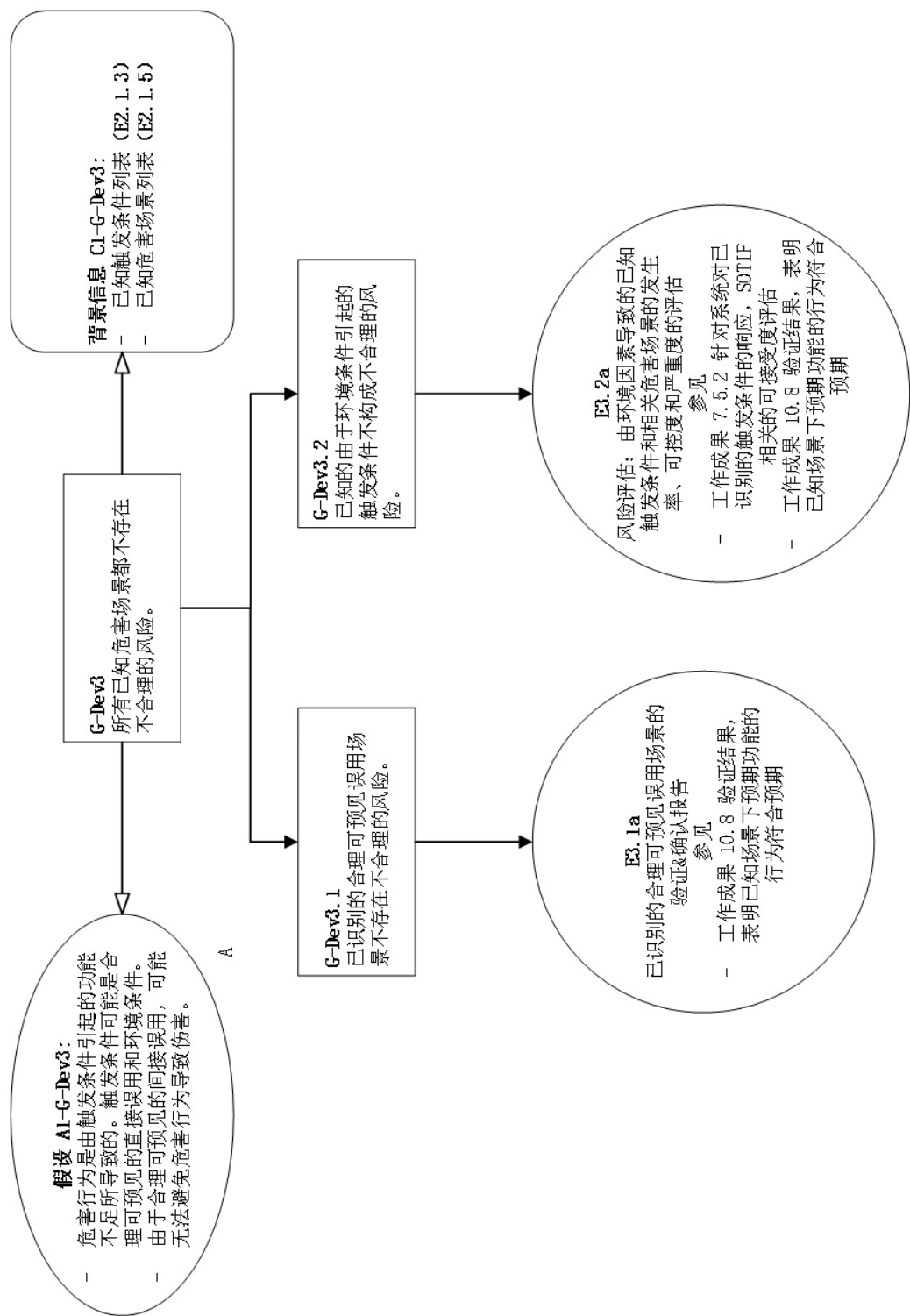
图A. 6 G-Dev2. 1. 4



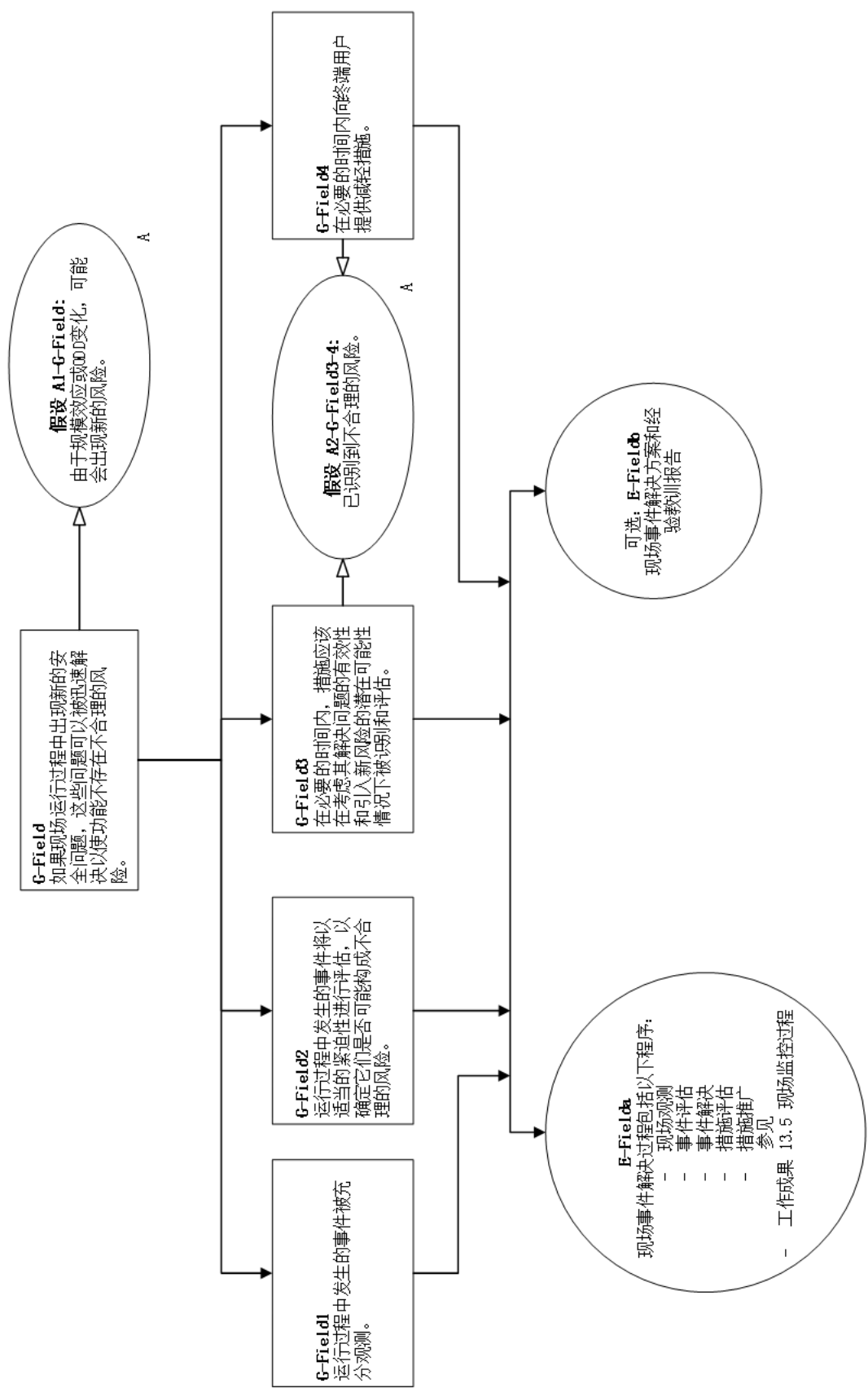
图A. 7 G-Dev2. 1. 5



图A.8 G-Dev2.1.3: 充分避免了定义不足



图A.9 G-Dev3: 所有已知潜在危害场景都不存在不合理的风险

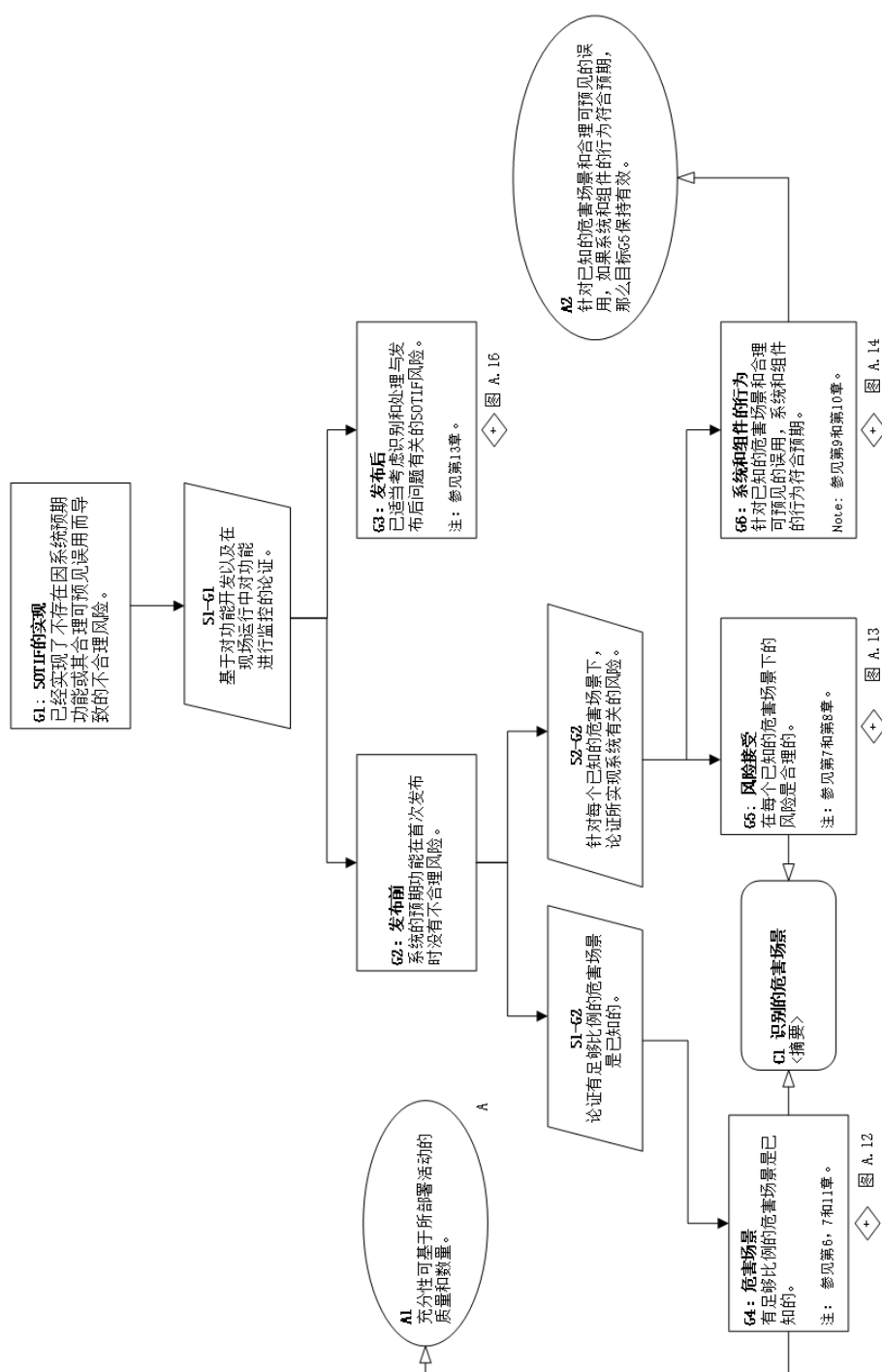


图A.10 G-Field：如果现场运行过程中出现新的安全问题，这些问题可以被迅速解决以使功能不存在不合理的风险

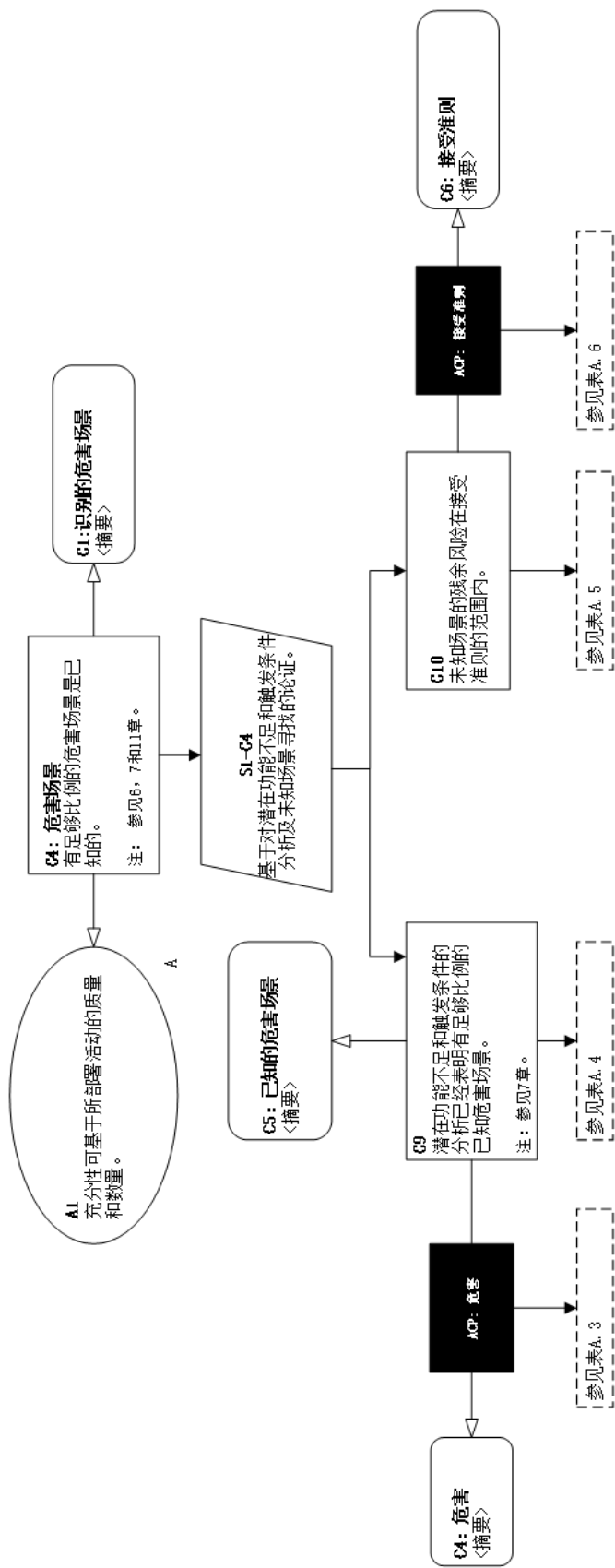
A. 1.3 GSN示例2

示例GSN(图A. 11到A. 16)显示了支持顶层目标“已经实现了不存在因系统预期功能或其合理可预见误用而导致的不合理风险”的论据结构。

所提出的论据结构具有通用性,适用于所有系统。其被分解到子目标,以进一步发展为特定的系统。在这一点上,参考了标准中提到的可以用来进一步阐释每个子目标和提供必要证据的主题。



图A. 11 G-Field: 已经实现了不存在因系统预期功能或其合理可预见误用而导致的不合理风险



图A.12 G4: 潜在危害场景

表A. 3 与 ACP 相关的主题:危害声明(已正确识别所有危害)

用于识别所有因功能不足导致的危害的方法的充分性
方法的定义
方法部署所耗费资源
风险评估的完整性和正确性
评审(根据第12章)SOTIF活动产生的证据, 以发现SOTIF实现过程中潜在问题的能力

表A. 4 与 G9 开发相关的主题(潜在功能不足和触发条件的分析已经表明有足够比例的已知危害场景)

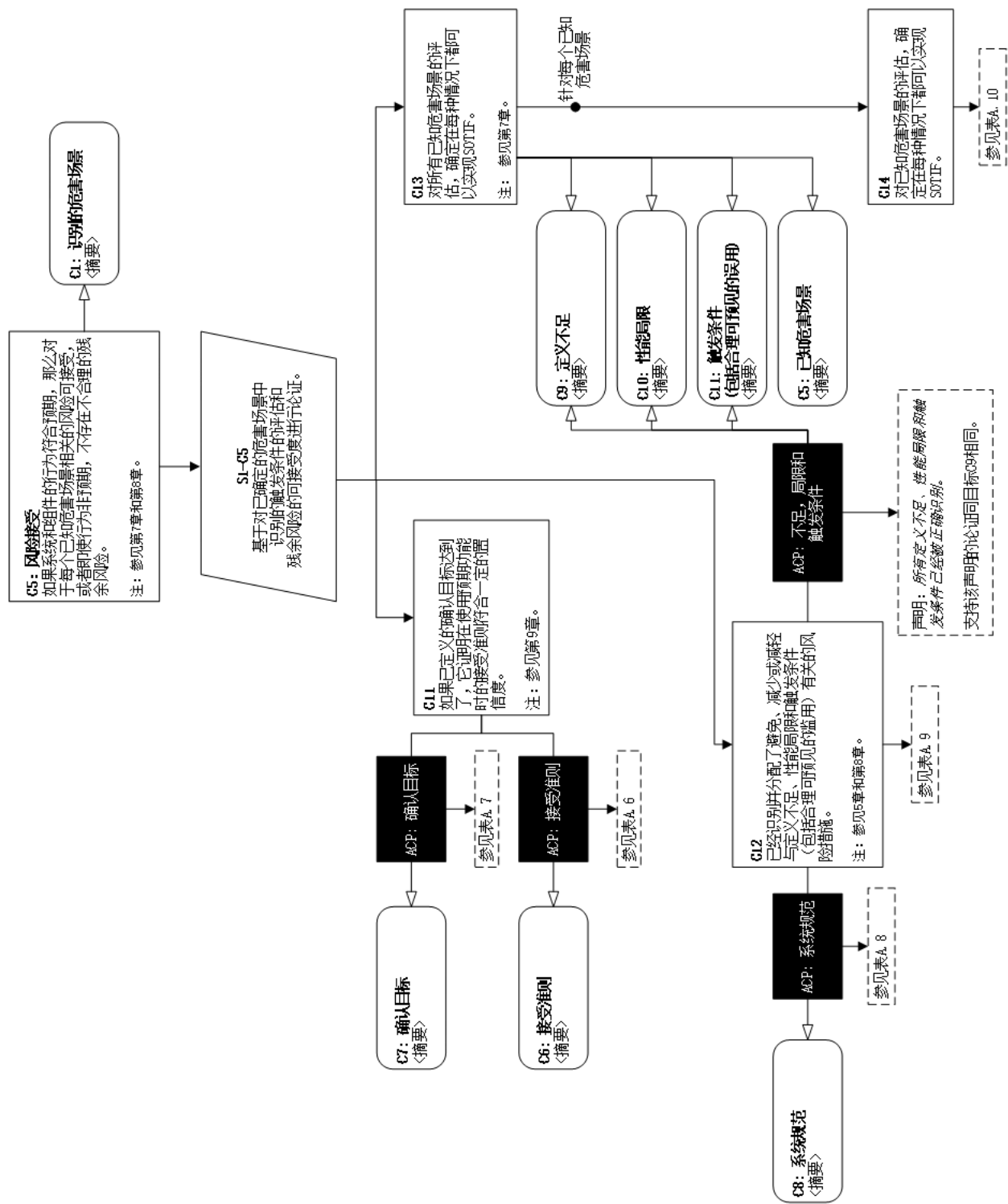
从类似项目中获得的知识
从现场经验中获得的知识
已知的潜在规范定义不足和性能局限
先前识别的环境条件和合理可预见的误用
用于识别所有潜在功能不足和触发条件的组合方法的充分性(表4)
每种方法识别特定潜在功能不足和潜在触发条件的能力(表4)
方法的定义(表4)
方法部署所耗费资源(表4)
与算法相关的潜在功能不足和触发条件的识别
与传感器和执行器相关的潜在功能不足和触发条件的识别
合理可预见的误用分析(表5)
评审(根据第12章)SOTIF活动产生的证据, 以发现SOTIF实现过程中潜在问题的能力

表A. 5 与 G10 开发相关的主题(未知场景的残余风险在接受准则的范围内)

车辆设计(例如安装位置)
用于揭示迄今未知场景方法的充分性(表11)
每种方法识别特定潜在功能不足和潜在触发条件的能力(表11)
方法的定义(表11)
方法部署所耗费资源(表11)
对新识别场景的处理

表A. 6 与 ACP 相关的主题:危害声明(已正确定义可接受准则)

符合已定义的可接受准则
工作被认为是充分的
适用的政府和行业标准法规
用于证明SOTIF的置信度的定义
对目标市场可用交通数据的使用(C. 2. 2. 4)
来自类似功能现场运行的已有准则的使用
选择目标的基本原理, 如GAMAB, ALARP, MEM



图A.13 G5: 风险接受

表A.7 与 ACP 相关的主题: 确认目标声明 (已正确设置确认目标)

暴露于一个场景子集
在评估触发条件时使用暴露度、可控性和严重度

表A.8 与 ACP 相关的主题: 系统规范声明 (系统规范已被完整且正确地定义)

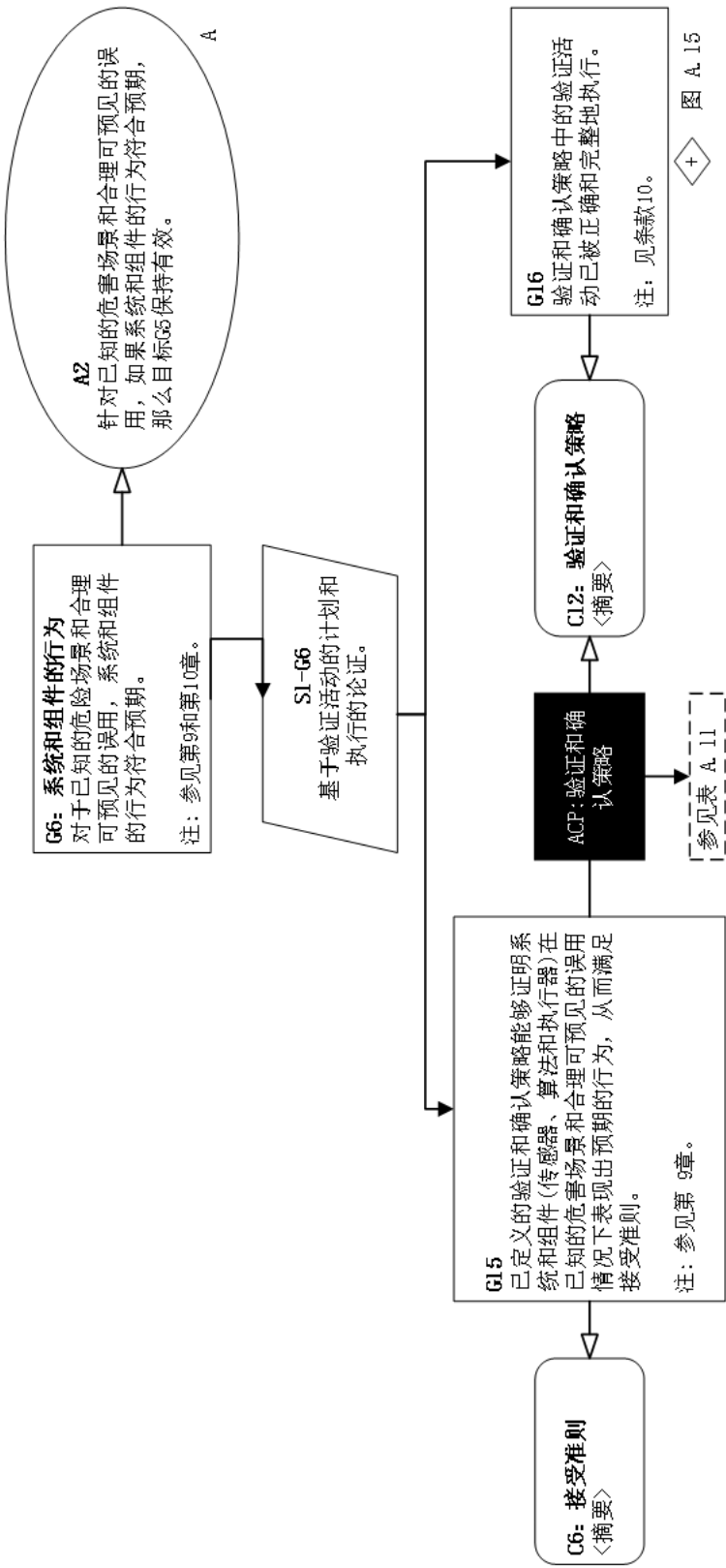
ODD定义的完整性和正确性
中间层决策逻辑描述的完整性和正确性
车辆和实现预期功能的要素（可包括系统、子系统和组件等）的描述的完整性和正确性
该功能在车辆动力学上的权限和驾驶自动化等级描述细节的完整性和正确性
性能目标的适当性
合理可预见的误用场景描述的完整性和正确性
接口和交互描述的完整性和正确性
假设的完整性和有效性
系统和子系统的局限性及其对策描述的完整性和正确性
系统架构（用来支持对策）描述的完整性和正确性
报警和降级概念描述的完整性和正确性
支持预期功能的数据收集信息描述的完整性和正确性
性能目标描述的完整性和正确性
已知潜在性能局限及其对策描述的完整性和正确性
迭代过程（用来保持规范的更新）有效性描述的完整性和正确性
分布式开发管理过程有效性描述的完整性和正确性
系统限制描述的完整性和正确性
鲁棒性（由最终系统架构提供）描述的完整性和正确性
评审（根据第12章）SOTIF活动产生的证据，以发现SOTIF实现过程中潜在问题的能力

表A. 9 G12 开发相关的主题

采取“避免”措施
采取“减少”措施
采取“减缓”措施
使用系统修改以避免或降低预期功能安全相关的风险
使用措施以限制预期功能
使用措施将驾驶权从系统转交给驾驶员
使用措施来减少或减缓合理可预见误用的影响
用于根据修改进行系统规范更新的流程的充分性

表A. 10 G14 开发相关的主题

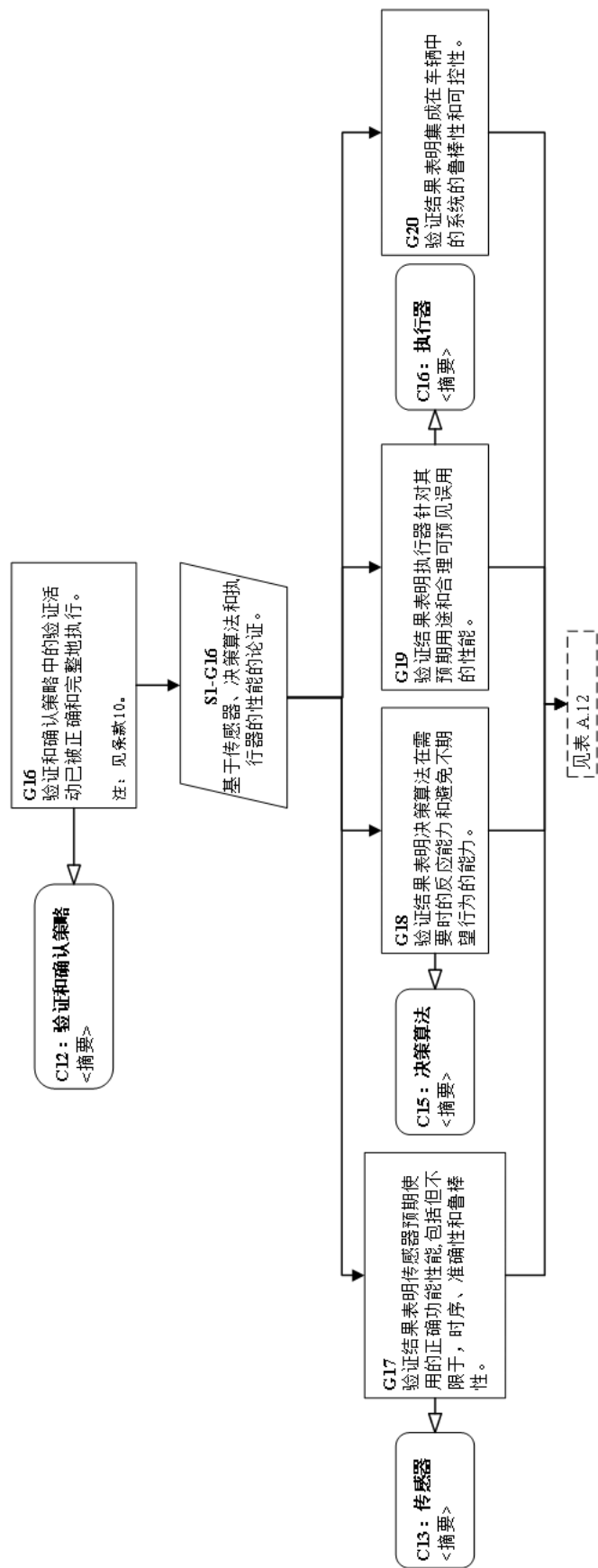
使用专家判断
残余风险与6. 5中规定的接受准则的比较
缺少已知的场景可能导致特定车辆出现不合理风险的情况



图A. 14 G6: 系统&组件行为

表A. 11 ACP 相关主题：验证和确认策略（验证和确认策略已被正确定义）

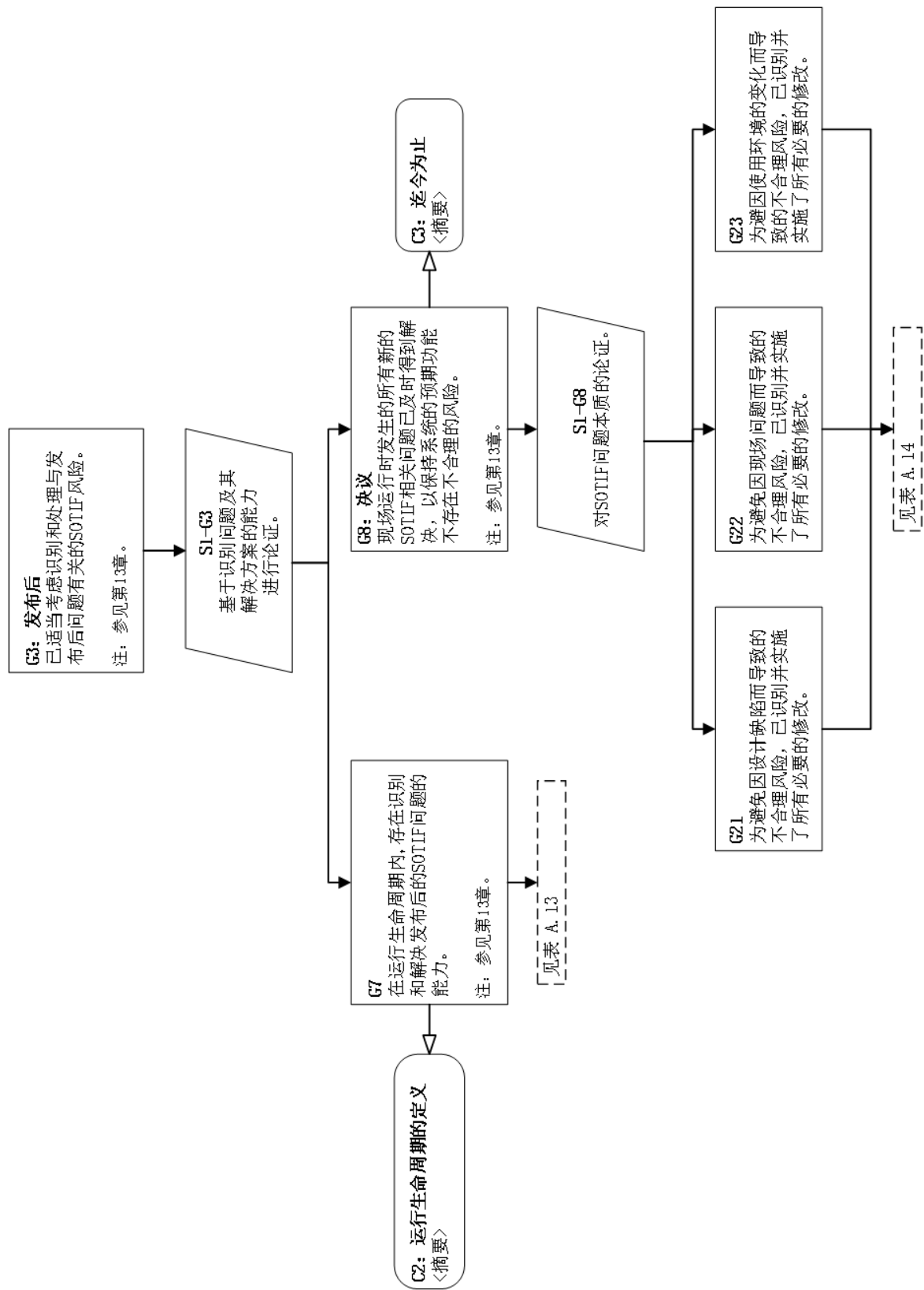
已知场景的覆盖率
场景子集的暴露度
使用暴露度、可控性及严重度来评估危害行为场景
用于指定验证和确认活动的方法的基本原理(表6)
策略用来验证传感器提供准确环境信息的能力的有效性
策略用来验证传感器处理算法对环境精确建模的能力的有效性
策略用来验证决策算法安全处理元素技术局限的能力的有效性
策略用来验证决策算法根据环境模型和系统架构做出适当决策的能力的有效性
策略用来验证系统或功能鲁棒性的有效性
策略用来验证不存在由于预期功能的危害行为而导致的不合理风险的有效性
策略用来验证人机交互界面（HMI）避免合理可预见误用的能力的有效性
策略用来验证后援移交场景有效性的有效性
所选方法的基本原理(表7、表8、表9、表10)
所选方法的充分性(表7、表8、表9、表10)
评审(根据第12章)SOTIF活动产生的证据，以发现SOTIF实现过程中潜在问题的能力



图A. 15 G16

表A. 12 G17, G18, G19, G20 开发相关的主题

车辆设计 (例如安装位置)
已知场景的覆盖率
接受准则的符合度
触发条件的覆盖率
所选方法的基本原理 (表7、表8、表9、表10)
所选方法的充分性 (表7、表8、表9、表10)
方法的定义 (表7、表8、表9、表10)
方法部署所耗费资源 (表7、表8、表9、表10)
评审 (根据第12章) SOTIF活动产生的证据，以发现SOTIF实现过程中潜在问题的能力



图A. 16 G3: 发布后

表A. 13 G7 开发相关的主题

车端和场外基础设施是否足以监控在用中的功能不足
识别和响应系统潜在不足的能力
识别和纠正设计不足的能力
识别和响应运行变更的能力
收集现场数据的能力
监控预期功能安全相关问题的能力，包括系统的误用
使用现场数据识别问题的能力
监控当前认知水准的能力
监控使用环境更改的能力
分析和评估已识别风险的能力
减轻已识别风险的能力

表A. 14 G21, G22, G23 开发相关的主题

识别和应对系统的潜在不足
识别和纠正设计不足
识别和响应运行变更
利用现场监控数据收集来加强用于预期功能安全活动的数据库
监控与预期功能安全相关的问题，包括系统的误用
使用现场监控以识别潜在的不足
对认知水平的监测，以识别潜在的不足
监控对使用环境的更改以识别潜在的不足
对识别出的风险进行分析和评估
风险的减轻

A. 2 GB/T 34590 标准功能安全和本文件之间相互作用的说明

A. 2. 1 总则

按照GB/T 34590. 3-XXXX第6章的要求，应用危害与可操作性分析（HAZOP）方法识别驱动电机系统的功能异常表现，见表A. 1。

为简单起见，对于所讨论的活动或工作成果，并未论述全部的方面，因而，本子章节不具有完整性。

A. 2. 2 GB/T 34590标准与本文件的范围对比

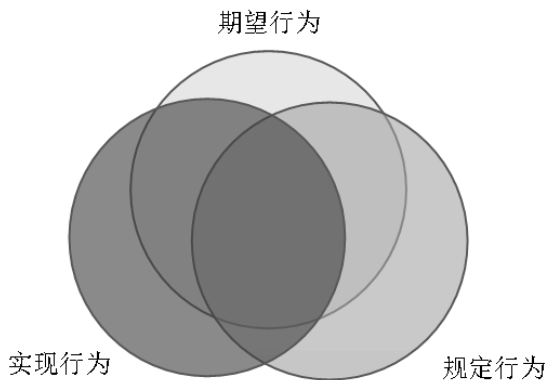
A. 2. 2. 1 总则

可以通过两种不同的方法来进一步解释两个标准之间的差异和共性：

- 三圈行为模型；
- 安全问题的因果关系分类视图。

A. 2. 2. 2 三圈行为模型

图A. 13中的三圈行为模型，详细说明了GB/T 34590标准范围和本文件范围的差异和重叠的部分[15]。



图A. 17 三圈行为模型

注1：这三个圆圈明显缺乏足够的重叠程度，仅是用于举例说明，并不代表真实情形。

在图A. 17中每个圆圈代表行为的一个不同方面。

——从不考虑任何技术限制的安全角度来看，期望行为是理想的（有时是想要致力达成的）行为。其反映了用户和社会对系统行为的期望。

示例1：从不发生或导致事故的自动驾驶功能。

示例2：自动紧急制动（AEB）系统的期望行为应该是 100%的真阳性制动和 0%的假阳性动。

注2：这三个圆圈明显缺乏足够的重叠程度，仅是用于举例说明，并不代表真实情形。

——规定行为是考虑了约束（如法律、技术、商业、客户接受度）的期望行为。

注3：根据第3章，预期功能被定义为规定功能。因此，预期行为作为预期功能的行为，是规定行为的同义词。

——实现行为是真实世界的系统行为。

比较GB/T 34590标准和本文件的范围，可得出以下结论。

——GB/T 34590 标准明确地针对电气/电子随机硬件故障问题。本文件并不明确针对这种问题。

然而，对随机硬件故障的反应，即紧急运行，可能涉及预期功能安全。

——确保实现行为符合规定是 GB/T 34590 标准的任务之一；而针对某些复杂系统（如 ADAS、AD 系统）的上述问题，则是本文件的任务之一。对于这些系统，GB/T 34590 标准未能提供足够的指导来确保这一点。该问题与开放性环境有关，即真实世界无法 100%准确地被描述，或对其的正确感知无法被 100%的确认。采用复杂的算法和传感器（如摄像头、雷达或激光雷达）来感知环境并进行分类，并从该信息中推导出控制行动的系统，在本文件考虑的范围中。

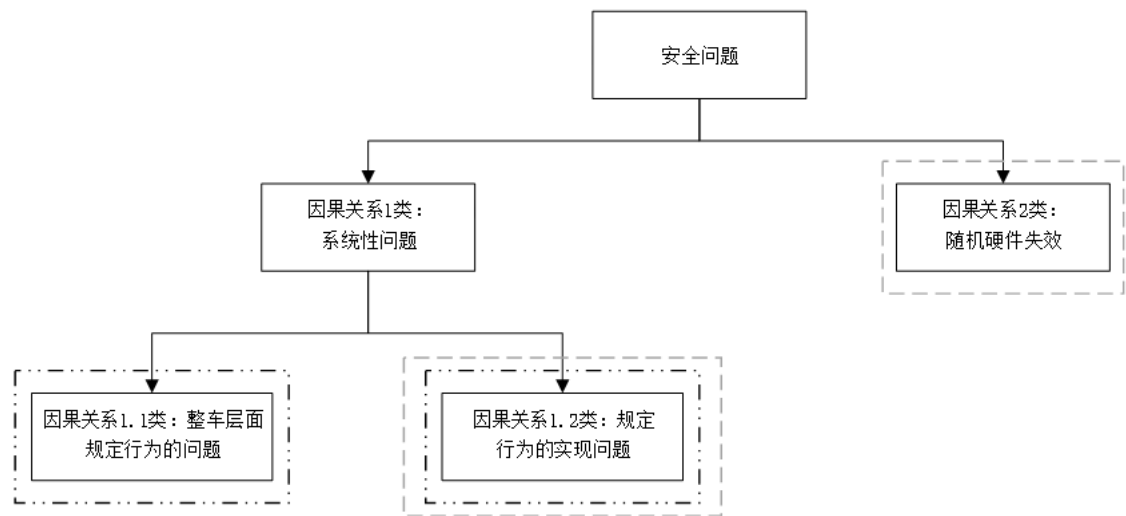
在图A. 17中每个圆圈代表行为的一个不同方面。

示例3：某个基于摄像头的系统具有行人检测功能。当行人身穿特定颜色图案的服装时，该算法可能会出现错误地分类行人的问题。对所有可能的服装颜色图案进行定义和测试是不可能的。本文件旨在描述 GB/T 34590 标准的附加要求。与 SOTIF 相关的电气/电子要素，在 GB/T 34590 标准中也被视为安全相关要素。

示例4：如果某个通过软件实现的物体检测算法可导致安全目标的违反或实现，则在 GB/T 34590.1 中，该算法被视为安全相关的要素。

A. 2. 2. 3 安全问题的因果关系分类视图

在安全问题的因果关系分类视图中，GB/T 34590标准和本文件范围的差异和共性之处如图A. 18所示。



□ 本文件重点关注的方面

□ GB/T 34590标准重点关注的方面

图 A.18 安全问题因果关系分类方案

注1：此分类方法仅关注GB/T 34590标准和本文件所述的电气/电子系统引起的安全问题。为了简单起见，该方法省略了其他安全问题（如电气危害）。

该方案包括以下分类：

a) 因果关系 1 类：系统性问题

该类别包含可能与系统性问题相关的安全方面。该类别可进一步分为：

——因果关系 1.1 类：整车层面规定行为的问题；

——因果关系 1.2 类：规定行为的实现问题。

b) 因果关系 2 类：随机硬件故障

该类别包含GB/T 34590标准描述的由随机硬件故障引起的安全问题。

c) 因果关系 1.1 类：整车层面规定行为的问题

该类别包含整车层面上规定行为引起的安全问题。本文件针对整车层面功能的规定行为引起的风险，对于这些功能，恰当的态势感知对安全至关重要。态势感知基于复杂的传感器和处理算法（如通过摄像头、激光雷达或雷达探测目标物）。在本文件中，造成这类问题的原因被认为是整车层面的设计规范定义不足。

注2：GB/T 34590标准中明确排除了标称行为的安全相关内容。

d) 因果关系 1.2 类：规定行为的实现问题

此类问题由性能局限、要素层面定义不足等其他各种设计和实现问题造成的。

这三种因果关系中，1.2类系统性问题属于GB/T 34590标准范围，因为此类问题与电气/电子系统、子系统、组件或其他要素的潜在系统性失效有关，也和预期功能安全要求相关的问题有关。

在要素层面，只有那些合适的态势感知对安全至关重要的预期功能的性能局限和规范定义不足包含在本文件的范围之内。在要素层面范围内的功能包括：

——感知：对环境的感知[例如：利用车辆内部和外部（如 V2X）数据探测周围静态和动态目标物、街道布局 and 自车位置]；

——规控：决策算法（即：根据前述感知结果得出控制动作的控制算法）；及

——执行：执行（即：执行上述决策算法得出的控制请求）。

注3：如果某个特定安全问题不能明确归类为预期功能安全问题或是功能安全问题，则可以同时应用两个标准来解决该问题。

A. 2. 3 本文件与GB/T 34590系列活动的协同

本文件与GB/T 34590标准产品开发活动的结合如图A. 19所示。由于两个标准处理不同安全方面，因此这两个过程都被视为产品的可靠安全论证。标准开发活动之间的结合对于在早期阶段对整车的设计以及包括系统、子系统、部件等要素进行可能的修改至关重要。

在开发过程初期，规范定义和设计（根据第5章）可以与GB/T 34590. 3（见A. 2. 4）的相关项定义协同开展。

注：第5章包含各抽象层级的功能和规范定义。这与相关项定义不同，后者定义顶层的功能。

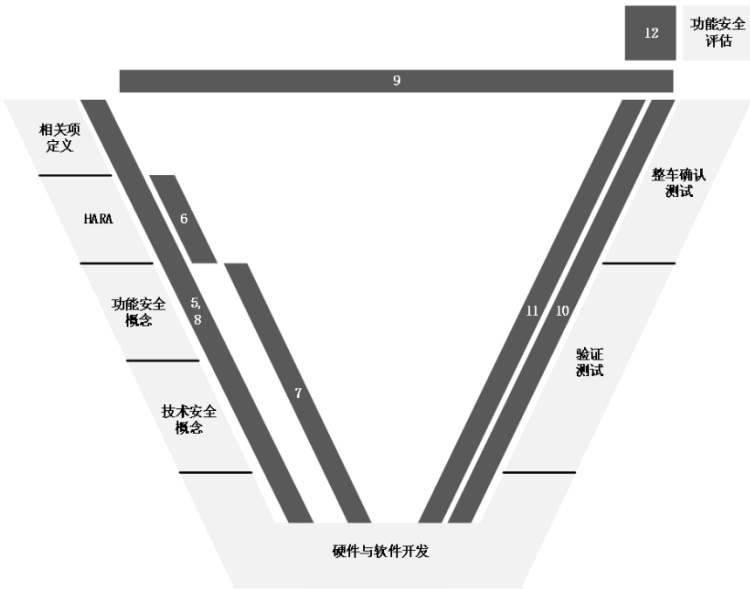
对预期功能引起的危害的识别和评估可以与GB/T 34590. 3（见A. 2. 5）的危害分析和风险评估（HARA）协同开展。对性能局限和潜在触发条件的识别与评估，需要考虑系统局限，并根据SOTIF评估其可接受性（见A. 2. 7）。这一阶段可与GB/T 34590标准（见A. 2. 6和A. 2. 7）的功能安全概念和技术安全概念协同开展。

用于降低SOTIF风险的功能修改（根据第8章）可以与GB/T 34590 V模型左侧的开发活动协同开展。

在评估硬件（HW）和软件（SW）组件层面的性能局限和潜在触发条件时，可与GB/T 34590标准的硬件和软件开发活动协同开展。4. 4. 2和4. 4. 3分别提供了分布式的SOTIF开发和SEooC流程方面的指导。本文件的A. 2. 9中描述了GB/T 34590. 8中的过程支持的主题。

SOTIF的验证和确认可与GB/T 34590标准V模型右侧的活动协同开展（见A. 2. 10）。SOTIF验证和确认策略是根据SOTIF开发前期阶段产生的信息定义的。

SOTIF成果评估和功能安全评估对开发活动进行了总结，并用于整个系统发布。现场监控可以与GB/T 34590. 7中要求的现场监控流程协同开展。



说明：

流程步骤

X GB/T XXXX 的流程步骤，X 表示相应章节

图 A. 19 本文件和 GB/T 34590 标准产品开发活动的可能交互

A. 2. 4 整车层面相关项定义和功能定义

本文件的起点为整车层面功能规范。GB/T 34590标准的起点则为相关项定义。

注1：相关项是实现整车层面功能或部分功能的系统或系统组。给定的车辆功能可能由多个相关项实现。在这种情况下，车辆功能和单个相关项本身的功能有所不同。

注2：相关项可以与多个车辆功能的实现相关，从而多个车辆功能（或其子集）的定义可以作为相关项定义的一部分。

注3：本文件阐述的在整车层面上定义的功能与GB/T 34590标准中一个或多个相关项实现的车辆功能相同。

示例1：本示例中的车辆功能“自动紧急制动（AEB）”通过雷达传感器、域控制器和制动系统[如电子稳定性控制系统（ESC）]实现（见图 A. 20）。



图 A. 20 统架构示例

GB/T 34590标准允许以不同的方式定义相关项。例如，车辆功能可以通过两个相关项实现（雷达传感器和域控制器作为一个相关项，ESC作为另一个相关项），或者整车功能可以通过一个相关项实现（雷达传感器、域控制器和ESC）。

在A. 2. 4至A. 2. 10中，相关项被定义为实现完整的整车功能，即相关项功能等同于整车功能。

注4：为了简单起见，本示例忽略了相关项实现的其他功能。

示例2：整车层面 AEB 功能定义：AEB 功能触发最大制动力：

- 一旦发现障碍物且碰撞无法避免（这意味着碰撞无法被阻止，但可以降低碰撞的严重度）；
- 最大速度降为 x km/h。

注5：为改进SOTIF而进行的更改（如功能修改、新要素的引入），也会对相关项定义产生影响。

A. 2. 5 HARA与针对预期功能引发的危害的识别和评估

A. 2. 5. 1 总则

GB/T 34590标准关注电气/电子功能，在危害分析和风险评估（HARA）中，分析功能异常行为在整车层面产生的危害。在整车层面，不论是电气/电子失效，还是不安全的预期功能（或者甚至是信息安全问题），导致危害的行为是相同的。然而，考虑到预期功能的权限限制（例如AEB的最大减速度限制），产生的危害幅度可能是不同的。在HARA中识别出的危害和功能异常行为，可能与SOTIF所考虑的相同或相似。

A. 2. 5. 2 GB/T 34590. 3 危害分析和风险评估（HARA）

HARA识别相关项的功能异常行为，并评估由此产生的风险。

示例1：AEB 相关项的功能异常行为：

- 非期望的自动制动：
 - 在规定的速度降限制范围内：ASIL X为对危害事件进行E、C和S评估的结果；

- 超出规定的速度降限制范围：ASIL Y是对危害事件进行E、C和S评估的结果（ $Y \geq X$ ）；

——自动制动过晚或丢失：

- 由于高可控性（制动是驾驶员的常规行为）和低暴露度（紧急制动是小概率事件），可将危害事件评定为QM。

注：（关于上述示例1）在其他具有更高驾驶自动化等级的系统中，系统可能承担了总体的制动任务，而不仅仅是紧急制动。在这种情况下，上述阐述可能不再有效。

SOTIF引发的功能修改可能对HARA中的参数产生影响。

示例2：AEB 功能在自动制动时限制了最大速度降，这增加了后方车辆的可控性，避免追尾，降低碰撞的严重程度。

A. 2. 5. 3 识别和评估由预期功能引起的危害

本活动从以下方面评估车辆功能：

——车辆功能的规定行为是否安全？

——车辆功能的非期望行为有哪些，其是否是可信伤害的来源？

——合理可预见的误用会带来什么风险？

示例：AEB 的风险识别和评估：

a) 在定义的用例中，整车层面的规定行为是否安全？

如果规定行为可能是事故的原因，评估给定环境中是否有更合适的行为。

根据AEB系统的规范，只有当碰撞不可避免时，才会进行干预。在这种情况下，驾驶员可以以最大能力进行制动。如果驾驶员不这样做，AEB系统将代替驾驶员进行制动。除非驾驶员想通过横向躲避来防止事故发生，这是最好的可能行为。在后一种情况下，制动甚至可能会适得其反，降低可用的横向加速度。因此，修改了整车层面的规定行为：当转向扭矩为 y Nm时，AEB干预被抑制或中止。通过这一修改，整车层面的规定行为被认为是安全的。

为了简单起见，本示例中省略了对该新增内容的进一步评估。

b) 车辆功能的非期望行为有哪些？这些行为是否是可信伤害的来源？

——误触发：在规定的速度降限制范围内非预期制动。

- 后方车辆没能及时作出反应，导致追尾。在这里，该系统引入了新的风险。这种非预期行为是可信伤害的来源，且与预期功能安全（SOTIF）相关。

——漏触发：在即将发生碰撞时不制动。

- 该系统仅起辅助作用，也就是系统既没有把驾驶员从指定任务中解放出来，也不会给驾驶员一种从制动任务中解放出来的印象，因为除非事故已经不可避免，否则系统不会制动。从预期功能安全（SOTIF）的角度来看，这种非预期行为不会给系统引入新的风险，也不会被视为可信伤害的来源。因此，这种非预期行为与预期功能安全无关。
- 在其他系统中，该系统可能会接管制动驾驶任务。在这种情况下，上述说法不再有效，这种非期望行为则与预期功能安全有关。

——在规定的速度降限制范围之外制动。

- 在规定的速度降限制范围内进行制动的能力，取决于车速测量的准确性和制动器的执行情况。
- 导致超出速度降限制范围的制动的与环境有关的潜在触发条件（如，来自前方的阵风，上坡梯度增加）是可能存在的，但这也基于相关项的控制回路可以快速的适应它们，并将过制动控制在较小的范围内。
- 如果已经建立的系统很好地解决了车速测量、制动控制回路和制动执行的性能局限。则不需要本文件中描述的预期功能安全（SOTIF）流程。这种非期望行为与本文件无关。

c) 合理可预见的误用会带来什么风险？

——误用场景：驾驶员将“避免与目标物碰撞的制动”任务转移至 AEB 系统。

- 在用户手册中明确提到，该系统仅辅助驾驶员，并不能防止碰撞，只是削弱了影响。
- 该系统以令人极不舒适的方式进行制动。

因此，驾驶员将制动驾驶任务全权转移给该系统的风险不是不合理的。

一般来说，可通过告知驾驶员该系统的局限性（例如通过用户手册）来降低误用的可能性。

请注意，包括广告和产品命名在内的销售材料不应导致用户的错误期望。

A.2.5.4 结论

应小心谨慎，确保由预期功能引起的危害识别和评估结果与危害分析和风险评估(HARA)结果一致。在A.2.5中使用的示例中，功能异常行为/非预期行为“非预期制动”和“在即将发生碰撞时不制动”就是这种情况。由预期功能引起的危害识别和评估发现的非预期行为，与危害分析和风险评估(HARA)所发现的错误功能行为，可能导致相同的危害。

针对预期功能引起危害的识别和评估，与危害分析和风险评估(HARA)不总是必然重合。考虑规范行为安全性的评估是典型的预期功能安全(SOTIF)话题。

在评估由相关项功能异常行为所引起的危险事件时，GB/T 34590 HARA只考虑了合理可预见的间接误用，并将其视作可能导致可控性降低或严重程度增加的原因。

在评估该系统危害行为引起的危害事件时，本文件同样考虑了合理可预见的间接误用。然而，本文件还考虑了合理可预见的直接误用，并视其为潜在触发条件。

这些活动的某些方面，例如可控性评估，既可以被视为预期功能安全(SOTIF)话题，也可以被视为功能安全话题。

A.2.6 功能安全概念和预期功能安全(SOTIF)功能规范

功能安全概念规定了故障响应(如紧急运行、过渡到安全状态等)。对于高级驾驶辅助系统(ADAS)和自动驾驶系统而言，这种故障响应也可能是一个预期功能安全(SOTIF)问题。对于这些系统，预期功能安全(SOTIF)定义必要的功能，从而以安全的方式执行规定的故障响应。功能安全的任务是在发生故障时确保定义的必要功能可以使用(例如通过故障容错)，或者确保故障发生的概率足够小(例如通过故障预防)。

定义安全故障响应其本身既可以被视为预期功能安全(SOTIF)任务，也可以被视为功能安全任务。

示例：在使用自动驾驶功能的情况下，故障响应可为如下示例：

——在当前车道上安全停车，

——开往下一个停车场。

注：通过适当的信息交换和(或)评审，可保持第8章规定的功能修改与从GB/T 34590中功能安全概念相关分解的要求一致。

A.2.7 技术安全概念和预期功能安全(SOTIF)

由于预期功能安全(SOTIF)活动，系统设计可能会发生变化(例如，通过引入新的传感器)，这可能会影响技术安全概念。

此外，由于功能安全活动，系统设计也可能会发生变化(例如，通过引入新的传感器)，这可能会影响预期功能安全(SOTIF)。

A.2.8 安全分析

确保功能安全和预期功能安全(SOTIF)的分析活动，侧重于功能链，并使用相同的设计作为起点，但具备不同的视角。功能安全分析解决了实现电气/电子要素规定行为过程中引入的系统性问题，以及电气/电子要素的随机硬件故障。

预期功能安全(SOTIF)分析(第7章)侧重于功能不足、其潜在的触发条件及其对车辆行为的影响。此外，分析还考虑了合理可预见的间接误用(第6章、第7章)。

GB/T 34590标准的安全分析结果可作为预期功能安全（SOTIF）分析的输入，反之亦然。

对整车层面的规定行为的安全性以及合理可预见的误用导致的风险的分析是预期功能安全(SOTIF)特有的。

A. 2. 9 支持过程

本文件未明确阐述有关开发流程本身的要求。开发流程是否适用对于实现安全性至关重要，现有标准（如IATF 16949和GB/T 34590标准）对此进行了说明。例如，假设有必要对GB/T 34590.8的支持过程进行调整，并将其应用于支持预期功能安全（SOTIF）的实现，如：

- 可对根据 GB/T 34590.8-XXXX 第 5 章规定实施的开发接口协议（DIA）加以详细说明，以解决预期功能安全（SOTIF）各方面的问题（见 4.4.2）；
- 可以将根据 GB/T 34590.8：XXXX 第 11 章规定使用软件工具的信心应用到相关的工具中，通过一些调整来实现支持预期功能安全（SOTIF）。

注1：除了明显的工具错误之外，仿真工具在一定偏差范围内表示真实世界的能力，可能与预期功能安全（SOTIF）具有特定相关性。

注2：对真实世界数据的测量准确性本身，可能与预期功能安全（SOTIF）具有特定相关性。

A. 2. 10 验证和确认

针对预期功能安全（SOTIF）相关要求的验证和确认策略（见第9章）以及具体测试用例（第10章和第11章）也可以考虑功能安全要求。

某些测试用例可以针对预期功能安全和功能安全问题，而某些测试用例仅针对功能安全方面（例如，安全机制探测和指示随机硬件故障的能力）或仅针对预期功能安全方面（例如，评估整车层面规定行为充分性的测试）。

A. 3 简化的预期功能安全（SOTIF）应用示例

表A.15通过考虑与设计运行范围相关的SOTIF危害及其减缓，给出了具备递增驾驶自动化等级的不同功能间的对比简化示例。

表A.15 设计运行范围相关的预期功能安全（SOTIF）危害与减缓措施的简化示例

系统示例	驾驶员辅助（L1-根据第3章表2）	部分驾驶自动化（L2-根据第3章表2）	有条件驾驶自动化（L3-根据第3章表2）	条件驾驶自动化（L3-根据第3章表2）	高度驾驶自动化（L4-根据第3章表2）
	自适应巡航控制	结合车道保持的自适应巡航控制	交通拥堵辅助	高速路辅助驾驶	自动驾驶出租车
系统描述	该功能通过传感器检测前车来增强标准汽车巡航控制。如果距离前车太近，则该功能将采取减速措施，以匹配前车的速度。	该功能使用传感器将车辆保持在车道中心的位置，并检测前车以调整车速，从而保持预设的车间时距。	在公路交通堵塞时，该功能使用传感器使车与前车保持安全的纵向距离。功能包括转向，以便保持在行驶车道上。	该功能使用多个不同的传感器，在交通中自主导航，执行高速公路驾驶的所有必要操作。	该功能使用多个不同的传感器，在定义的地理区域内自动驾驶从A点到B点的交通。

系统示例	驾驶员辅助（L1—根据第3章表2）	部分驾驶自动化（L2—根据第3章表2）	有条件驾驶自动化（L3—根据第3章表2）	条件驾驶自动化（L3—根据第3章表2）	高度驾驶自动化（L4—根据第3章表2）
	自适应巡航控制	结合车道保持的自适应巡航控制	交通拥堵辅助	高速路辅助驾驶	自动驾驶出租车
动态驾驶任务—横向和纵向车辆运动控制	驾驶员和系统	系统	系统	系统	系统
动态驾驶任务（DDT）—目标和事件探测与响应（OEDR）	驾驶员	驾驶员	系统	系统	系统
动态驾驶任务—后援	驾驶员	驾驶员	后援用户 ^a	后援用户 ^a	系统
运行用例	1) 保持跟车时距，引导车辆达到设定速度	1) 在车道上跟随前车，达到设定速度和跟车时距	1) 跟随前车，其行驶速度等于或低于x公里/小时，距离不超过y米	所有与高速公路相关的用例（跟随、车道保持、并道、超车等）	所有与城市和高速公路相关的用例（跟随、超车、并道、交通控制停车等）
	2) 当自主车辆的前方没有前车时，保持所需速度	2) 当自车的前方没有前车时，保持所需速度并沿车道行驶	2) 如果前车改变车道，则继续跟随下一辆紧随的前车；或者，如果没有前车，则要求驾驶员接管车辆控制权		
设计运行范围	当车辆以x公里/小时或以上的速度行驶时，该系统可运行。	当车辆处于可检测车道且以x公里/小时或以上的速度行驶时，该系统可运行。	当车辆在地理区域（地图区域）内，在有效车道上，并且在大多数环境条件下以低于x公里/小时的速度行驶时，该系统可运行（假设在恶劣环境条件下，如浓雾、大雨等，该功能会断开）。	系统在大多数环境条件下，在地图可用的高速公路上可运行（假设在恶劣环境条件下，如浓雾、大雨等，该功能会断开）。	系统在除极端天气（如规范中所定义）以外的所有环境条件下，在地理围栏内的高速公路和城市区域中可运行。
预期行为/功能示例	与前车保持安全的跟车时距。如果距离前车太近，则该功能将施加适当的制动力以保持安全的跟车时距。如果检测到距离前车较远，则该功能将应用加速，直到达到用户的预设速度。	保持车道边界，并与前车保持安全的跟车时距。如果距离前车太近，则该功能将施加适当的制动力以保持安全的跟车时距。如果检测到距离前车较远，则该功能将应用加速，直到达到用户的预设速度，并应用横向控制保持车道。	该系统邀请用户在如浓雾等恶劣环境条件下接管车辆，并在车辆退出设计运行范围前取得车辆控制权。	执行并线横向操作，同时为他人留出适当的时间和空间。	在遮挡区域谨慎行驶。

系统示例	驾驶员辅助（L1-根据第3章表2）	部分驾驶自动化（L2-根据第3章表2）	有条件驾驶自动化（L3-根据第3章表2）	条件驾驶自动化（L3-根据第3章表2）	高度驾驶自动化（L4-根据第3章表2）
	自适应巡航控制	结合车道保持的自适应巡航控制	交通拥堵辅助	高速路辅助驾驶	自动驾驶出租车
需要减缓的预期功能安全（SOTIF）危害示例	当接近桥梁时，系统因错误地将桥梁感知为道路中的静态金属物体而制动。	自车和前车在合流车道上运行。前车汇入预定车道，当前自车不再检测到前车，因此开始加速至先前预设的巡航控制速度。自车驾驶员无法在合并车道结束并驶离道路之前汇入预定车道。	当该系统进入大雾区无法以合格精度感知目标物而发出接管请求时，由于接管用户没有观察到视觉警报，因此接管用户没有接管车辆控制。	车辆未能成功合流，因为无法检测到具有灯光和颜色的车辆，导致自动驾驶系统错误地将车辆归类为标称天际线。	相邻车道上的大型车辆遮挡了交通信号灯，导致自动驾驶出租车感知不到交通信号灯，在红灯时进入了交叉路口。
预期功能安全（SOTIF）减缓措施示例	加强软件算法，以区分车辆和道路基础设施（即钢桥、钢制外罩）。	该功能限制加速权限。	车辆的设计能够检测到阻碍道路的雾情，并向后援用户提供视觉警报。如果后援用户无法接管控制，则该系统采用其他方法，通过刺激驾驶员的其他感官，如听觉，触觉，运动感知（如短制动脉冲）来通知驾驶员。	单独评估原始传感器数据的正交且独立的碰撞减缓算法，验证生成的路径在被较低级别控制器接受之前不会发生碰撞。	车辆通过感知数据使地图数据合理化，以便在进入交叉路口之前寻找交通信号灯状态，并理解出现的大型车辆遮挡了交通信号灯。选择适当的行为方式。
* 驾驶员和后援用户之间的区别在于驾驶员需要持续监督。而后援用户可能不会监督 OEDR，但需要在适当的时间内根据请求进行控制。					

就验证和确认而言，无论驾驶自动化等级如何，都有很多共性。

针对已知的潜在危害场景，SOTIF缓解措施的评估：

- a) 分析工作以揭示新的潜在触发条件；
- b) 在演示缓解措施的已知场景下执行该功能。

这可以通过使用子系统和系统在封闭路线、仿真环境和开放道路的组合测试来实现。

针对未知的潜在危害场景，SOTIF缓解措施的评估：

- a) 影响 V&V 策略的分析工作以揭示未发现的潜在触发条件；
- b) 在封闭路线、仿真环境和开放道路中对 ODD 进行持续暴露以实现确认目标，从而证明未知潜在危害场景的残余风险是可接受的。

在扩大ODD时（例如将功能导出到其他城市或国家），需要识别和评估ODD和OEDR的变化。这可能导致需要重复进行测试和仿真活动。

A. 4 规范定义和设计的简化示例

自动驾驶功能规范定义和设计是预期功能安全开发的输入，通常包含大量信息，并且在开发过程中不断得到补充和细化。本附录以导航辅助驾驶NOA（Navigate on Autopilot）功能为例，给出了一个经

过简化的、非完整的功能规范定义和设计示例，用以示范如何在项目开发的前期，快速生成可用于支持预期功能安全开发的规范定义和设计。

注：本附录目的是提供一种规范定义和设计的初始编写方式和方法的参考，用以支持规范定义和设计的首次建立或对相关内容的总结性呈现，因而本附录内容未涵盖规范定义和设计的全部条目，给出的条目也不具备完备性，相关条目中给出的数据仅为了便于理解，非推荐值。

表 A. 16 用于支持预期功能安全开发的规范定义和设计示例

1. 功能描述	
1.1 功能名称	导航辅助驾驶NOA（Navigate on Autopilot）功能
1.2 功能描述	导航辅助驾驶NOA（Navigate on Autopilot）功能，能够基于驾驶员选择的地图起点和终点导航输入，实现高速/高架道路上车辆自动巡航控制，包括：车道/匝道居中行驶、匝道自动汇入/汇出、自主换道、自主超车等子功能。在此过程中，驾驶员须全程关注车辆运行情况，在必要时（车辆可能有/无提醒）随时接管车辆。
注1：为了简化，以下挑选NOA功能中匝道自动汇入（主道）子功能场景为例进行论述。	
1.3 功能开启条件	NOA功能已开启； 驾驶员已授权自动运行匝道汇入； 导航和路径规划需要匝道汇入。
1.4 功能主场景	<p>如图1所示，车辆以匝道允许的最高车速（例如：60km/h）进入平行加速车道，车辆通过传感器（例如：毫米波雷达、摄像头、GPS等）探测加速车道及主路行车道的情况，在确保足够安全的情况下自动控制转向变道，完成车辆汇入主路行车道。</p>  <p>图1：匝道自动汇入场景示意图</p>
1.5 其他常见功能场景	当相邻车道（行车道或具有2条加速车道的内侧车道）有障碍物车辆时，自车需要通过加速或减速实现躲避； 当加速车道有障碍物时，自车需要通过制动进行避障。
1.6 功能接管	若相邻车道（行车道或具有2条加速车道的内侧车道）一直存在障碍物车辆或其他原因（如：大雨等天气原因）导致无法自动汇入时，车辆会通过声音和光电报警提醒驾驶员接管，如果驾驶员一直不接管，车辆会在达到加速车道终点渐变段前完成主动停车。
2. 设计运行范围ODD	
2.1 空间	高速公路入口匝道，主路宽3.75米，平行加速车道有效长度 ≥ 230 米，路宽3.5米，道路水平，附着良好（附着系数0.9）
2.2 交通	匝道限速60km/h，主路可能有行驶的车辆，车速60—120km/h，车长 ≤ 13 米
2.2 时间和天气	能见度良好、天气良好情况下全天可用
2.4 其他	若有

注2：对于ODD的描述，还可以参考场景描述的方法，例如本文件附录B等。

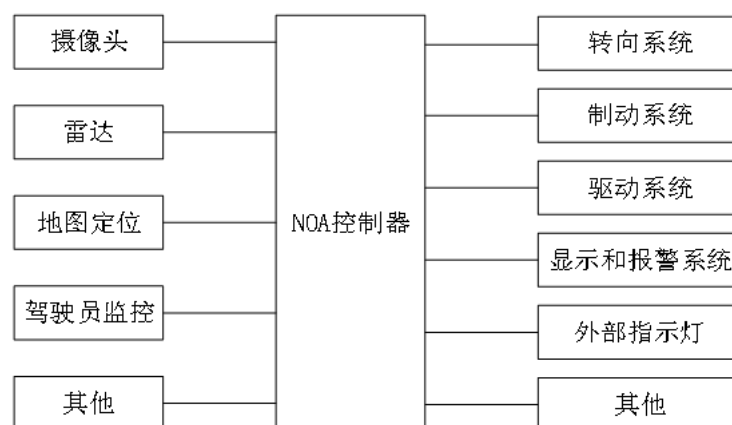
注3：ODD是设计时确定的驾驶自动化功能的运行条件，但场景和使用案例考虑是执行驾驶任务过程中，一定连续时间内车辆和外部的交互活动，这里的外部包括：环境、道路、其他交通参与者等。可见，ODD考虑的是设计上已知的范围，而场景和使用案例还要考虑未知因素。

3. 系统设计

3.1 系统组成

3.1.1 感知	摄像头、雷达（毫米波雷达和/或激光雷达）、地图及定位、驾驶员监控传感器（摄像头和/或手握方向盘检测）等。
3.1.2 决策	NOA控制器。
3.1.3 执行	转向系统、制动系统、驱动系统、HMI显示和报警系统、转向灯、制动灯等。

3.2 系统架构图



4. 性能目标

4.1 感知方面	摄像头探测距离150米（夜晚取决于前照明条件）； 雷达探测距离：前向150米，角雷达80米； 地图和GPS定位准确性1米； 等。
4.2 决策和执行方面	无障碍物情况下，自车主动完成汇入（打开转向灯至车辆中心行驶到主车道中心线）时间6秒（提前3秒打开转向灯提示）； 车辆汇入过程中，侧向运动加速度最高0.15g； 为躲避相邻车道障碍车辆时，自车主动加速能力最高0.3g，减速能力最高0.3g； 紧急避障减速能力0.9g（若附着条件允许）。

5. 已知功能或性能不充分及对应的触发条件

5.1 感知方面	感知不足1：角雷达感知范围有限，汇入前可能无法准确探测到左后方来车或探测到的时间过晚，导致自车与主路后方来车有碰撞风险； 感知不足1的触发条件：匝道曲率较大、加速车道较短、隔离带遮挡等。
5.2 其他方面	若有

6. 整车层面的安全策略&报警和降级概念

6.1 安全策略	安全策略1：自车打开转向灯到开始转向的时间间隔 $\geq 3s$ ，以充分提醒主车道车辆； 安全策略2：自车与相邻车辆的间距 $TTC \geq 0.5s$ ；
----------	--------------------------------------------------------------------------------------

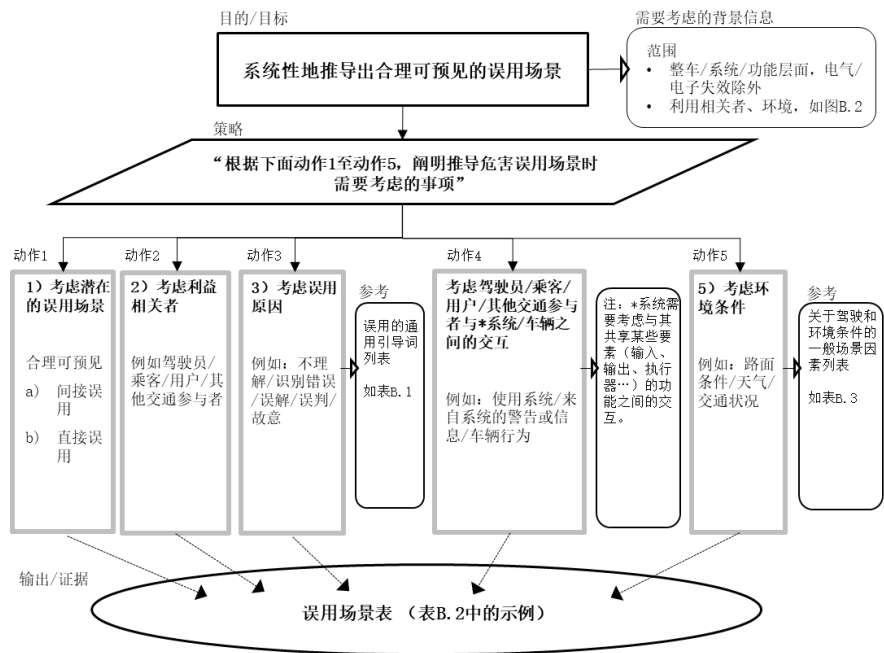
	安全策略3：车辆汇入过程中，侧向运动加速度最高0.15g；为躲避相邻车道障碍车辆时，自车主动加速能力最高0.3g，减速能力最高0.3g。
6.2 报警和降级策略	报警和降级策略1：当相邻车道一直有障碍物车辆时，自车会通过声光提醒驾驶员接管，驾驶员接管手力矩为1Nm；提醒时间。 报警和降级策略2：如果驾驶员不接管（合理的人员接管延迟 $\leq 3s$ ），车辆应在达到渐变段前完成主动停车，主动停车减速度0.3g（紧急避障减速度最高仍为0.9g）。
7. 其他	若有

附录 B
(资料性)
场景和系统分析指南

B.1 推导 SOTIF 误用场景的方法

B.1.1 概述

对于与SOTIF相关的系统，在执行安全分析时，考虑潜在的合理可预见的误用至关重要。可以从各种来源获得包含SOTIF相关误用的场景，例如：经验教训、专家知识、设计人员的头脑风暴等。B.1给出了系统性地推导SOTIF相关误用的示例方法，以支持SOTIF安全分析。图B.1给出了该示例方法的总体概念，并概述了与SOTIF相关的误用示例。人为因素分析的方法在参考文件[16]中进行了描述。



注：关于图B.1中每个元素的符号形状的含义，请参考表A.1。

图 B.1 SOTIF 相关误用场景的系统性推导（示例）

B.1.2描述了包含SOTIF相关误用的场景的考虑要点和示例场景因素表。

B.1.2 误用的安全分析方法的流程

以下描述了推导SOTIF相关误用时可考虑的要点：

a) 潜在的误用场景

考虑两种类型的误用情况：

- 在识别危害事件时，将“合理可预见的间接误用”与潜在的系统危害行为结合起来考虑；及
- “合理可预见的直接误用”作为一种潜在的触发条件，可能直接引发危害行为。

b) 利益相关者

考虑是谁引起了导致危害的SOTIF相关误用（如，驾驶员、乘客、用户、其他交通参与者）。

c) 误用原因

在考虑SOTIF相关误用的原因时，从典型的人为误用过程（识别、判断和行为）中得出的通用引导词可能是有用的。

表B. 1中描述了可能的引导词示例。

表B. 1 人为错误引导词

过程	引导词	示例
识别	1. 不理解	由于用法复杂或信息不足而无法正确操作。
	2. 错误识别	由于信息过多而无法正确识别。
判断	3. 判断错误/误判	由于错误印象或误解而导致的误判（如，自行车架的安装改变了GNSS天线的环境）。
行为	4. 走神/失误	由于注意力不集中（分心、疲劳、惯性自满等）而导致的失误。
	5. 故意	违反社会规则、普遍认同的人类行为、正确的操作方法（根据用户手册）。
	6. 无能力	难以操作。

- d) 驾驶员/用户、系统和车辆之间的交互
- 误用的原因可能是驾驶员/用户与系统/车辆接口之间交互的错误传达或时间限制（见图B. 2）。例如，可推导出以下接口主题：
- 驾驶员进行的系统操作（使用）：从驾驶员到系统/车辆的接口。
- 示例1：预期由驾驶员的语音指令激活的系统，也可能由于乘员之间的对话中说出的“关键词”而被意外激活。
- 系统发出的警告通知：从系统/车辆到驾驶员的接口；及
 - 系统/车辆行为：从系统/车辆到驾驶员的接口。

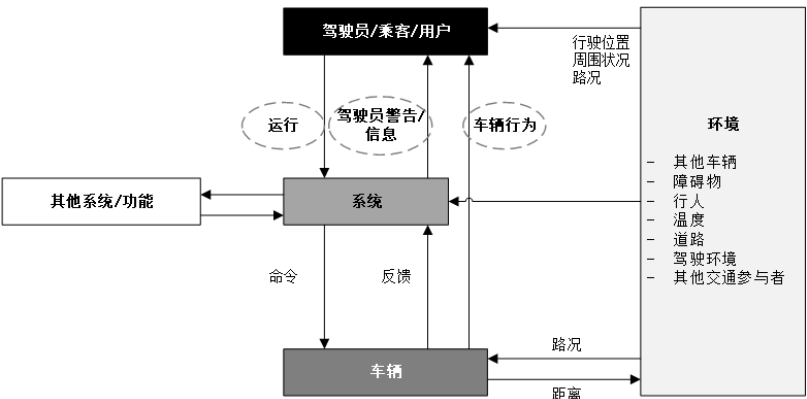


图 B. 2 驾驶员/用户、系统和车辆之间的交互示例

- 注1：图B. 2中的方框和箭头具有以下含义：
- 方框：可能与系统交互的外部因素；
 - 箭头：可能的交互。
- e) 考虑用例和场景的环境条件
- 在推导SOTIF相关误用时，可以考虑环境的影响，包括路面条件。
- 示例2：表 B. 3 或表 B. 4 描述了用例场景中考虑的一些环境条件。
- 注2：表B. 3或表B. 4既可用于功能不足场景分析，也可用于包含SOTIF相关因素的场景分析。作为替代方案，误用案例可以与危害识别活动（见第6章）和此处的驾驶情况类别相关联。
- 在考虑上述要点1)至5)来推导包含SOTIF相关误用的场景时，可以考虑使用如表B. 2的场景表。

表B. 2 基于与 HAZOP 类似的引导词方法的误用场景表示例

1) 潜在的SOTIF 相关误用场景	2) 利益相关者	3) 误用原因		4) 驾驶员与系 统/车辆之间的 交互	5) 环境条件 (见表B.3)	导出的危害误用 场景
		过程	引导词			
在执行2级DDT 时，例如在高速 公路上进行车道 保持辅助和自适 应巡航控制，由 于某个性能局 限，车辆无法估 计车道边界位 置。如果车道边 界信息丢失，系 统无法检测车辆 是否会偏离车 道，则会通知驾 驶员。	驾驶员...	识别	1. 不理解	操作（使用）
				车辆行为
				警告/信息	高速公路，弯 道，车道白线突 然不清晰。	因为不知道警告 的含义，驾驶员 没有接管车辆的 控制权，车辆驶 离车道。
			2. 错误识别	操作（使用）
				车辆行为
				警告/信息
		判断	3. 判断错误/误 判
		行为	4. 走神/失误
			5. 故意 驾驶员离开座位
			6. 无能力 驾驶员没有集中 注意力 驾驶员睡着
...

注3：HAZOP和STPA（系统理论过程分析，其应用示例见B.4）等方法可用于推导SOTIF相关误用的场景。

注4：图B.1的方法无意对所有组合进行全面分析。图B.1中概述的方法仅作为示例，可用于启发特定SOTIF开发所需的分析。仅选择了影响危害事件的因素进行分析。对危害事件没有影响的因素可以记录为不适用。

B.2 SOTIF 安全分析方法的场景因素构建示例

本子章节给出了开发场景的示例方法以支持危害识别（第6章），安全分析（第7章）以及为已知和未知触发条件创建验证/确认场景（第10和第11章）。

采取以下步骤来识别和评估因各种条件（如部件特征、工艺、物理现象和环境条件）而影响系统性能的潜在触发条件：

- a) 为了进行分析，可把系统功能分解为以下要素：感知、规控、执行。
- b) 根据触发条件的每个要素的影响因素（参考表 B.3 和表格 B.4）构建潜在功能不足的场景。

注1：可以在SOTIF相关场景的生成中包含GB/T 34590中的HARA场景生成表。

注2：也可以在参考[17]中找到关于如何为正在考虑的操作推导一组代表性的具体测试场景的建议。

表B.3 场景因素示例（未穷举）- 案例-1

分类	因素
天气	良好
	多云
	雨；“小雨”，“大雨”

分类	因素
	雨夹雪
	雪（积雪）；“小雪”，“大雪”
	冰雹
	雾；“浓雾”，“薄雾”
	风
一天中的时间	清晨
	白天
	傍晚
	夜晚
道路/车道形状	直道
	弯道
	下坡
	上坡
	倾斜道路
	阶梯路
	不平坦道路
	比利时砖路
	窄路，宽路
	有中间带的道路
	井盖
	道路交汇
	道路分支
	坑
道路特征	隧道
	地下通道
	桥梁
	高架桥
	立交
	菱形路
	收费站
	闸门
路面条件	干燥
	潮湿
	低附路面
	交叉道
	水槽路
	砾石路
光照	日光直射（刺眼）
	无月之夜
	月光之夜

分类	因素
	街灯
	背光
	黄昏
自车条件	传感器的不规则干扰（例如：撞击导致传感器视野的变化）
	传感器变化（例如：装配松动）
	传感器起雾
	传感器脏污（灰尘，泥浆，雪，冰等）
	车辆姿态（例如，当车辆因突然制动而倾斜时传感器视角发生变化）
	车辆情况（例如，当自车牵引大型拖车时，传感器视野被遮挡）
	实际车辆重量（例如，带牵引）
	重量分布
	轮胎（例如：温度，胎面纹路或橡胶硬度）
	制动盘（例如：结冰或过温）
自车操作	车辆加速
	车辆减速
	车辆匀速行驶
	车辆停止
	高速行驶
	低速行驶
	车辆转向中
	车辆突然偏离路径
	过路口
	右转或左转
	施工区绕行
	接近交叉路口
	（交通）环岛
	上下匝道
	穿越铁轨
周围车辆： ——前车 ——侧边车辆 ——对向来车	周围车辆的位置
	前车减速
	前车突然减速
	前车加速
	前车突然加速
	插入车辆
	走走停停的跟车
	有车辆在自车右方同向行驶
	有车辆在自车左方同向行驶
	有对向车辆驶来
	对向车辆打开了远光灯
	摩托车穿行

分类	因素
	自行车
	来自周围车辆的强干扰（例如：来自周围车辆雷达的干扰）
其他交通参与者	行人
	卡车
	摩托车
	异形车辆
道路外目标（周围）	侧壁
	标志（各种位置方向）
	柱
	隧道
	停车位
	在高架桥下
	路缘
	护栏
	塔架
	停在路边的车辆
	跳出来的动物
	铁路交叉口
	工地
	标记的人行横道
	路边的水洼
道路上的目标： ——道路标记	博斯点，猫眼，辉石（嵌入式）反光块
	实线-白色，黄色
	虚线-白色
	人行横道
	跳动路面
	减速带
	信息（箭头、限速、让行、减速等）
	没有车道标记
	断线
	变浅的车道标记
	多车道标记
道路上的异物	动物尸体（路杀）
	垃圾、轮胎胎面等
	颗粒物、灰尘、污垢、沙子和泥土
	建筑材料、沥青、混凝土、钉子、螺钉和其他通常锋利的物体
	从移动车辆上意外或蓄意掉落的固态物体
	在交通碰撞中从车辆上脱落的碎玻璃、塑料和其他固体材料

表B.4 场景因素结构（非穷举）示例 - 案例-2

第1层因素	第2层因素	第3层因素	第4层因素		
道路几何形状和拓扑	道路类型	高速公路			
		乡村路			
		城市路			
	道路几何形状	直道			
		弯道			
	道路海拔	水平			
		上坡			
		下坡			
	道路横断面	车道数量			
		车道标记			
	道路表面	粗糙度	沥青		
			混凝土		
			铺装路面		
			碎石路面		
		损坏情况	裂缝		
	道路交叉口		坑洞		
			发散		
			汇合		
交织					
		路口			
		道路设施和限制	边界	柱	
				护栏	
				混凝土屏障	
噪音屏障					
隧道	架空间隙				
桥梁	架空间隙				
	实体在桥下移动				
交通标识		交通灯			
		警告			
		限速			
临时物理限制	车道重新分配				
	车道标记				
	道路作业标识				
	道路作业路障				
可移动实体	实体类型	车辆	小轿车		
			卡车		
			大巴		
			轻轨		
			摩托车		
			应急车辆		

			农用车
			脚踏车
		行人	婴儿
			幼儿
			成人
		动物	
		目标	
	操作	巡航	高速
			低速
		速度变化	减速
			加速
		跟随	
		靠近	
		通过	
		变道	左
			右
		转向	左
			右
		掉头	
		安全停车	
	相对位置	左	
		右	
		前	
		后	
第5层因素			
环境条件	一天中的时间	清晨	
		白天	
		傍晚	
		夜间	
	气候条件	温度	
		能见度	
		风	
		云	
		降水	雨
			冰雹
			雨夹雪
			雪
	光照条件	日光	
		月光	
	路面条件	干燥	
		潮湿	

		雪覆盖	
		结冰	
第6层因素			
数字信息	V2X信息		
	数字地图数据		
<p>注：此表中层定义如下：</p> <p>第1层 街道布局和路面条件；</p> <p>第2层 交通引导设施，例如标识，障碍和标记；</p> <p>第3层 临时建筑工地的拓扑和几何叠加；</p> <p>第4层 道路使用者和目标，包括基于操作的交互；</p> <p>第5层 环境条件（例如：天气和时刻），包括其对第1到第4层的影响；</p> <p>第6层 数字信息，包括其对第1到第5层的影响。</p>			

示例1：用例构建：气候 = 雨，一天中的时间 = 白天，道路形状 = 直道、下坡，路面条件 = 潮湿，自行车操作 = 车辆停止，其他车辆 = 右侧对向来车，行人 = 无，道路外目标 = 无。

注1：表B. 3和表B. 4并不全面。因此，在构建场景时可考虑其他因素，诸如当地的驾驶习惯与基础设施。

注2：当开始SOTIF分析来识别潜在的危害场景及其触发条件时，下列性能局限/触发条件分类可能有用：

- a) 感知局限。
例如，气候，一天中的时间，道路/车道形状，自行车条件，周边车辆，其他交通参与者和道路外目标都是可能的触发条件。
- b) 交通相关条件；及
例如，道路/车道形状，路面条件，周边车辆，自行车操作，事故，其他交通参与者和道路外目标都是可能的触发条件。
- c) 自行车相关问题（影响自行车行为或性能的问题）。
例如，自行车传感器安装位置容易堆积制约性能的碎屑或灰尘。

注3：触发条件不仅可以由单个因素组成，还可以是多个因素的组合。

注4：构建场景的过程中，可以基于场景要素与特定功能、系统/组件或SOTIF活动(ODD的定义、验证与确认计划等)的相关性形成优先度子集。表B. 5展示了一个基于定性规则的场景优先度子集，该子集用于对基于雷达的功能开展确认活动。

在该示例中，考虑对于一个仅依赖雷达的系统，不受晚上或白天的影响，因而在相关要素中不予考虑。

表B. 6基于表B. 5，展示了一个基于定量规则的场景优先度子集，该子集通过分析相关要素的发生概率，可支持对基于雷达的功能开展确认活动时，计划确认场景的先后顺序，及分析确认活动的覆盖率。

表B. 5 基于定性规则的场景优先度因素子集示例（例：考虑基于雷达的功能确认）

分类	相关要素	优先度子集
气候	雨	子集1
道路特征	隧道	
时刻	任意 / 不相关	
道路外目标	标志（位置很高）	
...	...	
...	...	子集n

表B. 6 基于定量规则的场景优先度子集示例

分类	相关要素	发生概率	优先度子集	优先度子集发生概率	场景总体覆盖率
气候	中雨	a	子集1	a*b*c	a*b*c+ d*e*f+...
道路特征	直行隧道	b			
...	...	c			
气候	大雨	d	子集2	d*e*f	
道路特征	转弯隧道	e			
...	...	f			
...	子集n	...	
注：相关要素的发生概率需要基于目标场景的统计数据或研究，准确定义发生概率可能难以实现，但基于这些分析和论证，为SOTIF确认计划的合理制定及场景的覆盖率论证提供了支持，因而是有益的。					

B.3 用于识别和评估潜在触发条件和功能不足的安全分析的示例

B.3.1 系统性地识别触发条件的分析方法

随着驾驶自动化水平的提高，触发条件变得更加复杂和难以识别，需要多种分析技术结合道路测试，以充分地探测已知和未知的危害场景。在进行识别触发条件的分析时，可以考虑以下方法：归纳分析、演绎分析、探索性分析、探索性仿真（采用本示例中使用的高级组合技术或其他被认为合适的技术）和探索性驾驶（采取充分的安全措施）。

归纳和演绎分析有助于发现功能不足、输出不足及触发条件方面的危害事件的成因，并探索它们的因果关系。但是，当使用新技术（如机器学习）或ODD包含大量场景时，不能声称这些分析足以找到所有相关的不足和触发条件。

随着驾驶自动化水平的提高，在系统实现了一个错误的置信状态但原因尚不明确的情况下，增加探索性分析方法可能会有所帮助。例如，高度自动化的驾驶系统错误地认为自己处于无碰撞路径上，或者错误地认为自己能够或已经避免了碰撞。错误置信状态的根源可能来自单个或多个要素。例如，由于高威胁目标靠近其他低威胁目标而被错误地归类为低威胁目标，或者由于某些物理限制，车辆无法执行指定路径。诸如系统理论过程分析（STPA）之类的分析可以作为一种合适的技术，因为这种分析将系统、场景和人之间的交互视为危害来源。

最后，对于识别触发条件，探索性仿真和探索性驾驶是有用的自下而上的工具。但是，每种方法都有其局限性。当应用方法和准则来评估SOTIF的实现时，可考虑这些方法的局限性。

B.3.2 因果树分析示例

基于第6章中识别的危害事件，可以使用适当的演绎风险评估方法（类似于用于功能安全的经典故障树分析方法），来确定潜在定义不足、性能局限和触发条件。

注：因果树分析是一种确定事件根本原因的合适方法，可用于识别和理解特定危害事件的触发条件。

当识别了系统不足和触发条件时，就能够确定导致危害的事件组合，并确定足以引起危害的最小割集。结果可用于识别重要的潜在相关性和最严重的不足，并确定为减轻风险而采取的措施是否充分，见7.4。此外，结果可用于将确认活动按优先级排序或分组。

示例：在某ACC系统的范围内分析突然意外减速的危害事件。该系统由一个调节器组成，基于驾驶员的请求速度和用于检测障碍物及测量车前距离的立体相机的输入，该调节器可控制给发动机的动力和执行制动。功能不足树模型如图B.3所示。基于功能不足分析，顶事件G0的最小割集可用以下的等价布尔代数函数来表示： $TOP = B01 + B02 + (B03 \times B04) + (B05 \times B06)$ 。

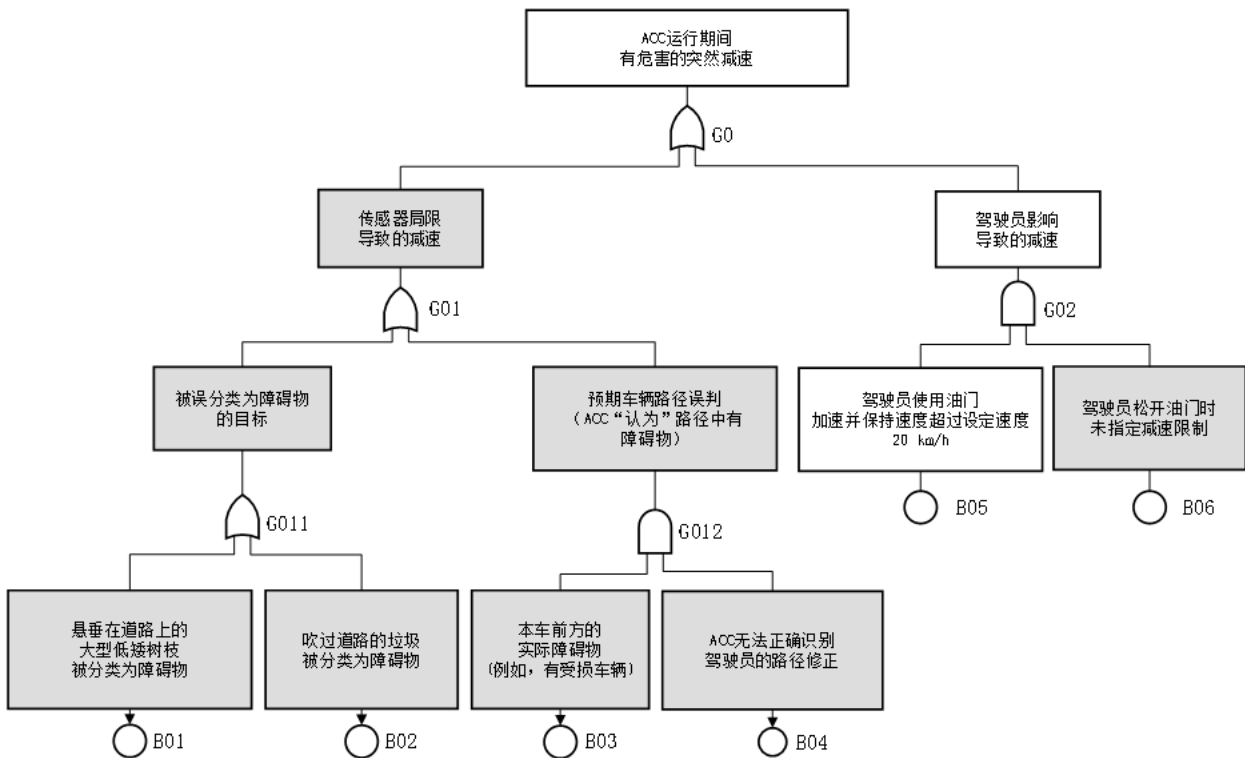


图 B. 3 因果树分析

除了演绎分析外，通常还进行归纳分析，通过分析功能、架构和详细设计以及通过评估系统实现过程中新识别到的危害来提高安全分析的完整性。

B. 3. 3 SOTIF归纳分析示例

B. 3. 3. 1 SOTIF 归纳分析工作流程

图B. 4所示的SOTIF分析工作流程旨在描述活动，以支持：

- 识别和评估可能导致危害行为的潜在功能不足，危害行为可以由驾驶场景的已知特定条件引发；
- 识别和评估可能引发危害行为的潜在触发条件，危害行为由已知潜在功能不足所致；及
- 识别修改措施以避免或减轻 SOTIF 相关的风险。

考虑各个方面的顺序(从潜在功能不足到潜在触发条件,或从驾驶场景的特定条件到潜在功能不足)取决于分析人员的偏好。

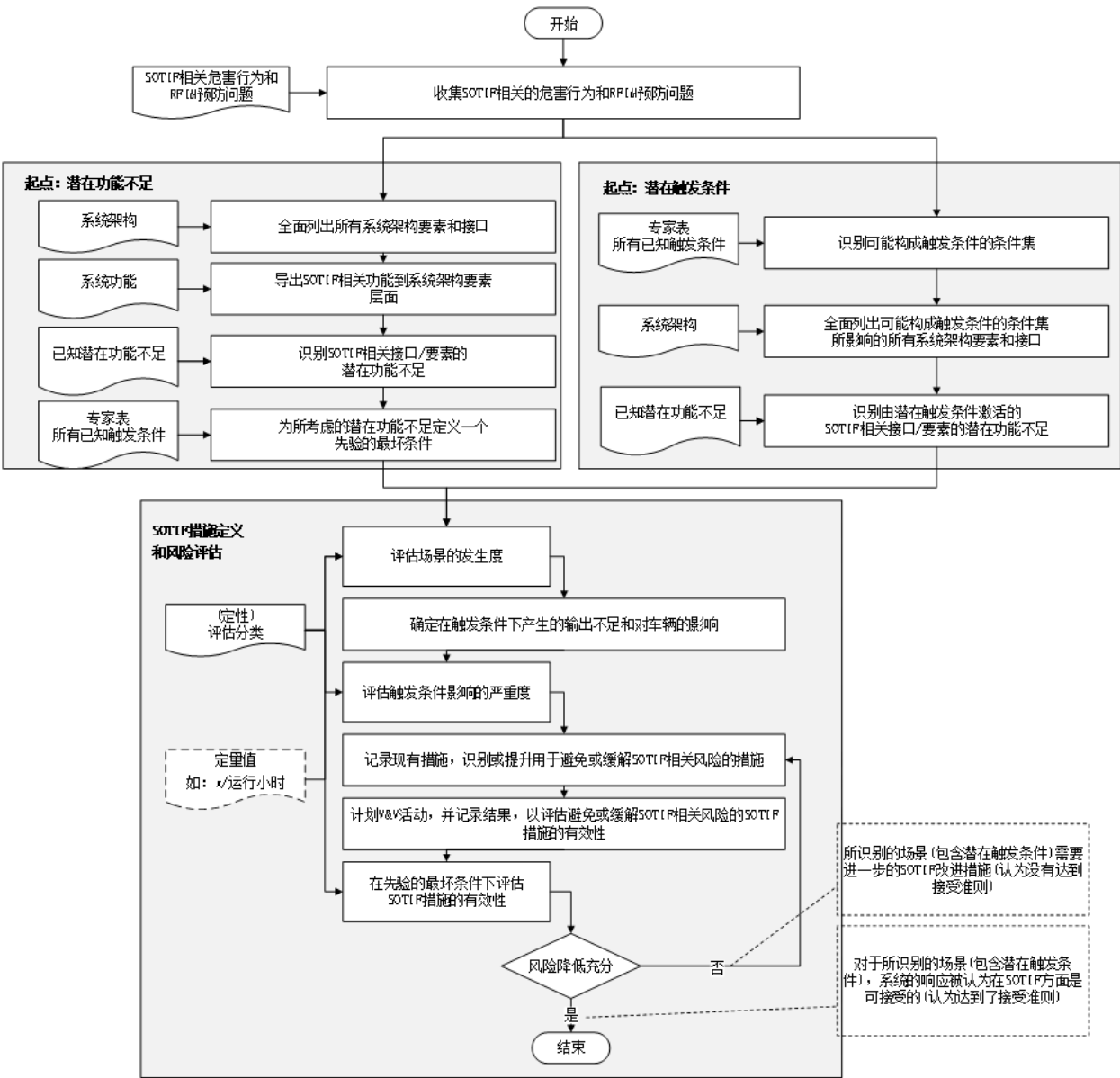


图 B. 4 SOTIF 归纳分析工作流程

SOTIF相关风险分析可基于可能性和影响的定性评价量表或基于定量值, 如误报率、每个运行小时的触发条件数量。这些结果可用于优先评估某些场景或要素。

注: 考虑了定性评价的统计分析和图表, 如帕累托分析、风险矩阵, 可用于支持确定如7.4中定义的触发条件的可接受性。然而, 由于评估标准的差异性, 使用预先定义的评级来确定定性分析的可接受性是不合适的。

B. 3. 3. 2 从潜在功能不足到触发条件的 SOTIF 分析示例 (基于系统的分析)

这种归纳分析旨在首先识别系统要素潜在的功能不足, 其次识别可能激活这些已识别的潜在不足的场景条件, 这些场景条件可能导致输出不足、危害行为或RFIM预防问题。

注1: B. 3. 3. 2中使用术语“RFIM(reasonably foreseeable indirect misuse) 预防问题”表示系统无法避免或减轻合理可预见的间接误用。

以下示例展示了紧急制动系统不同要素的归纳分析。呈现在表B.7、表B.8和图B.5中的分析并非详尽无遗的。其更倾向于说明对“感知-策划-执行”模型中涉及的不同类型系统要素的SOTIF分析，系统要素为：

- 摄像头硬件传感器成像器（硬件单元 HW43）；
- 摄像头硬件加速器或“IP”（硬件单元 HW32）；
- 摄像头软件分类功能（软件单元 SW11）；
- 制动扭矩执行系统（系统 SYS 12）。

这些系统要素用于系统功能“在出现迎面而来的或横向穿越的目标时制动”（SYS23.1）。如果检测到的目标在定义的目标列表（参考#RRR）中，并且在定义的紧急条件下（参考#CDNXX），则需要紧急制动。

每个系统要素都有其自身潜在的功能不足，结合“先验的”最坏条件，可能导致危害行为、RFIM预防问题或输出不足。

注2：功能不足是系统要素的属性，而“先验的”最坏条件是所考虑场景的属性。

对于每个元组（系统要素、相关潜在功能不足、相关潜在触发条件），进行SOTIF相关风险分析，旨在识别改进SOTIF的措施，验证其有效性，并用适当的理由评估残余风险。

表B.7 从潜在功能不足到触发条件的 SOTIF 分析示例

ID	潜在导致SOTIF相关危害事件的系统要素				潜在触发条件 已知潜在功能不足的先验的最坏条件				潜在触发条件 影响		处理输出不足的措施 (包括现有的和新提议的)			接 受 理 由
	系统架构功能	分 配 到 系 统 或 硬 件 / 软 件 要素	SOTIF 相关接 口 / 要 素	系 统 设 计 中 已 知 潜 在 功 能 不 足	场 面 特 征 (环 境 条 件、道 路 / 市 区 基 础 设 施)	驾驶场景(动作、 事件、目标和数 值)	驾驶员、其他驾 驶员、道路使用 者的行为	发 生 度	未通过 任何措 施解决 输出不 足的整 车层面 影响	危 害 事 件 的 严 重 度	设计中改 善SOTIF的 措施	提 供 系 统 响 应 证 据 或 设 计 措 施 有 效 性 证 据 的 验 证 措 施	措 施 有 效 性	
ID1.1	系统要素实现功能SYS23.1：在紧急条件下（参考#CDNXX）出现迎面而来的或横向穿越的目标（目标列表：参考#RRR）	硬件单元 HW32：摄像头IP	IP结果	影响距离估计的图像分辨率限制	白天，干燥的道路	以90km/h直行	前方车辆在本车车道上轻微偏出（前车距本车>100m）	根据评级规则完成	误报：迎面对标检测导致非预期的车辆减速 <-X m/s²	根据评级规则完成	使用不同技术的传感器：激光雷达、雷达	测试报告TC#225通过，负责：A组	根据评级规则完成	见表B.8
ID1.2		硬件单元 HW43：摄像头传感器硬件	传感器结果	在弱光条件下较差的图像渲染	晚上，干燥的道路	在弱光条件和干燥道路上的所有操作	没有进一步条件	根据评级规则完成			使用不同技术的传感器：激光雷达、雷达	测试报告TC#226通过，负责：B组	根据评级规则完成	见表B.8
ID1.3	时制动	软件单元	目标分类结果	边角案例CC#52	CC#52条件：	高峰时间，高交通流量，繁忙的	没有进一步条件	根据评级			新架构新算法			见表B.8

		SW11: 目标分类		下低性能	场面中有大量的移动目标需要处理	十字路口，一群骑自行车的人，一群骑摩托车的人，挂着许多旗帜的风景。在汽车前面但不在其运行轨迹上（例如，由于弯道）的移动目标。		规则完成			行动：OPL#227 C组		根据评级规则完成	
ID1.4		系统 SYS12: 制动扭矩执行系统	制动扭矩	执行器在温度<-10℃且低电压<9.5V下时间响应慢	冬天，积雪，温度<-15℃	电池电压低由于接近慢行车辆，AEB干预	没有进一步条件	根据评级规则完成	非预期的丢失 减速<-Z m/s ² AEB干预情况下较小的AEB减速	根据评级规则完成	新执行器行动：OPL#228 D组		根据评级规则完成	见表B.8
ID1.5		软件单元 SW11: 目标分类	目标分类结果	意外的/未训练的目标的错误分类	罕见的目标，不寻常的目标	在某些活动（例如足球比赛、游行）、假期（例如圣诞节、嘉年华）、汽车装饰、汽车装载期间驾驶	驾驶员在车顶行李架上安装了一些悬垂的物体，伸入到摄像头图像中（例如挂着一些纺织材料或绳子的梯子或运动器材）	根据评级规则完成	误报：迎面目标检测导致非预期的车辆减速<-X m/s ²	根据评级规则完成	在驾驶周期开始时，检查摄像头检测到的异常目标，对检测到的目标进行持续的	仿真或确认试验	根据评级规则完成	见表B.8

											合理性检查，在汽车用户手册中，指示驾驶员不要让任何物体伸入到摄像头视野内		
--	--	--	--	--	--	--	--	--	--	--	--------------------------------------	--	--

SOTIF分析表B. 7分为四大列，分别记录和分析：

- a) 在适当的抽象层面（例如在系统架构的最低层面）进行描述的潜在导致输出不足的系统要素，即潜在的所有系统要素；
- b) 与 a) 中列出的系统要素相关的潜在触发条件，在外部或内部环境层面进行描述；
- c) 在没有任何 SOTIF 措施情况下的这些潜在触发条件的影响，在最高抽象层面进行描述，如整车层面；
- d) 用于处理 a) 中列出的输出不足的现有和计划措施，在适当的抽象层面进行描述，如实现层面。

表B. 8 从潜在功能不足到触发条件的 SOTIF 分析示例（续）

ID	接受理由
ID1. 1	<p>在多条且不同的窄路上进行定向试验和耐久性试验（<道路>标记为“窄的”、<车速> >90km/h、<一天中的时间>在整个驾驶数据集中为白天）证明：</p> <p>——出现以下情况的概率：摄像头 IP HW32 的图像分辨率限制，在没有 SOTIF 措施(通过停用雷达和激光雷达)时会发生目标检测误报而导致非预期地车辆减速，从而影响距离估计，这个概率被确认为相当低：0 个车辆非预期减速事件；0 次探测到“潜在目标”；</p> <p>——如果激活，雷达和激光雷达的组合被确认是有效的措施：</p> <p>当在融合算法中有雷达和激光雷达信息时，在图像分辨率限制会影响距离判断的相同条件下的重复试验呈现出更好的反应时间（-x%）和更高的置信度估计，以确认没有目标。证据：TC#225通过。</p> <p>问题点可以关闭。</p>
ID1. 2	<p>在夜间的定向试验和耐久性试验（<一天中的时间>在整个驾驶数据集中为“夜晚”或“黄昏”）证明：</p>

ID	接受理由
	——出现以下情况的概率：摄像头传感器 HW43 的图像渲染限制在低光照条件下，在没有 SOTIF 措施(通过停用雷达和激光雷达)时会发生目标检测误报而导致非预期地车辆减速，从而影响图像，这个概率被确认为相当低：0 个车辆非预期减速事件；6 次探测到“潜在目标”，但由于不合理性未被决策算法采用；当在融合算法中有雷达和激光雷达信息时，在相同条件下的重复试验呈现出更好的反应时间（-x%）和置信度，以确认没有目标。 问题点可以关闭。
ID1.3	极端用例CC#52是一组在耐久性试验和仍在进行的车队试验期间未遇到的特殊条件。 然而，极端用例CC#52并不能被归类为“不可能”，因为其已在交通场景仿真环境中重现。C组的替代算法显示出轻微的性能提升（更高的置信度估计），尽管在这种仿真环境中并不显著。 问题点仍然未决，以确认是否需要新架构或新算法。
ID1.4	D组最近进行的测试识别到，当前制动扭矩执行器（变体A）的规范定义不足。在低电压值（仍在规定范围内）和瑞典北部的低环境温度（-30° C至-15° C）下，发现非预期地丢失减速（<-Z m/s ² ）。 相同测试证明了具有鲁棒性的制动执行器原型（变体B）在达到确认目标方面的有效性。需求规范已更新。 问题点仍然开放，将用变体B型的发布版本重复相同的测试。
ID1.5	仿真结果未决

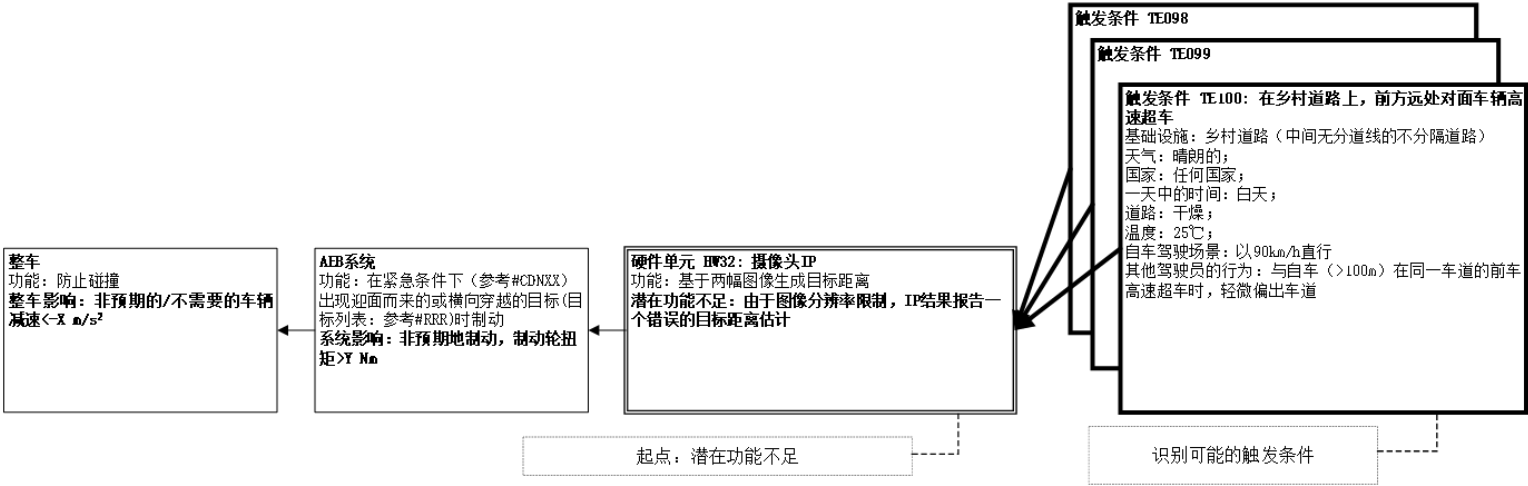


图 B.5 以潜在功能不足为起点的 SOTIF 因果树，用于解释表 B.7 和 B.8

B.3.3.3 从触发条件到潜在功能不足的 SOTIF 分析示例（基于场景的分析）

该SOTIF归纳分析旨在首先识别可导致输出不足、危害行为或RFIM预防问题的驾驶场景的条件，其次识别受这些潜在触发条件影响的系统架构功能或要素。

以下示例展示了一个紧急制动系统要素的归纳分析，其场景条件“绘制在路上的行人标识”可导致一个危害行为。呈现在表B. 9、表B. 10和图B. 6中的分析并非详尽无遗的。其更倾向于说明对“感知-策划-执行”模型中涉及的不同类型系统要素的SOTIF分析，系统要素为：

- 雷达硬件要素（硬件单元 HW53）；
- 摄像头硬件加速器或“IP”（硬件单元 HW52）；
- 摄像头软件分类功能（软件单元 SW11）；
- 制动扭矩执行系统（系统 SYS 12）。

这些系统要素用于系统功能“在出现迎面而来的或横向穿越的目标时制动”（SYS23.1）。如果检测到的目标在定义的目标列表（参考#RRR）中，并且在定义的紧急条件下（参考#CDNXX），则需要紧急制动。

该分析倾向于识别可能受到相同潜在触发条件影响的系统要素功能不足。例如，在如下的示例中，摄像头IP（硬件单元HW52）的算法可能在出现“绘制在路上的行人标识”的情况下，触发一些目标检测误报，尽管只是特殊的边角案例（CC#536）。

注：功能不足是系统要素的属性，而潜在触发条件是所考虑场景的属性。

对于每个元组（潜在触发条件，相关的系统要素的潜在功能不足），进行SOTIF相关风险分析，旨在识别改进SOTIF的措施，并用适当的理由评估残余。

表 B.9 从触发条件到潜在功能不足的 SOTIF 分析示例

ID	潜在触发条件				潜在导致SOTIF相关危害事件的系统要素				潜在触发条件影响		处理输出不足的措施 （包括现有的和新提议的）			接受理由
	来自专家表的已知危害用例			发生度	受触发条件影响的系统架构功能	受触发条件影响的系统架构要素	SOTIF相关接口 / 要素	系统设计中的潜在功能不足	未通过任何措施解决输出不足的整车层面影响	危害事件的严重度	设计中改善SOTIF的措施	提供系统响应证据或设计措施有效性证据的验证措施	措施有效性	
	场面特征 （环境条件,道路/市区基础设施）	驾驶场景 （动作、事件、目标和数值）	驾驶员、其他驾驶员、道路使用者的行为											
IDA. 1	基础设施：画在路上的行人标识 天气：良好 国家：所有一天中的时间：夜间弱光 道路：干燥 温度：25℃	以50km/h直行(市区)	跟随车辆靠近自车（<5m）	根据评级规则完成	系统要素实现功能SYS23.1:在出现迎面而来的或横向穿越的目标时制动	硬件单元HW52:摄像头IP	IP结果	无法区分烟雾和实体目标	误报：检测到行人导致非预期的车辆减速<-Xm/s2,导致跟随车辆追尾碰撞	根据评级规则完成	使用不同技术的传感器：激光雷达，雷达	TC#234 通过负责：A组	根据评级规则完成	见表B. 10
IDA. 2						硬件单元HW63:雷达要素	雷达结果	该场景下无			无		不适用	见表B. 10
IDA. 3						软件单元SW11:目标分类	目标分类结果	该场景下无（假定1002输入比较/表决没有不足）			基于完全冗余和多样算法的表决（HW52, HW63, SW11）	参考VC2 通过	根据评级规则完成	见表B. 10
IDA. 4						系统 SYS12：制动扭矩执行系统	制动扭矩	该场景下无			无		不适用	见表B. 10

SOTIF分析表B. 9分为四大列，分别记录和分析：

- a) 在外部或内部环境层面描述的潜在触发条件，如已知潜在触发条件或随机潜在触发条件；
- b) 在暴露于 1) 中列出的潜在触发条件下，可能导致输出不足的系统要素，在适当的抽象层面进行描述（例如在系统架构的最低层面）；
- c) 在没有任何 SOTIF 措施情况下的这些潜在触发条件的影响，在最高抽象层面进行描述，如整车层面；及
- d) 处理 2) 中列出的输出不足的现有和计划措施，在适当的抽象层面进行描述，如实现层面。

表 B. 10 从触发条件到潜在功能不足的 SOTIF 分析示例（续）

ID	接受理由
IDA. 1	在加拿大不列颠哥伦比亚省本纳比市的三角洲大道上，于布伦特伍德公园和圣十字小学之间驾车进行定向测试： ——出现以下情况的概率：摄像头 IP HW52 识别到虚目标，在没有 SOTIF 措施(通过停用雷达，激光雷达和基于光流的机制)时会导致非预期地车辆减速，这个概率被确认为相当低：0 个车辆非预期减速事件；1 次探测到“潜在目标”，但由于确认时间不足未被决策算法采用。实际上，即使在低速行驶时，图像也仅在很短的时间内没有失真，这不足以检测道路上的行人。 ——如果激活，雷达和激光雷达的组合被确认为有效的措施： 当在融合算法中有雷达和激光雷达信息时，在相同条件下的重复试验呈现出更高的置信度，以确认没有目标。 证据：TC#234通过。
IDA. 2	不适用。雷达要素不会受限于道路标识的误判。
IDA. 3	对于该特殊场景，在软件单元SW11中没有识别到系统设计弱点。 然而，基于能够确认目标存在的多个多样算法的决策算法，被视为一个解决SW11单元功能不足的非常有效的措施，如有。
IDA. 4	不适用。制动扭矩执行器不会受限于道路标识的误判。

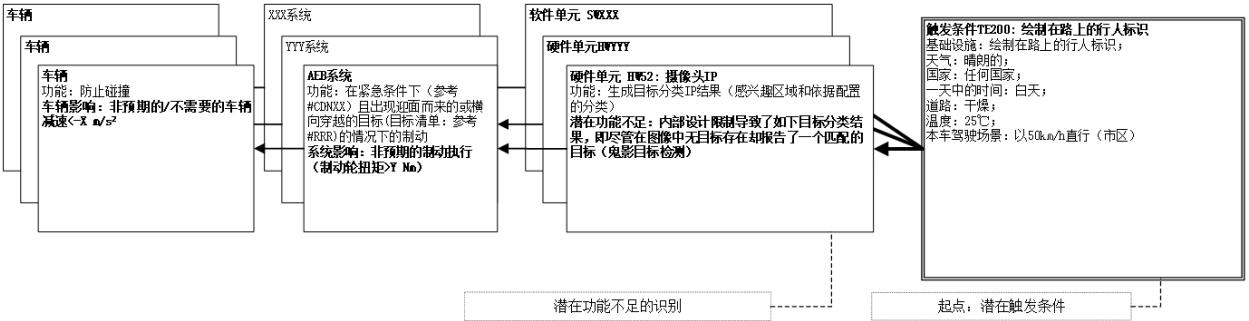


图 B. 6 以潜在功能不足为起点的 SOTIF 因果树，用于解释表 B. 7 和 B. 8

B. 4 在 ADAS 和自动驾驶车辆的 SOTIF 研究中应用 STPA 方法

B. 4. 1 介绍

STPA（系统理论过程分析）（参考[19]和[20]）是一种安全分析方法，用于评估复杂系统的安全和识别安全约束及要求。很多已发表的论文描述了STPA如何应用于汽车系统、ADAS和自动驾驶（参考[21]，[22]，[23]和[24]）。STPA对SOTIF很有用，因为可以解决功能不足、系统在不适用的环境中使用、人员误用等问题。

B. 4章节提供了一个简化的SAE J3016定义的L3级别高速领航驾驶系统的示例，该示例用STPA（系统理论过程分析）方法对第6章（危害识别）和第7章（触发条件的识别和评估）进行SOTIF分析。高速领航驾驶（HP）可在限定的环境条件下长时间控制整个车辆动态，而无需受到人类驾驶员的即时监督。但是，仍然需要人类驾驶员，并能够在指定的时间范围内接管车辆，这个时间范围通常在几秒钟到不超过最大指定时间。

B. 4. 2 STPA步骤1：定义分析的目的和范围

STPA的第一步识别出要防止的利益相关者的损失。一旦识别出STPA损失，就确定了STPA整车层面的危害。这些整车层面的状态或条件，连同一组特定的最坏环境条件，将会造成损失。表B. 11提供了高速领航驾驶的STPA损失和STPA整车层面危害。

表 B. 11 损失和危害识别的示例

情景/场景 (节选自HARA)	损失	潜在结果 (伤害)	整车层面危害 (来自HARA)
夜间行驶在高速公路上， 视野差，车速快。 逐渐接近前方一个速度较 慢的摩托车。	[L1]失去生命或人员伤亡	严重或致命的伤害	[VH1] 自车违反了与其它车辆之间的 最小距离阈值/要求。
...	[L2]	[VH1] ...

注：B. 4的其余部分包含规范示例。在这些示例中使用了“应”的表达方式。在本B. 4章节中，“应该”这个表达仅为示例的要求，而不是用来符合本文档的。

请注意，随后的STPA步骤系统性地分析了每个系统控制器（包括人类）的控制行为，以识别在特定场景下会潜在导致整车层面危害的特定行为及其原因。对于整车层面危害，一组整车层面的SOTIF要求（作为HARA的一部分）被识别出来，见表B. 12。

表 B. 12 危害和对应整车层面的安全约束

危害	整车层面的SOTIF要求 (整车层面安全约束)
[VH1] 自车违反了与其它车辆之间的最小距离阈值/要求。	[SC-1] 自车应该确保和其他车辆之间的安全距离。
...	

B. 4. 3 STPA步骤2：控制结构的建模

通过分析系统和功能规范来识别系统的控制层次体系及其接口环境，这被称为“控制结构”。提取控制器命令（也被称为“控制行为”）和来自受控过程与环境的反馈来进行分析。

高速领航驾驶的控制结构示例如图B. 7所示。

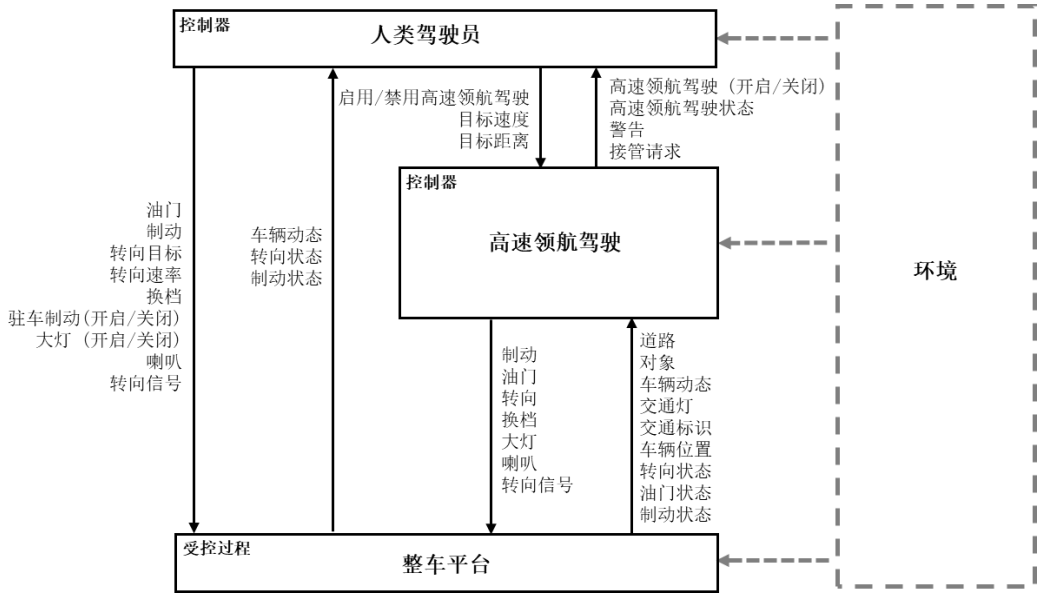


图 B. 7 高速领航驾驶的高层面控制结构

由于空间有限，B. 4中的STPA并没有更深入，对此类功能的更细层面控制回路模型感兴趣的读者请参考[25]中的图5。

B. 4. 4 步骤3：识别不安全的控制行为

STPA方法的下一步是识别不安全控制行为(UCA)，即在特定情况和最坏环境下会导致整车层面危害的行为。UCA及其相关危害和HARA被用来完成第6章的危害识别和风险评估。一个不安全的控制行为由五个要素组成，如图B. 8所示。

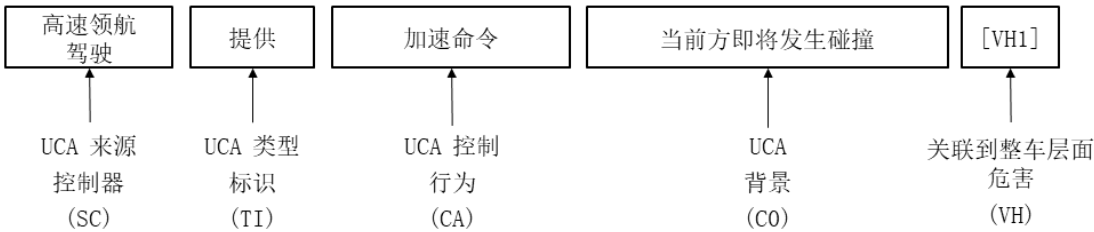


图 B. 8 不安全控制行为的 5 个要素

一些高速领航驾驶制动命令的不安全控制行为示例见表B. 13。

表 B. 13 高速巡航功能控制器发出制动命令时的不安全控制行为示例

控制行为	没有提供	提供	提供的过早，过晚，或顺序错误	提供的时间过长，或停止的过早
制动命令	UCA-1：当前方即将发生碰撞的时候，高速领航驾驶没有提供制动命令。[VH1]	UCA-2：当前方即将发生碰撞的时候，高速领航驾驶提供的制动命令不足。[VH1] UCA-3：当驾驶员提供油门命令时，高速领航驾驶提供了制动命令。[VH2]	UCA-4：当前方即将发生碰撞的时候，高速领航驾驶提供的制动命令过晚。[VH1]	UCA-5：高速领航驾驶在碰撞发生后过早地停止提供制动命令。（例如：在驾驶员还没有完成手动接管的时候就停止了制动命令。）[VH1]

请注意，每个不安全控制行为都可能导致至少一个整车层面危害（否则不会不安全），但也可能导致多个整车层面危害。

对于给定的UCA，可以定义控制器安全约束，以确保防止UCA的发生。控制器安全约束对控制器行为指定了需要满足的声明或不变量，以防止UCA的发生。

一些控制器安全约束（关于制动相关的UCA）见表B. 14。

表 B. 14 将 UCA 转换为要求（安全约束）

不安全控制行为	安全约束
UCA-1：当前方即将发生碰撞的时候，高速领航驾驶没有提供制动命令。[VH1]	SC-1：当前方即将发生碰撞的时候，高速领航驾驶应提供制动命令。[UCA-1]
UCA-2：当前方即将发生碰撞的时候，高速领航驾驶提供的制动命令不足。[VH1]	SC-2：当前方即将发生碰撞的时候，高速领航驾驶应提供足够的制动命令。[UCA-2]
UCA-3：当驾驶员提供油门命令时，高速领航驾驶提供了制动命令。[VH2]	SC-3：当驾驶员提供油门命令时，高速领航驾驶不应提供制动命令。[UCA-3]
UCA-4：当前方即将发生碰撞的时候，高速领航驾驶提供的制动命令过晚。[VH1]	SC-4：当前方碰撞即将发生的时候，高速公路巡航系统应在至少TBD秒前提供制动命令。[UCA-4]
UCA-5：高速领航驾驶在碰撞发生后过早地停止提供制动命令。（例如：在驾驶员还没有完成手动接管的时候就停止了制动命令。）[VH1]	SC-5：高速领航驾驶不能在驾驶员手动接管前停止制动命令。[UCA-5]

B. 4. 5 STPA步骤四：识别原因场景

STPA的最后一个核心步骤是确定导致危害的原因场景和相应的因果因素（如：触发条件，见7.3）。表B. 15概述了高速领航驾驶UCA-1的原因场景，以识别因果因素。

识别会导致当前不安全控制行为的系统控制器自身要素或其他要素的一个或多个输出不足的组合，是这个分析的第一步。这种一个或多个输出不足的组合在表B. 15中被称为“不足条件”。下一步，识别导致已确认不足条件的因果因素。这些因素可能是输出不足、功能不足或触发条件。

表 B. 15 识别因果因素

原因场景	UCA（危害行为）	不足条件	因果因素（触发条件，功能不足，输出不足）
CS-1	UCA-1：当前方即将发生碰撞的时候，高速领航驾驶没有提供制动命令。	IC-1：由于以下信息反馈不足，高速领航驾驶错误地相信不会立即发生碰撞：相对位置、相对速度、相对加速度、面向障碍物的方向。	CF-1:传感器安装不正确，传感器焦点或位置损坏，传感器被遮挡等。 CF-2：由于总线繁忙、消息优先级或仲裁不足、电磁干扰等而延迟且未及时收到反馈。 CF-3：反馈被认为是不正确的（被高速领航驾驶忽略），因为与其他反馈相冲突（例如，其他反馈表明车轮转速为零）。
CS-2	UCA-1：当前方即将发生碰撞的时候，高速领航驾驶没有提供制动命令。	IC-2：由于以下信息反馈不足，高速领航驾驶错误地相信不会立即发生碰撞：制动已施加。	CF-4：高速领航驾驶收到“已施加足够的制动或转向”的不正确反馈。

原因场景	UCA（危害行为）	不足条件	因果因素（触发条件，功能不足，输出不足）
CS-3	UCA-1：当前方即将发生碰撞的时候，高速领航驾驶没有提供制动命令。	IC-3：由于以下信息反馈不足，高速领航驾驶错误地相信不会立即发生碰撞：障碍物的大小或类型。	CF-5：高速领航驾驶收到的反馈不足，表明障碍物类型不会构成碰撞危险。 CF-6：高速领航驾驶收到没有碰撞障碍的反馈[因为传感器遮挡、传感器安装在错误位置/方向、传感器离线、障碍物在传感器视图视野之外、错误地识别不利天气条件（缺少算法功能）、未校准等]。
CS-4...	UCA-2：高速领航驾驶...	IC-4：高速领航驾驶...	CF-7：高速领航驾驶...

注：在此STPA示例中，同时考虑了SOTIF相关的问题和功能安全相关的问题。

B. 4. 6 识别控制措施和缓解措施，改进系统设计并导出要求

完成本文件中的STPA核心活动后，STPA的剩余活动可分配给相应的流程步骤，对于解决SOTIF相关风险的功能修改见第8章，相应的，如果是失效相关的原因，见GB/T 34590. 4-XXXX，第六章。这涉及制定适合满足 STPA 安全约束的可实施要求。

附录 C

(资料性)

预期功能安全验证和确认指导

C.1 验证和确认策略目的

系统的功能不足是SOTIF问题的根源。验证和确认策略旨在表明在已知和未知场景下的残余风险足够低，并符合6.5中定义的量化目标。本附录给出了导出和测试确认目标的概念。

一旦定义了确认目标，就可以根据第9章和第11章设计确认测试计划，以表明不存在由于已知和未知的危害场景(区域2和3)导致的不合理风险。确认通常包含物理测试(封闭道路测试、开放道路测试)和仿真测试的某种组合。作为第9章定义的确认策略的一部分，量化目标通常需要在物理测试和仿真测试之间进行分配。

确认活动包括在各种运行条件下测试车辆。它可以是软件在环(SIL)、硬件在环(HIL)和实车测试的混合。他可以包含结构化测试(例如设计且用于封闭道路测试)、专用分析和仿真。但关键的是，特别是对于区域3，需要基于确认策略，在足够全面的运行条件下进行充分的测试，以尽可能的暴露潜在的未知不安全场景。

针对区域3的测试场景可以包括：

- a) 已识别用例的已知参数的随机组合(例如，恶劣天气和特定交通条件的组合)；
- b) 已知场景的随机组合；
- c) 在开放道路测试中未识别的可能触发危害系统行为的特定场景。

仿真可用于快速探索各种相关场景。然而，仿真可能会受限于对环境、传感器和车辆模型的基本假设。对现实世界建模的准确性是安全论据的一部分。此外，仿真只能基于已识别的参数[C.1 1)]或已识别的场景[C.1 2)]。

实际道路测试能够使用真实的输入来测试系统，但他受限于实际测试的公里数/小时数/场景数以及在实际测试期间遇到的实际场景快照的随机性[C.1 3)]。通过实际道路测试，可能会发现以前未知的参数。

类似功能及其相关潜在危害场景的先验知识可以用来裁剪确认策略，例如，从类似系统的历史现场数据中吸取的经验教训。还可以使用策略来减少所需的测试量，同时仍满足确认目标。

附录C的结构如下：

- 附录 C.2 讨论了使用危害行为比率来满足接受准则，并且给出了定义和评估接受准则与确认目标的示例；
- 附录 C.3 说明了如何使用统计数据和安全裕量。
- 附录 C.4 给出了如何在传感器验证和确认中使用的各种类型测试的示例；
- 附录 C.5 讨论了如何使用约束随机测试和重要性抽样来减少仿真测试的数量；及
- 附录 C.6 讨论了如何使用系统的物理架构来证明减少测试量的合理性。

C.2 确认目标的导出

C.2.1 使用危害行为比率来满足接受准则

接受准则通常非常小，例如为 $10^{-8}/h$ 。要确认这些非常低的取值，通常需要付出巨大努力。因此，找到一种既能降低确认目标，同时又证明达到了接受准则的方法就很重要。一种可能的方法是考虑相关危害行为的比率 R_{HB} 。

C.2.1的目的不是定义接受准则,而是从接受准则中得出危害行为的比率,进而使用该比率来定义确认目标。

第6章中识别和评估了由预期功能的危害行为引起的可能危害事件及其后果。每种已识别的危害行为都与第6章中定义的该行为的接受准则有关。每种危害行为的确认目标是由与该危害行为关联的接受准则导出的。

注1: C.2.1中不考虑导出接受准则的方法或者支持接受准则的论据。假设接受准则是由公认的且可接受的方法确定的一个比率。

符合已定义的接受准则的 R_{HB} 取值可以由以下步骤导出:

- 识别导致伤害H的事故/事件,这些事故/事件是由所分析的危害行为引起的(例如,由于非预期的制动而导致的追尾);
- 识别对于这些事故/事件的接受准则 A_H (其取值是由初始的接受准则结合安全裕量导出的);
- 识别潜在危害场景,在这些场景下,会发生已识别的事故,这些事故是由所考虑的危害行为导致的(例如,高速行驶中后车近距离跟随)。假设所考虑的危害行为在这种场景下已发生,则暴露在这种场景下的条件概率为 $P_{E/HB}$;

注2: 潜在危害场景包括危害行为的触发条件。

- 假设危害行为发生在已暴露的场景下,识别在此场景下这一危害行为的不可控的概率 $P_{C/E}$;
- 假设可控性行为未成功,识别出已识别的事故/事件 A_H 造成的严重度分布。该分布描述了在这些事故/事件中严重度的某一程度的发生概率 $P_{S/C}$ 。

注3: 根据所使用的接受准则, $P_{S/C}$ 可用于表示一定程度的严重度的可能性(例如,X%的涉事人员受重伤),也可用于表示严重度至少在一定程度以上的可能性(例如,Y%的涉事人员至少受轻伤)。

注4: 对于相似的危害事件,可以检查已明确的参数 $P_{E/HB}$, $P_{C/E}$ 和 $P_{S/C}$ 与GB/T 34590中危害识别与风险评估中的E,C和S参数值的一致性。GB/T 34590.3中关于暴露频率与持续时间的考虑也可以适用于SOTIF危害行为。

假设危害行为并不总是导致伤害,接受准则 A_H 可以按公式(C.1)分解为:

$$A_H = R_{HB} \times P_{E/HB} \times P_{C/E} \times P_{S/C} \dots\dots\dots (C.1)$$

危害行为比率 R_{HB} 是可以容忍的比率,即在给定时间内发生这种危害行为的概率。 R_{HB} 是直接由触发条件的发生率决定的,这些触发条件可以激活功能不足并导致危害行为。因此,它可用于导出适用的确认目标[公式(C.2)]:

$$R_{HB} = \frac{A_H}{P_{E/HB} \times P_{C/E} \times P_{S/C}} \dots\dots\dots (C.2)$$

注5: 在触发条件独立于暴露的危害行为导致伤害的场景的情况下,则条件概率可以简化为概率的简单乘积。

示例: 一个伤害H已经被识别到并关联到接受准则 $A_H = 10^{-8}/h$ 。从现场数据可知,导致这种伤害的危害行为中 $P_{C/E} = 10\%$ 是不可控的。 $P_{S/C} = 1\%$ 的伤害达到了接受准则指出的严重度。基于驾驶时间,用户处于危害行为会导致伤害的场景的概率为 $P_{E/HB} = 5\%$ 。基于这些取值,用于确认目标计算的危害行为比率如公式C.3给出的:

$$R_{HB} = \frac{A_H}{P_{E/HB} \times P_{C/E} \times P_{S/C}} = \frac{10^{-8}/h}{0.05 \times 0.1 \times 0.01} = 2 \times 10^{-4}/h \dots\dots\dots (C.3)$$

使用 $R_{HB} = 2 \times 10^{-4}/h$ 作为确定确认目标的新起点可以减少确认工作量。使用公式(C.7)和相关假设,如果在5000小时的测试中没有遇到危害行为,则可证明接受准则以63%的置信度被满足。

C.2.2 AEB系统中可接受误激活率的定义和确认的示例

C.2.2.1 目标

C. 2. 2提供了如何根据公布的交通事故统计数据计算SOTIF建议的最小确认距离(公里)的示例。长期车辆测试/车队测试被选作确认方法。使用统计方法和4步分析来计算目标里程。以下列出了步骤清单,并为每个步骤制定了其部分目标。

- a) 危害事件的可能原因(C. 2. 2. 2):
 - 对于目标系统,识别由功能不足引起的危害事件;
 - 明确发生危害事件的场景的已知参数以及这些参数的相关组合。
- b) 危害事件建模(C. 2. 2. 3):
 - 考虑激活系统功能不足的代表性参数;
 - 对危害事件(事故)的场景进行建模。
- c) 交通统计分析(C. 2. 2. 4):
 - 识别与上一步导出的场景相关的基本统计变量的分布;
 - 根据可用统计数据计算确认里程基准。
- d) 测试场景定义(C. 2. 2. 5):
 - 根据任务概要和所考虑的危害场景,选择用于确认目标应用的测试场景;
 - 对于这些场景,定义最小确认工作量。C. 2. 2. 5 以驾驶距离(公里)的形式界定了最小确认工作量。

注1: C. 2. 2与区域2和区域3都相关。SOTIF分析(第6章和第7章)和SOTIF验证假定在量产车辆部署之前执行。

注2: C. 2. 2基于文献[26]。

该评估应尽可能减少由于驾驶员对故障注入的心理预期而产生的影响,可以通过在评估过程中实施认知分心任务的方式来实现。

该评估应考虑目标市场驾驶员性别、年龄、地域等因素对测试结果的影响。

C. 2. 2. 2 危害事件的可能原因

对制动系统具有一定权限的车辆控制系统(例如AEB)可能会因错误的执行而使驾驶员或其他道路使用者处于危险之中。例如,由于物体识别功能不足引起的错误紧急制动激活会在不需要时迅速使车辆减速并将其完全停止。

根据本文件(见第4章,图2),识别和评估激发危害行为的触发条件,例如,由于非预期的AEB启动而导致后方车辆追尾。这里提到的性能局限可能是由多种外部因素触发的。

对于此示例,接受准则是由 AEB 功能引起的危害事件的可能性等于或小于由人类引起的相同危害事件的可能性,见公式(C. 4)。

$$P_{ha,AEB} \leq P_{ha,hu} \dots\dots\dots (C. 4)$$

式中:

$P_{ha,AEB}$ ——AEB 功能引起危害事件的概率;

$P_{ha,hu}$ ——人类引起危害事件的概率。

注: C. 2. 2. 2并未说明该接受准则是否足够合理可以向公众发布。

危害的可能性取决于场景,特别是取决于场景中对安全至关重要的参数的取值(例如触发条件)。例如,对于基于摄像头的系统,安全至关重要的参数是光照条件;对于基于雷达的系统,安全至关重要的参数是是否存在雷达光束反射材料等。然而,在区域3 (“未知危害场景”)中,影响安全的所有参数及其取值都未知。根据已知的依赖关系定义场景,并估计其风险。

C. 2. 2. 3 危害事件建模

C. 2. 2. 3~C. 2. 2. 5 的示例考虑某个系统，该系统能够按照图 C. 1 所示的减速曲线执行 AEB，并存在如下的潜在设计限制：

- AEB 系统可以控制制动以最大减速度 $9m/s^2$ 进行减速，以响应移动目标；
- 制动上升时间取决于制动系统预填充，限制为 $15m/s^3$ ；
- AEB 功能在速度大于 $5km/h$ 时可用；
- 允许最大降速为 $50km/h$ ；
- 传感器和制动系统中的安全机制将防止 AEB 控制减速超出设计的速度范围。

图 C. 1 显示在初始速度为 $50km/h$ （相当于 $13.9m/s$ ）时，AEB 减速导致的自车速度的理想变化曲线。

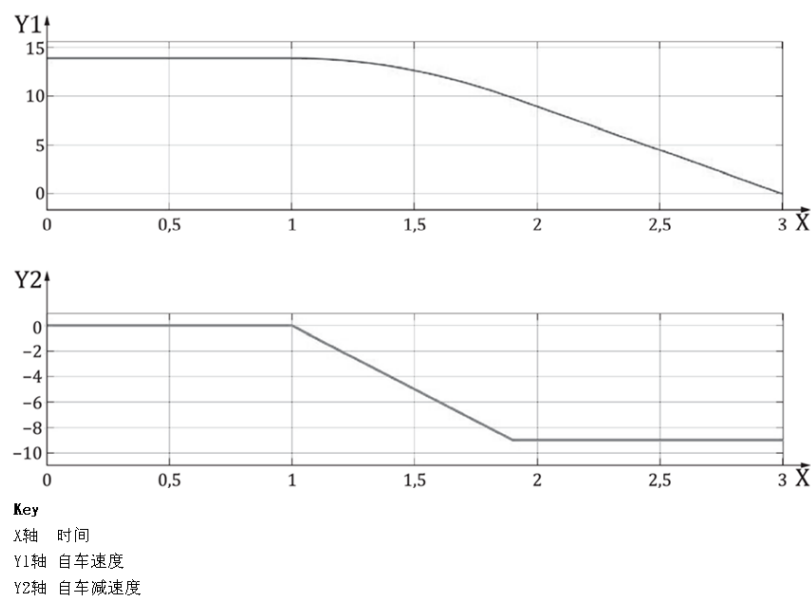


图 C. 1 AEB 减速曲线

预期功能安全相关的危害和危害场景如下：

- 危害行为：在设计意图范围内的非预期 AEB 制动超过 340ms。
- 危害场景：非预期的 AEB 制动时间超过 340 ms 且存在跟车距离较近的后车。在这种条件下，非预期制动会造成追尾碰撞。

危害事件可以被建模为在直路上跟车场景下的一阶效应(见图C. 2)^[26]。

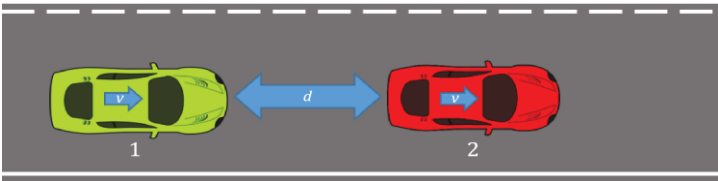


图 C. 2 危害事件模型中使用的跟车场景

该场景基于以下假设：

- 初始时，两辆车都以同样的速度行驶；
- 与速度相关的跟车距离具有已知的概率分布^{[26] [27] [28]}；
- 即使驾驶情形不需要，2 号车 AEB 激活紧急制动；

- 所有 AEB 制动事件遵循图 C.1 所示的制动曲线；
- 后方驾驶员感知到危害情形，并通过制动做出反应。反应时间服从已知的概率分布。

使用蒙特卡罗仿真对图C.2中所示的场景(“场景1”)进行分析,使用后车的跟车距离和反应时间作为输入变量,以估计危害事件(追尾)的概率。发现场景的结果在很大程度上取决于AEB非预期激活时车辆的速度。仿真以初始速度*v*作为输入,以碰撞概率作为输出。

图C.3表明,由于跟车距离较短,在较低的速度下碰撞的概率更高。由于跟车距离增加和最大降速阈值的存在,当速度在50km/h以上时,碰撞概率下降。在没有减速阈值的情况下,图C.3将有所不同(单调增加)。

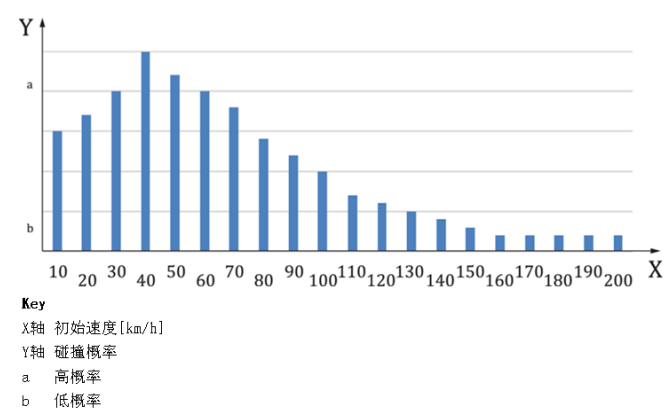


图 C.3 场景 1 中在不同初始速度下的追尾概率

C.2.2.4 交通统计分析

假设对于AEB来说,导致伤害的最常见事故是在跟车场景下两辆车之间的追尾碰撞(图C.2中的“场景1”)。通过分析,可以确定追尾碰撞的最大可容忍(可接受)发生率,即公式(C.4)中的 $P_{ha,hu}$ 。

国家道路安全管理部门提供的交通统计数据(例如,美国的NHTSA GES数据^[8],按照事故地张贴的限速进行分类)可以提供现场现有碰撞发生率的概况。

交通统计通常提供以下数据:

- 现场乘用车数量(*N*);
- 每辆乘用车每年平均行驶里程(*K*);
- 或者,可以提供每年行驶的车辆总里程数(*M*)。如果没有提供该参数,可以使用公式估计:
 $M = N \times K$;
- 每年现场相关事故(追尾)数量(*A*)。

对所考虑的变量采用统计模型,通过进一步分析,提高了估计的可信度。根据这些信息,可以计算人类驾驶员在两次碰撞之间的平均行驶里程(基准, *B*):

$$B = \frac{M}{A} \dots\dots\dots (C.5)$$

式中:

- B* ——人类驾驶员在两次碰撞之间的平均行驶里程(基准, *B*);
- M* ——每年行驶的车辆总里程数;
- A* ——每年现场相关事故(追尾)数量。

为了获得最坏情况估计, *M*取其上限值, *A*取其下限值。安全论据要求有证据证明配备AEB的车辆可以运行至少*B*公里而不会造成事故,或者AEB系统功能不足造成的事故概率低于每公里 $1/B$ [与公式(C.4)相比]。

注1: 上述准则只是一个理论概率量度,用于在决定向市场投放产品时,评估可以容忍的风险。因此,即使达到了此

确认目标，当实际市场中发生不期望的AEB时，判断是否有必要采取对策也需要基于（作为示例）系统架构、ODD和系统规范进行额外的分析和考虑。

注2：公式(C.5)中的基准可以被看做是系统确认的下界。考虑交通统计的不确定性，该基准可以通过乘以因子 k_1k_2 来增加或者减少。此时基准定义为： $B = k_1k_2(M/A)$ 。

示例1：将基准 B 乘以因子 $k_1 > 1$ 可用于保守地证明：与交通统计事故相比，AEB功能不会导致事故数量增加。

示例2：交通统计数据包括合理和不合理的制动事件。对于误报的AEB制动，只有导致危害事件（追尾碰撞）不合理制动才与定义基准有关。 k_2 被定义为危害事件的概率， $k_2 = 1/n$ 可被用于调整 n 个实际制动事件中只有一个由于不合理的制动而导致危害事件的情况。

注3：C.2.2.3中描述的仿真可用于估计由于不合理的制动 k_2 引起的危害事件的概率。

C.2.2.5 测试场景定义

如果以必要的置信度满足了接受准则，可能没有必要通过驾驶里程数等于或者超 B 来证明达到了可接受的残余风险水平。车辆任务概述（见表C.1）和系统行为数据可用于细化数据收集和确认策略。

仿真（见第C.2.2.3）表明，当速度为 50km/h 时，AEB导致的风险最高。我们将场景1（图C.2）分为三种：

——场景1.1： $v = 0 \sim 40\text{km/h}$ ；

——场景1.2： $v = 40 \sim 80\text{km/h}$ ；

——场景1.3： $v > 80\text{km/h}$ 。

表C.1使用公开数据分析了2010年至2017年美国发生的追尾事故的严重度概率分布^[26]。在这些数据中，碰撞概率和相关严重度级别按张贴的道路限速被划分为：

——城市道路（ $0 \sim 40\text{km/h}$ ）；

——乡村道路（ $40 \sim 100\text{km/h}$ 之间）；

——高速公路和州际公路（限速超过 100km/h ）。

比较图C.3碰撞概率最高的区域与表C.1中严重度的分布，我们发现人类和AEB系统引起的追尾碰撞的区域是一致的。最高风险区域对应于场景1.2。

注：在超过 80 km/h 的速度下，潜在的AEB激活违反了系统的限制。例如，这可以通过GB/T 34590标准中建议的外部措施来实施，因此被视为不在C.2.2的范围内。

表 C.1 不同限度下追尾事故严重度风险概率分布

限速 (km/h)	0~40	40~80	80~100	>100	全部速度
追尾事故(包含车尾对车尾)占比	9.4 %	69.9 %	12.8 %	7.9 %	100.0 %
未受伤	80.0 %	73.3 %	74.6 %	72.9 %	74.1 %
非致残性伤害	18.9 %	24.7 %	22.7 %	25.0 %	24.0 %
致残性伤害	1.1 %	1.8 %	2.3 %	1.6 %	1.8 %
死亡	0.055 %	0.52 %	0.33 %	0.55 %	0.13 %

假设有统计数据，则可以为场景1.2重新计算基准（公式C.6）：

$$B_{40..80} = \frac{M_{40..80}}{A_{40..80}} \dots\dots\dots (C.6)$$

$B_{40..80}$ 限速为 $40 \sim 80\text{km/h}$ 时，人类驾驶员相邻两次碰撞之间的平均里程（基准， B ）；

$M_{40..80}$ 限速为 $40 \sim 80\text{km/h}$ 时，每年行驶的车辆总里程数；

$A_{40..80}$ 限速为 $40 \sim 80\text{km/h}$ 时，每年现场相关事故（追尾）数量。

对于对风险影响未知的参数，数据收集可包括各种驾驶条件，例如：

- 天气状况：AEB 系统可在一组代表性天气状况下进行测试。这包括干燥、雾天、雪天、雨天、阴天等；
- 一天中的时间：根据传感器的类型，数据收集可以包括一天中的不同时间，如夜晚、黄昏等。

此外，数据收集还可以包括从传感器局限和特性特有局限分析中得出的相关驾驶情况。

表C.2给出了车辆任务概述的示例。该规范基于天气、速度和其他参数的真实状况。它还可以基于涵盖场景发生率的数据，其中场景发生率可以通过仿真或者估计得到。

表 C.2 车辆任务概述示例

一天中的时间	
类型	占比
白天	50%
夜晚	35%
黄昏	15%
车速	
速度[km/h]	占比
0..50	60%
50..80	40%
>80	0%
天气状况	
类型	占比
干燥/晴天	65%
雨天	7%
雾天	5%
雪天	5%
阴天	10%
大雨	5%
其他天气状况	3%

C.2.2.6 对基准的考虑

可以使用C.2.2所述的基于交通统计的方法定义目标平均碰撞间隔时间(MTBC)基准，该基准可用于在量产或现场运行之前确认自动驾驶系统的鲁棒性。然而，这种方法的主要考虑因素是：

- 可扩展性：除非对系统架构作出具体考虑，否则将此方法应用于完全自动驾驶车辆可能是不可行的。对于 C.2.2 中的 AEB 示例，将特性适用速度范围扩展到高速公路速度(例如：130 km/h)，将增加基准确认里程，因为在高速下追尾频率较低；及
- 系统架构独立性：可使用系统架构的考虑因素来优化目标确认里程。对于复杂的特性，其中使用超过一个子系统来冗余验证特定的控制行为，则可以通过观察影响整车 MTBC 的单个度量(例如，基于摄像头的目标探测子系统误报率或基于雷达的目标探测子系统误报率)来优化从交通统计导出的 MTBC；
- 对确认路线的依赖：基于系统局限分析设计的特定驾驶路线可以更准确的定义 MTBC，从而减少所需收集的数据量。

C.3 预期功能安全适用系统的确认

图C.4阐述了一种可能的模型，说明如何结合覆盖目标值和约束的随机测试来进行V&V迭代，用于发现未知危害场景或功能不足(即减少区域3)，以支持SOTIF开发(图11)。在V&V启动之前的初始状态(最左边的圆圈)中，安全分析过程中已经确定了一些潜在的功能不足，即灰圈所代表的区域2。除此以外，可能依然存在一些在此阶段未能发现的功能不足(黑圈部分，未知的危害场景(区域3))。虚线方框表征了全部功能中已使用的功能(例如在ODD中使用的功能)。

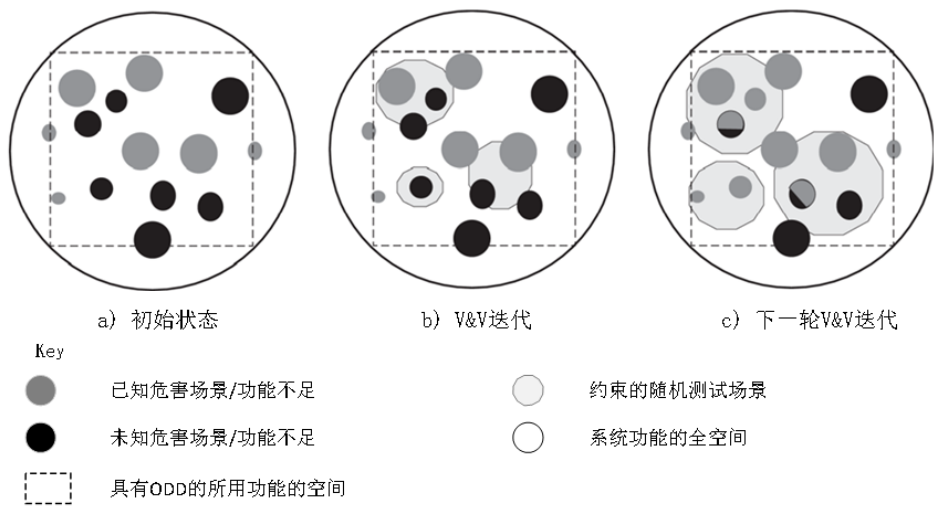


图 C.4 SOTIF 开发测试迭代

总体V&V目标是在给定的ODD边界内最小化未知危害场景的发生。一种方法是使用已知场景作为约束随机生成新场景测试的基础，从而逐步增加测试覆盖空间。这些新场景/测试增加了覆盖的测试空间，因而能够暴露更多未知的危害场景[图 C.4 b)]。

下一轮 V&V 迭代建立在前一轮的基础上。通过扩展已覆盖的随机空间，那些暴露的未知场景现在变为已知，并作为进一步扩展覆盖度的基础。先前已知的场景也可以被用于创建更多随机测试和场景。

这个迭代过程持续进行直到充分地覆盖已用功能空间。迭代的结果是区域3中发现的危害场景转变到区域2中[(图C.4 c)]。一些没有覆盖到的危害场景可以通过减少ODD来减少。

图C.4的模型也可以应用于SOTIF适用系统的典型车辆软件开发策略里。伴随着软件的测试活动和潜在的危害行为被消除，潜在危害行为之间的平均里程数预计会上升。但是随着新特性/功能被引入或使能，每个潜在危害行为之间的平均时长或里程或许会下降，然后由于引入新特征/功能所带来的漏洞被追踪并解决，这些数值会而再次上升。最终达到特定用例和功能的确认目标阈值，并认为确认活动得到满足。这一概念如图C.5所示。

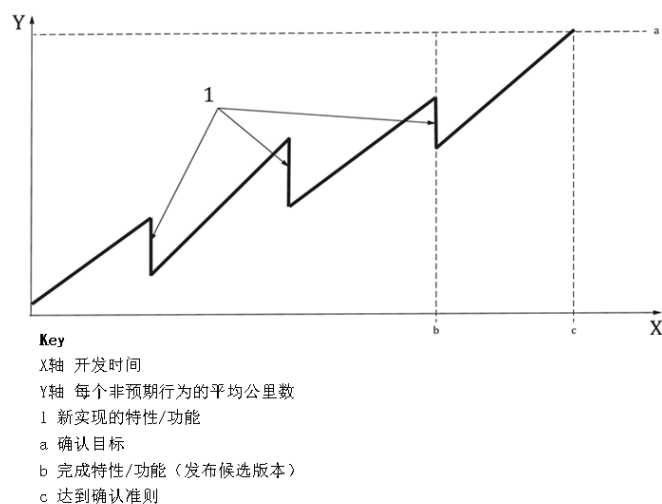


图 C.5 开发期间潜在危害行为率的预期概况

例如，在测试之前，系统所有者需定义如下内容：

- 确认目标(停止规则)；
- 测试工作在不同测试模式(实车测试、硬件在环、软件在环等)上的分配；
- 定义潜在危害行为、定义重启距离计数的标准。

SOTIF适用系统的确认过程从选择接受准则开始(见6.5)。从这个接受标准推导出确认目标。确认目标可以根据系统用例(例如辅助驻车、自动紧急制动、车道保持、自动平行停车、低速自动停车场辅助、高速公路自动驾驶辅助、自动出租车)的碰撞统计数据和安全裕量来计算。

以下可以用于形成目标：

——要使用的统计数据；

示例1：已报告的碰撞事故。

——统计数据中的人为表现；

示例2：NHTSA 2015年碰撞事故统计为每500000英里1次^[29]。

——安全裕量；

——统计的置信度限值。

示例3：对于特定用例，人类驾驶员在事故之间平均行驶距离为B公里。出于安全原因，定义了额外的裕量 $y > 1$ 。所选的SOTIF适用系统的接受准则是潜在危害行为之间的平均公里数 $B \times y$ 或目标事故率 $A_H = 1/(B \times y)$ 。终止规则假设事故服从泊松分布。使用确认目标 τ ，如果在连续的 τ 驾驶时间或驾驶里程内，没有发生潜在危害行为，则可以证明系统的故障率小于或者等于 A_H ，且置信度为 α ，其中 τ 在公式(C.7)中给出^[31]：

$$\tau = -\ln(1 - \alpha)/A_H \quad \text{..... (C. 7)}$$

注1： τ 可以是时间单位或距离单位，具体取决于事故率的单位。

注2：对于 $\alpha \approx 0.63$ ， $\tau = \frac{1}{A_H} = B \times y$ 。

注3：分布可以随时间而改变。例如，对于现有ADAS系统如AEB，有必要通过比较某一系统广泛引入前后的事故率来控制其统计分布。

实际上，需要通过行驶进行确认的公里数或者小时数 τ 可能会非常大，因此在某些情况下不实用。通过类似的系统、MIL、SIL及HIL模拟行驶公里数，使用专家知识，可以降低实际行驶要求。基于仿真能力，可以将测试例在现实世界测试和仿真测试之间进行适当分配(例如，仿真仅在特定场景下有效)。

合理的改变现实和仿真确认测试的条件(例如,不同的天气条件、一天中所处的时段、道路条件、交通条件、行人条件等),以试图发现罕见的实际行驶工况。

C.4 感知系统的验证和确认

C.4.1 感知系统的验证和确认框架

C.4.1.1 总则

C.4.1提供了一个示例方法,可用于增量地验证和确认给定感知系统的性能。感知系统在任何驾驶自动化等级的自动驾驶汽车的SOTIF中都发挥着重要作用。该示例方法可适用于ADS搭载车辆中使用的任何类型的感知技术(例如雷达、摄像头、激光雷达、超声波等)。

在任何开发阶段可能引入的不同类型问题,这些问题都会影响感知系统的性能。因此,感知系统经历如图C.6中描述的增量验证和确认过程是有价值的。

注1:这些步骤的顺序是以递增形式呈现,但在执行这些步骤时不强制要求顺序。

注2:这些步骤可以跨多个公司共享(见 4.4.2)。



图 C.6 感知验证和确认的示例步骤

感知系统验证和确认过程可以包括多个步骤:

- 台架验证 (BV): 在受控环境下对感知系统探测能力进行初步验证;
- 算法性能验证 (APV): 使用更大规模的数据验证感知系统性能;
- 整车集成验证 (VIV): 在目标车辆中集成后,验证感知系统性能;
- 封闭道路验证 (TTV): 在封闭道路上针对多个参考用例验证感知系统性能;
- 开放道路确认 (ORV): 在开放道路中针对所有相关场景确认感知系统性能。

C.4.1.2~C.4.1.6 展示了使用 SIPOC(供应商、输入、过程、输出、客户)图表的分析示例。SIPOC 是一种以表格形式汇总一个或多个过程的输入和输出的工具,用于定义一个从开始到结束的过程[32]。SIPOC 是一种用于质量管理和过程改进的分析方法,但是其他方法也可用于感知系统验证和确认过程的分析。

C.4.1.2 台架验证

可以定义台架验证活动,以验证组装的感知系统在参考环境下的探测能力(台架测试)。该测试有助于验证感知系统在受控环境中针对特定生产误差的鲁棒性(例如,不同的可容忍的雷达天线灵敏度或不同的摄像头焦距)。表 C.3 提供了这些类型测试的示例。

表 C.3 台架验证

种类	供应商(S)	输入(I)	处理(P)	输出(O)	客户(C)
定义	工程	探测要求: (示例:分辨和分离能力,精确度)	根据产品规范,在受控环境下验证感知系统的探测性能。	验证通过: 在受控环境下性能验证通过的感知系统。	工程团队(进一步测试) OEM/TierX 供应商

种类	供应商(S)	输入(I)	处理(P)	输出(O)	客户(C)
	制造	组装的系统 (SMV之后)		验证失败： 报废的感知系统(返工或处理)	
示例1	工程	雷达探测要求(KPI)	在消声室中使用雷达 目标发生器验证雷达	验证通过： 对参考数据具有验证通过的探测	工程团队(进一步测试)
	制造	组装的雷达 (SMV之后)	正确的探测能力。	能力的雷达 验证失败： 报废的雷达(返工或处理)	OEM/TierX供应商
示例2	/工程	摄像头探测要求(KPI)	在屏幕前播放录制的	验证通过： 对参考数据具有验证通过的探测	工程团队(进一步测试)
	制造	组装的摄像头 (SMV之后)	数据或合成片段以验证 正确的探测能力。	能力的摄像头 验证失败： 报废的摄像头(返工或处理)	OEM/TierX供应商

C. 4. 1. 3 算法性能验证

可以定义算法性能验证活动来验证感知系统算法对一组参考数据的探测能力(例如重用仿真或先前收集的数据)。这种测试有助于验证在相同硬件上新发布的软件版本上没有性能回退：

- 代码不同阶段暴露的系统行为不足和可能的功能不足；
- 在重复的过程中获得更好的鲁棒性；
- 避免问题在开发过程中再次出现；
- 为分析根本原因提供稳定的基线。

算法的性能验证步骤可以在目标硬件上执行(例HIL测试)，也可以通过注入以前记录或合成的数据在模拟器上执行(例如SIL测试)。由于这两种方法的差异，表C. 4没有提供将这一验证步骤应用于不同感知系统的示例。C. 4. 4描述了一种可用于算法性能验证的技术。

表 C. 4 算法性能验证

种类	供应商(S)	输入(I)	处理(P)	输出(O)	客户(C)
定义	工程	参考数据(预采集数据或模拟数据)	针对一组参考数据 (数据注入或仿真) 验证正确的算法性能。	验证通过：已验证	工程团队(进一步测试) OEM/TierX供应商
		探测要求/KPI		通过的感知系统算法	
	生产制造	算法和仿真软件(软件在环的情况下) 组装的系统(硬件在环的情况下)		验证失败：修改或 重新设计感知系统 算法	

C. 4. 1. 4 整车集成验证

可以定义整车集成验证活动，以验证感知系统能够在目标车辆上运行，并且没有非预期的性能降级/改变，这个验证步骤有助于更好地理解以下内容：

- 感知系统能够使用目标车辆提供的信息(车载信号比如车辆动态信号等)；及

——在与目标实现相关的规范定义不足时(例如,与摄像头相关的挡风玻璃反射率,雷达集成在保险杠后时保险杠喷涂的类型和厚度,或位于雷达前的不正确的介电材料),感知系统可以运行且无性能降级。

表C.5展示了整车集成验证的示例。

表 C.5 整车集成验证

种类	供应商(S)	输入(I)	处理(P)	输出(O)	客户(C)
定义	工程	车辆性能规范	验证感知系统在目标车辆中使用 ——按照规范工作	验证通过: 整车集成验证通过的感知系统	工程团队(进一步测试) OEM/TierX供应商
	制造	组装的感知系统 车辆(目标环境)		验证失败 1: 修改或重新设计感知系统 验证失败 2: 修改或重新设计感知系统	
示例 1	工程	车辆通信协议	验证感知系统能够使用车载信号: ——以适当的延迟接收车辆动态信息; ——电信号在规范范围内。	验证成功: 车辆上集成验证通过的感知系统。 验证失败 1: 修改或重新设计感知系统; 验证失败 2: 修改或重新设计感知系统或车辆接口。	工程团队(进一步测试) OEM/TierX供应商
	制造	组装的感知系统 车辆(目标环境)			
示例 2	工程	雷达的预期降级	对雷达系统的降级进行测试: ——不正确的定义保险杠形状/曲率导致雷达性能降级(雷达在保险杠或标志后面) ——不正确的喷涂规格导致性能降级雷达前的保险杠或标志的油漆厚度或类型不正确 ——不正确的介电特性导致性能降级(不正确定义保险杠材料,不正确的标志设计……)	验证成功: 集成雷达到汽车保险杠后。 验证失败 1: 修改或重新设计的感知系统; 验证失败 2: 修改或重新设计感知系统或汽车保险杠。	工程团队(进一步测试) OEM/TierX供应商
	制造	组装的感知系统 车辆(目标环境)/ 车辆的一部分(目标设计)			
示例3	工程	摄像头的预期降级	对摄像头系统的降级进行测试: ——摄像头集成在挡风玻璃后 验证摄像头-支架-挡风玻璃组装。	验证通过: 摄像头集成到挡风玻璃后。 验证失败 1: 报废感知系统(用于返工或处理) 修改或重新设计感知系统; 验证失败 2: 修改或重新设计摄像头布置。	工程团队(进一步测试) OEM/TierX 供应商
	制造	已安装的感知系统 车辆(目标环境)/ 车辆的一部分(目标设计)			

C.4.1.5 封闭道路验证

可以定义封闭道路验证活动，以针对一组特定的参考用例(场景，包括特定的触发条件)来验证感知系统的探测能力。虽然用例(场景)本身通常是“技术不可知”(不依赖于感知系统的特性)，但是可以选择或优先考虑一组特定于技术的用例(场景，包括特定的触发条件)来验证以下几个方面：

- 特定用例的感知系统性能(如车辆安全性能评估项目如:Euro NCAP, JNCAP, NHTSA, KNCAP, C-NCAP, Latin NCAP 协定的那些测试场景或相似的测试场景)；
- 在特定的场景下的感知系统的验证，旨在发掘感知系统的局限性(以雷达角精度为例)；
- 自车传感器与其他自车传感器或其他车辆传感器之间的相互作用(例如雷达相互干扰)。

表C.6描述了一个封闭道路验证的示例。

表 C.6 封闭道路验证

种类	供应商(S)	输入(I)	处理(P)	输出(O)	客户(C)
定义	工程	用例列表 感知系统已知的性能 局限	在与终端功能相关的特定用例 中验证感知系统性能	验证通过：验证通过的感知系统性能 验证失败：修改或重新设计感知系统	工程团队(进一步测试) OEM/TierX供应商
	制造	组装的(车内)感知系统(VIV后)			
示例 1	工程	用例列表 感知系统已知的性能 局限	验证感知系统在给定时间内区分行人和停放车辆的能力，该测试是车辆安全性能评估计划提出的AEB Euro NCAP 被遮蔽的弱势道路使用者场景的一部分	验证通过：已验证的感知系统性能 验证失败：修改或重新设计感知系	工程团队(进一步测试) OEM/TierX供应商
	制造	组装(车内)的感知系统(VIV后)			
示例 2	工程	基于雷达的感知系统 干扰频率			工程团队(进一步测试) OEM/TierX供应商
	制造	组装的(车内)感知系统(VIV后)			

C.4.1.6 开放道路确认

可以定义开放道路确认活动，以确认感知系统在目标环境中的性能。确认阶段的目标可以包括：

- 在多个市场、各种环境条件中持续收集典型数据；
- 在通常很少见，并且在正常驾驶中较少出现但可影响感知的情况下收集的特定数据，例如：
 - 视觉感知：黄昏或黎明时的数据；
 - 雷达感知：雨和雨滴飞溅条件，撒盐道路；
 - 激光雷达感知：不利的天气条件；
 - 所有感知：隧道入口/出口；
- 在不常见但可能增加危害行为的场景中收集特定数据，如：
 - 车辆行驶在车流量很小且无前车的道路上，会增加道路内目标选择错误和探测到鬼影目标的可能性；
 - 超越一排卡车车队时，车队的阴影遮盖了超车道；及
 - 扫雪机经过时溅起的雪，可能会导致一个或多个感知系统突然致盲；

- 基于系统局限的特定数据收集，如：
 - 技术局限性(如，雷达无法正确识别钢架桥)；及
 - 功能/算法局限性(如，没有交通车辆时的灯光控制)；
- 不同驾驶习惯；
- 在不利条件下的专项测试，比如：
 - 天气；
 - 基础设施质量；
 - 交通情况(混乱相对于有序)；
 - 驾驶动态特性(横向和纵向)；
 - 道路周边的复杂情况(存在多处光源或复杂的道路设施)；及
 - 交通状况(弱势道路使用者较多的路况，相对于高速公路)。

开放道路确认的示例见表C. 7。

表 C. 7 开放道路确认

种类	供应商(S)	输入(I)	处理(P)	输出(O)	客户(C)
定义	工程	用例列表 感知系统已知的性能 局限(在 TTV 或 APV 后或在多次 TTV 或 APV 后不断更新)	根据目标市场，目标功能以及 感知局限，确认感知系统 在 目标用例的性能。	确认通过：在所有相关条 件下确认感知系统性能 确认失败：修改或重新设 计感知系统	工程团队(进一步测 试) OEM/TierX供应商
	制造	组装的(车内)感知系 统(VIV 后)			

C. 4. 2 随机传感器模型

复杂的驾驶自动化系统可能需要大量的无法在实际环境中实现的测试。作为实际环境测试的补充，在虚拟环境中的仿真可以作为测试活动的一个重要部分。因为现代的传感器比较复杂，容易受到复杂的，通常是随机现象的影响，所以传感器的仿真是测试活动非常重要的环节之一。

基于物理学的详细传感器模型，需要大量的建模工作和巨大的算力。而随机传感器模型具有以下优点：

- 无需了解传感器实现的每一个细节；
- 易于应用蒙特卡罗方法测试不同的参数和情况；
- 仅需中/低算力。

该方法可以基于参数或非参数法：参数统计是统计学的一个分支，它假设样本数据来自服从一组固定参数集的概率分布的总体。最著名的基本统计方法都是参数统计方法。非参数模型的不同之处在于，参数集不是固定的，如果收集到新的相关信息，参数集可能会增加或者减少。由于参数模型依赖于一个固定的参数集，它比非参数方法对给定的总体作出更多的假设。当假设正确时，参数方法将得到比非参数方法更准确和精确的估计，即具备更强的统计功效。然而，当假设不正确时，参数方法失败的可能性更大，因此该方法也不是非常稳健的统计方法。

对于参数化方法，传感器模型通常反映传感器的功能结构：

- 将传感器分解为若干功能模块；
- 每个模块负责对探测/测量过程的特定效果进行建模；
- 每个模块独立建模；
- 每个模块都有一组可配置的参数；

——模拟器的输出是所有步骤模型的组合。

另者，非参数方法没有对传感器内部结构进行详细的建模，而是将其建模为一个黑盒子，因此他侧重于传感结果的统计表示^[33]。

对于摄像头传感器模型，用来估计传感器参数的统计试验的典型功能架构如图C.7所示。输入是一个来自真实世界的数据库。将该输入并行输入摄像头测试台架和随机传感器模型。对比模型的响应和摄像头测试台架的响应。用一个关键的绩效指标激励模型，并用优化函数更新模型参数，直到差值最小。

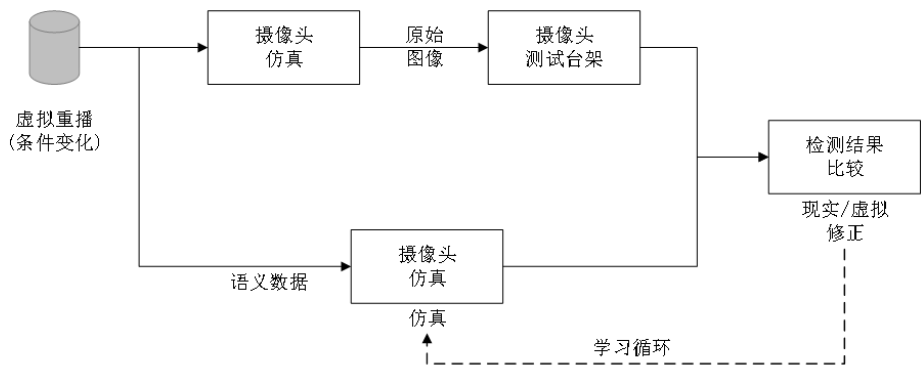


图 C.7 架构示例-摄像头传感器模型校正

在传感器模型校准后，该模型采用尚未用于仿真模型校准或训练的数据进行确认。经过仿真模型确认后，模型既可以分析独立的传感器，也可以分析车辆仿真框架中的传感器。每次收集到更多真实数据后，就可以进一步改进模型中的相关参数。

C.5 场景参数化及场景抽样指导

C.5为仿真及基于场景的验证和确认提供资料性指导，以支持第10章和第11章的目标。

仿真测试是确认工作的重要部分。在确保仿真是对系统和环境的准确表达之后，预先记录或构建的场景就可以用来对已知场景下的系统进行确认。

基于记录的场景和仿真生成的新测试用例也可以用于测试未知场景。图C.8描述了一个基于构建随机测试以生成新测试用例的示例。该示例通过改变标称场景中的一个或多个方面来生成新的测试用例。

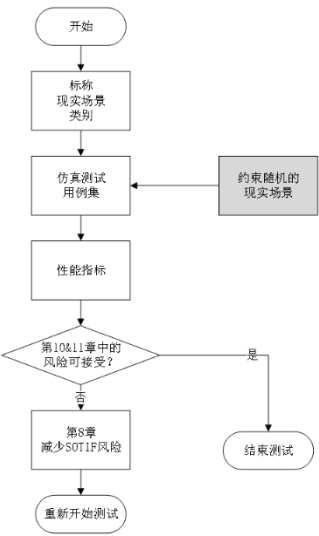


图 C.8 基于约束随机的 SOTIF 测试

这种类型的测试可以是随机的，也可以是结构化的(例如参数取值按照固定步长增加)，或者可以是两者的组合，例如多场景顺序组合或多子场景并行组合。

如果已知变量参数的分布，则可以进一步细化仿真过程。通常，可以基于自然驾驶数据来确定参数分布。下面给出一个来自参考文献[34]的示例来说明。自车在直道上以规定距离跟随前车。ACC负责自车的纵向控制，假设两辆车沿同一方向直行。前车制动，则自车也会制动，以防止碰撞。图C. 9显示了该场景的示意图，其中轿车代表自车。

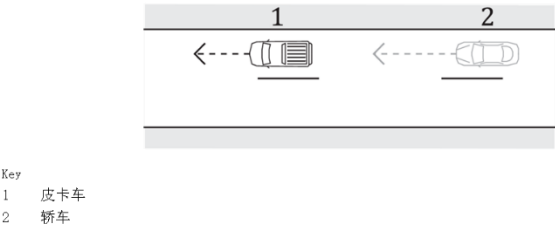


图 C. 9 交通场景示意图

如图C. 10所示，可以使用三个参数对该场景进行参数化：

- v_{end} ：制动后的速度；
- t_{brake} ：达到速度 v_{end} 之前制动所需的总时间；
- Δv ：前车的速度减少总量。

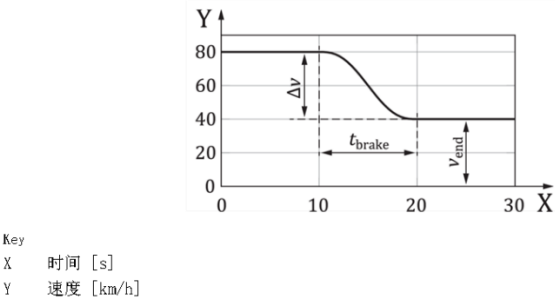


图 C. 10 场景中前车的制动曲线

所得联合分布的边缘概率分布如图C. 11中粗线所示。在这种情况下，核密度估计(KDE)可以用来估计基础分布，当然也可以使用其他技术。直方图表示原始数据，粗线表示参数的KDE的边缘概率分布。

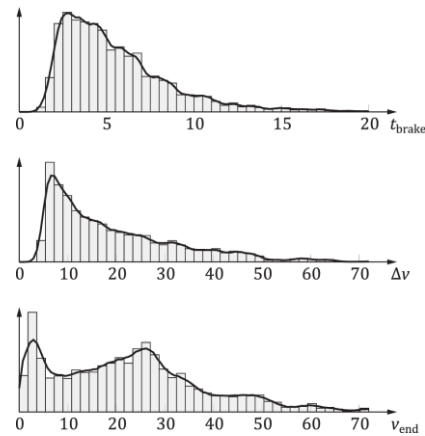


图 C. 11 测试场景的三参数

一旦从真实世界场景信息中获得了参数的分布，则图C. 8所示过程可以被细化为图C. 11。

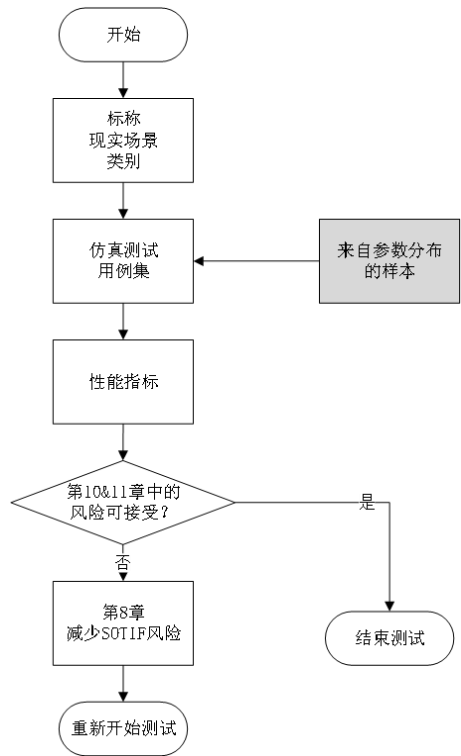


图 C. 12 基于参数分布的场景测试

使用估计的参数分布来生成测试场景可能会导致生成许多低价值的测试用例，因为高价值的测试更可能落入分布的特定区域。为了避免低价值的测试带来不必要的计算负担，测试参数可能会更偏向于在这些特定区域频繁抽样，称为重要性抽样[34]。图C. 13使用图C. 8中的示例，其中，当前车具有较大的速度减少量时更具有风险，因此抽样偏向较大值。

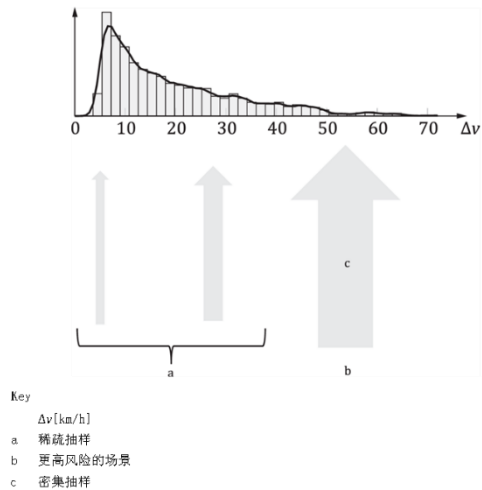


图 C. 13 重要性抽样示例

图C. 12的过程现在可以进一步增强为图C. 14。

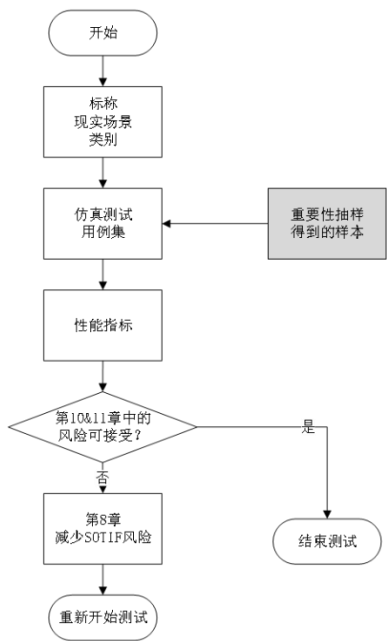


图 C. 14 基于使用重要性抽样参数分布的场景测试

总之，测试开始于为被测功能选择相关的场景类别。根据测试类型 (约束随机、分布抽样或重要性抽样)，将生成一组用于仿真的测试用例。根据适用于所选场景类别的相关度量来判断测试结果。根据测试阶段和被测试功能的相关风险，选择相关的测试用例生成类型。例如，重要性抽样可以遵循分布抽样方法，以增加可信性。最后，根据仿真结果和可接受准则来评估风险，也可采用第8章中方法以降低风险。图C. 15给出了结合所有三种测试类型的流程图。

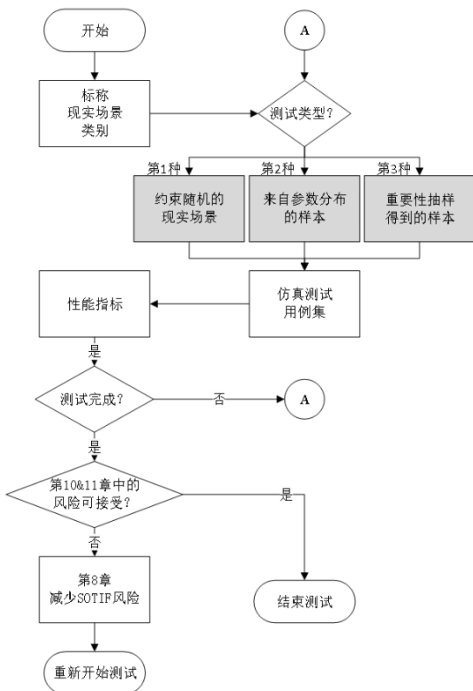


图 C. 15 基于仿真流程的场景示例

驾驶自动化系统的确认涉及大量的仿真场景，以尽可能多地覆盖真实世界的情景。这些仿真可能需要大量道路特性和驾驶场景的数据。收集和/或构建这些数据可能是一项巨大的工作。在公司内部或公司之间使用存储和交换此类情景数据的标准，有助于系统及全面地探索各种情景。

还可以使用标准支持各种工具或工具组件的集成。如FMI^[35]等标准使得将不同仿真元素组合起来成为了可能。

仿真相关标准包括但不限于：

- OpenDRIVE^[36]：道路网络的逻辑描述；
- OpenCRG^[37]：道路表面描述；
- OpenSCENARIO^[38]：驾驶场景描述；
- Open Simulation Interface (OSI)^[39]：传感器数据的连接；
- NDS^[40]：高精度地图数据；
- CityGML^[41]：城市三维模型；
- FMI^[35]：动态模型的模型交互和联合仿真。

C.6 减少确认测试的考虑

C.6.1 评估测试场景的覆盖率

通过正确定义测试计划(包含仿真，实际道路或两者结合)可以降低证明满足确认目标所需的测试量。

如果可以根据一个变量的不同模态划分客户使用场景集，掌握用户的使用情况就可以提供对变量每种模态的概率估计。

如果额外的论证(如来自仿真、专家判断)能够表明系统在特定模态下处理特定场景的能力比其他模态(例如白天和夜晚对比)高得多，则该论证可证明将确认工作集中在最严重的模态上并减少在不太严重的模态上是合理的。

注：测试场景的覆盖率取决于很多的定量确认目标类型。除了 $P_{\text{伤害}} < \sigma$ 类型的目标(通过具有代表性的单位时间或单位距离内的驾驶计算危害， σ 是很小的正数)，其他类型的量化确认目标可能会包含其他方面，如公平性(特定人群中的个体不会受到更多的伤害)。例如，为了表明自动化车辆不会将基于某些特征的特定群体置于比其他群体更高的被伤害的风险中，测试用例的设计应当包含这些特定群体。

C.6.2 与定量目标相关的组件级别的充分条件

在使用模块化设计时，关注与整车层面确认目标相关的组件层面的充分条件是有益的。组件的充分条件是这样定义的，如果他满足，那么，在设定的伤害概率的置信度水平下，确认目标也就达到了(假设系统其余部分的功能已知)，如 $P_{\text{伤害}} < \sigma$ 充分条件可通过(保守地)假设世界和系统其余部分比实际表现更差(不会更好)获得。

可以从对设计运行范围的详细认知中获得有用的充分条件，包括它的概率方面，以及具有各个组件及其依赖关系的系统架构。为了证明在组件层面满足充分条件，考虑到有关问题和相关系统方面的现有知识(例如车辆动力学、传感技术的物理特性)，可以通过利用实际可管理的数据来进行统计证明。关注组件层面的充分条件不仅可以降低确认成本，还可以在设计运行范围或系统组件发生变化时复用大部分安全分析。该方法的另一个优点是，通过在组件层面满足充分条件就可以更实际的实现论证，满足更细化的定量目标，例如考虑公平性。

C.6.3 系统架构的考虑

C.6.3.1 总则.

11.3讨论了为表11中描述的每种应用方法选择合适的累积测试里程。给定一个整体系统确认度量，恰当地定义系统架构可以减少所需的测试里程。

C.6.3.2 示例：使用满足充分条件的统计模块化安全论据

接受准则通常在数学上表示为 $P_{\text{伤害}} < \sigma$ 或 $E[\text{伤害}] < \sigma$ ，其中伤害是通过对单位时间或单位距离的代表性驾驶活动的计算得出的， σ 为较小的正数值。任何涉及统计因素的论证都会包含来自随机抽样不可避免的不确定性。量化不确定性有助于支持这一论证。

虽然接受准则是在整车级别定义的，但是安全论证的一个非常有用的特性是支持模块化设计，这意味着组件级的分析和设计运行范围分析可以组合成最终的整车级别安全论证。模块化安全论证的潜在好处是降低了确认成本。与整车级随机道路测试相比，单独的组件级和设计运行范围分析成本更低且可重复使用。为此，C.6.3.2提供了一个具有以下两个合理特征的安全论证示例；论证是：

- 结构化的，利用模块化组件级分析；
- 对量化达到接受准则不确定性的统计具有严谨性。

这个简单的示例旨在传达主要思想并作为具体说明，并不力求在使用真实系统时要求的那种真实性和细节。

系统和设计运行范围描述：

- 一辆自动驾驶汽车被设计为在只存在静止物体的直路上，以恒定速度 v 行驶。与该系统相关的环境条件(例如照明、降水、路面摩擦等)是固定不变的。
- 该车配备了自动制动功能，其规范如下：
 - 目标检测和深度估计的组合算法以固定频率提供最近物体的距离估计；及
 - 如果最近物体的距离估计值低于阈值 c ，则汽车开始制动，直到完全停止；实际制动距离是一个常数 $b < c$ 。
- 在每一次制动事件后，一旦移除静止物体，车辆就会重新安全启动(重新加速)，以速度 v 驶向他至少 c 的下一个静止物体。保持问题中车速恒定的假设，可以避免额外的复杂数学运算。
- 传感器和算法在测试和实际使用期间是固定的。

该车辆的设计目标是沿直路行驶并避免与任何物体相撞。或者该系统可以视作为更复杂的自动驾驶系统的一部分，具有障碍物检测和自动制动功能。

根据单位行驶距离的预期碰撞次数定义接受准则

$$E[\text{单位距离的碰撞次数}] < \sigma$$

其中 $\sigma > 0$ 是预先指定的目标水平。

概念如图 C.16所示，并使用以下标记法：

- b 是到完全停止的总制动距离；
- $c - b$ 是车辆和物体之间的纵向缓冲距离；
- m 是当车辆穿过车辆和物体之间的缓冲间隔 $c - b$ 时，感知系统以固定的频率产生的距离估计的数目；
- l 表示车辆在感知算法进行连续两次感知之间以速度 v 行进的距离；
- L 代表车辆到障碍物的实际距离， D 是感知算法估计的车辆到障碍物的距离；
- D^m 是物体感知系统，在 $b \leq L < b + 1$ 即刹车前最后一次感知时，确定的车辆到障碍物的估计距离；

—— V^h 是车辆碰撞物体的速度，若车辆停止并且恰好没有撞到物体，则 $V^h = 0$ 。

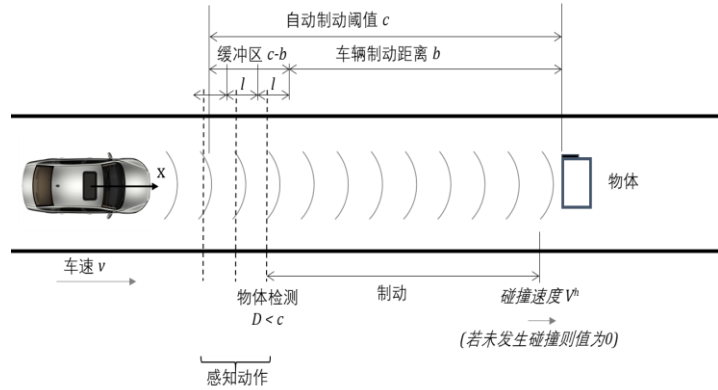


图 C.16 车辆运行示例

在此示例中，可以通过分解组件和 ODD 特性来获得预期的碰撞次数：

$$E[\text{单位距离的碰撞次数}] = P(V^h > 0) E[\text{单位距离物体数}],$$

$P(V^h > 0)$ 是，假定自车前方存在随机障碍物，且由于未能感知到障碍物，在离障碍物太近而无法避免碰撞之前，车辆没有制动的概率。在以上等式中出现的量是现实世界中物体分布的概率量。

存在障碍物时，碰撞概率 $P(V^h > 0)$ ，当 $b \leq L < b + 1$ 时，等于自车在距离障碍物过近之前没有制动的概率，也等于没有检测到障碍物（即 $D > c$ ）的概率。
因此，考虑到与目标的一次接触，

$$P(V^h > 0) = P(D > c \text{ 对所有 } L \in [b, c] \leq P(D^m > c))$$

因此，在这个特定的示例中，基于上面的定量安全措施可以界定为：

$$E[\text{单位距离的碰撞次数}] \leq P(D^m > c) E[\text{单位距离物体数}]$$

因此实现特定的检测性能是：

$$P(D^m > c) \leq \sigma / E[\text{单位距离物体数}] \text{ 满足接受准则的充分条件。}$$

可以用置信区间分别估计 $P(D^m > c)$ 和 $E[\text{单位距离碰撞}]$ ，反过来也可以借助以上不等式，通过获得置信度大于或者等于 $1 - \alpha$ 的声明 $E[\text{单位距离碰撞}] \leq \sigma$ 来论证整车级别的安全。由于可以通过仿真、结构化（例如轨迹）测试和随机道路测试的组合来估计 $P(D^m > c)$ ，并且通过道路测试以外的其他来源（例如交通数据、普通驾驶数据、区域成像）来估计 $E[\text{单位距离碰撞}]$ ，整车级随机道路测试的数量可以明显小于通过直接实现确认论据所需的整车级道路测试数量。

示例：接受准则是平均 100000 公里的驾驶中少于 1 次碰撞，且置信度水平至少 $1 - \alpha$ ，即 $E[\text{每公里碰撞数}] < 1/100000$ ，至少有置信度 $1 - \alpha$ 。假设估计每 100 公里有少于 1 个静止的道路物体，置信度至少为 $1 - \alpha_1$ ，即 $E[\text{每公里碰撞数}] < 1/100$ ，至少有 $1 - \alpha_1$ 的置信度。此外，假设人们还估计检测性能 $P(D^m > c) < 1/1000$ ，置信度至少为 $1 - \alpha_2$ ，其中 $\alpha_1 + \alpha_2 = \alpha$ 。然后，使用上面建议的上限， $E[\text{每公里碰撞数}] < 1/100 \times 1/1000$ ，置信度为至少 $1 - (\alpha_1 + \alpha_2) = 1 - \alpha$ 。这里的组合置信水平由基本概率规则¹⁾证明。

1) 假设任意两个事件 A 和 B，其中 $P(A) \geq 1 - \alpha_1$ 且 $P(B) \geq 1 - \alpha_2$ ，则 $P(A \cap B) = P(\Omega) - P(\Omega, (A \cap B)) = 1 - P((\Omega, A) \cup (\Omega, B)) \geq 1 - (P(\Omega, A) + P(\Omega, B)) \geq 1 - (\alpha_1 + \alpha_2)$

适当的试验设计有助于估计 $P(D^m > c)$ 和 $E[\text{单位距离物体数}]$ 。

示例： $P(D^m > c)$ 可以通过对遇到的物体随机抽样来估计， L 在区间 $[b, b + 1]$ 中均匀分布。
建模、充分条件和统计估计技术可以进一步细化^[42]。

C. 6. 3. 3 冗余和独立性的考虑

如果系统架构设计为：
——定义冗余通道来实现系统中给定的子功能；
——在给定的驾驶条件集合下，每条通道都能独立地实现该项子功能；
——任一通道的正确行为都能充分地保证预期功能的安全性，
那么可通过应用安全分析计算潜在系统危害行为发生的概率，降低每个通道的确认等级。
可通过如图C. 17所示的双通道系统来说明各通道确认等级的降低。

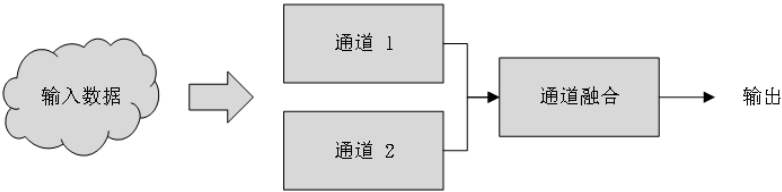


图 C. 17 双通道系统架构

对于图 C. 17所示的系统而言，假设通道1和通道2实现相同的功能，并且每条通道都能够避免潜在危害行为。
同时也假定通道融合要素不存在功能不足并能够融合两个通道的信息，这样对输入数据进行了正确评估的任一通道，都能充分避免输出产生的危害行为。在这些条件下，任一通道中的潜在功能不足就是多点功能不足。
通道1和通道2之间可能存在共模，例如相同的功能不足。在这种情况下，特定的触发条件可能会激活这种功能不足并引发危害行为。
基于这些假设，可以对导致危害事件的系统行为进行建模，如下图C. 18因果树（见B. 3. 2 因果树分析指南）。

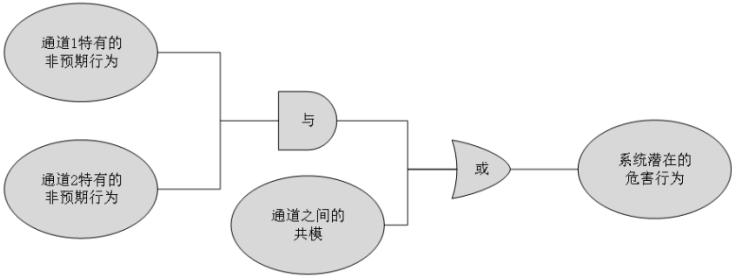


图 C. 18 系统行为模型

与纯粹的系统黑盒测试相比，单独测试通道1和2可以减少测试时间。这种减少很大程度上取决于通道1和通道2之间的共模因子。共模是由于通道1和通道2有相同的触发条件或者是具有统计意义相互依赖同时发生的不同触发条件而导致的。通过考虑两个通道之间的多样性，可以定性地估计这个因子。也可以通过专项仿真或分阶段测试计划进行估计。
理想情况下，可以认为通道是独立的（即因子=0）。并且可以通过专项独立性调查支持该独立性声明。这是量化策略的定性延伸。建立两个要素独立性的技术包括：

——分析：

- 包括已知现象的通道相关分析；
- 在通道中使用不同的传感器组；
- 在通道中使用异构的传感器原理和/或算法；

——专项试验和确认；

- 通过测试以表明系统能够处理假定的共因或相关性；
- 系统地设计耐久性确认测试，以充分测试所有已知或假定的传感器、组件和通道的弱点；

——专用方法；

- 用于覆盖通过理论或观测相关性得到的共因局限的方法；
- 从使用类似传感器或功能的其他系统中观测的功能不足的分析；
- 对开发过程中或通过现场监控观察到的单通道功能不足进行分析，提供证据证明其他通道不受此问题影响。

附录 D
(资料性)
关于 SOTIF 特定方面的指南

D.1 驾驶策略规范指导

D.1.1 目标和结构

D.1节的目标是指导驾驶策略设计，并提供实施示例。
驾驶策略是整车预期功能安全策略(VLSS)在决策层的实施。

示例1：因存在超出设计运行范围或性能局限，从正常状态切换到降级状态的相关要求属于整车预期功能安全策略范围。

在定义了整车预期功能安全策略之后，就可以通过分析某些影响设计和规范的关注领域来定义驾驶策略。在该过程中，可以考虑目标车辆的设计运行范围(ODD)和驾驶自动化等级^[2]。D.1提供了一些(但并不详尽的)示例，说明了如何推导整车安全策略，以及驾驶策略的要求。

整车预期功能安全策略是确ADS搭载车辆整体安全的总体规范，会影响ADS搭载车辆的所有构件设计。

如果实施了整车安全策略和驾驶策略，则在规范定义和设计中予以记录(根据第5章，并考虑第6章、第7章和第8章)，并根据验证与确认策略(第9章)予以验证。有许多方法可以实施驾驶策略，以适应和维持SOTIF。

示例2：参考文献[43]和[44]。

注：在驾驶策略中，可考虑与车辆基础设施(感知系统、执行器、人机界面)的依赖关系，并根据其他道路使用者的交互，考虑对道路安全的影响。

在开发过程中，可以通过是否违反驾驶策略来衡量ADS操控车辆充分应对其他道路使用者造成的危险情况的能力(这种衡量方法采用参考文献[45]中的道路驾驶技能概念表示)。

D.1.2 驾驶策略设计

D.1.2.1 驾驶策略设计示例概述

图D.1给出的示例是基于感知-规划-执行模型(4.2.3)的自动驾驶辆简化架构。

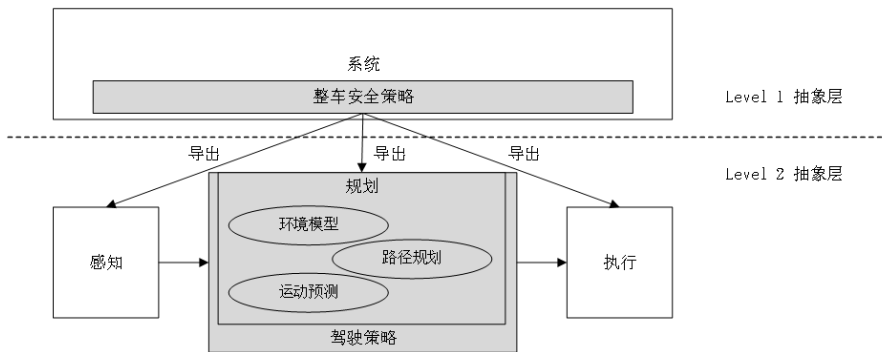


图 D.1 具有驾驶策略的 ADS 搭载车辆的简化架构示例

在所示架构中，驾驶策略属于规划子系统。规划子系统负责分析来自感知子系统的信息。规划子系统可包含多个子元素，其目的首先是以要求的精度水平重建ADS搭载车辆周围的环境，再根据驾驶策略(DP)决定系统的下一步行为。其中一种方法是设计适当响应。

注1：适当的响应被定义为驾驶策略在其他道路参与者做出可合理预见的行为时维持预期功能安全所需的一组纠正行为。适当的响应有两大属性：

- 可以根据交通场景中自车与其他交通参与者的关系，对ADS操控车辆进行评估；
- 在任何不需要切换到最小风险状态的运行条件下，通过统计证明是安全的。

示例1：通过适当的响应实现的修正动作包括加速、减速或转向指令，具体取决于交通场景。

驾驶策略的实施也可以依靠车外系统。

示例2：车外系统可能会影响驾驶策略的设计。例如依赖地图的更新了解道路基础设施的变化。

驾驶策略是为了确保在ADS搭载车辆的控制系统实施控制行为期间，ADS操控车辆在与其它道路使用者交互中，能够按照当地交通规则和公序良俗安全驾驶。驾驶策略可以作为监督手段实施，也可以直接纳入决策实施。在设计驾驶策略过程中，可以考虑以下两大类指标：

- 领先指标：对给定用例的指标，并可在验证与确认阶段(例如，在封闭道路或在开放道路测试)验证的指标。这些指标包括可通过仿真、封闭道路和开放道路中测试的违规行为、道路驾驶技能和系统脱离^[45]；及
- 滞后指标：根据统计数据得出的指标，在预期功能安全发布后，为风险可接受提供依据。滞后量度的有效性可在现场系统运行期间进行监测。见第13章运行阶段活动。

示例3：领先指标可包括车外系统的假设，如地图和定位，车对基础设施通信(V2I)，车对车通信假设(V2V)，以及对其它道路参与者的假设。

根据D.1.1给出的定义，驾驶策略的主要作用是监督驾驶自动化系统，将影响道路安全的危害行为的风险降到最低。为此，可以根据下述几项基本原则，制定驾驶策略规范：

- 可衡量；
- 反映车辆动力学和基本物理原理；
- 使用当前的技术；
- 反映交通规则；
- 将危险场景引发者与响应者分开(不惩罚规避行为)；
- 奖励可预测性和预测能力。

为有效实现预期功能安全，也可使用驾驶策略来预测和缓解感知子系统、执行子系统或车辆乘员人机界面的功能不足。

— 设计运行范围(ODD)的感知：在规定的设计运行范围之外使用ADS操控车辆，会增加危害行为的风险。驾驶策略可根据感知系统(包括传感器组和外部基础设施，如地图)提供的信息，避免自动驾驶功能在设计运行范围范围之外运行。在该用途中，感知系统需要符合额外的预期功能安全相关要求，即感知系统不会由于假阳性错误(车辆实际不在ODD范围内而感知系统报告车辆在ODD范围内)而造成风险，因为这样的错误会在ADS操控车辆超出ODD范围运行时导致伤害；

- 对于既可自动也可手动驾驶的系统，ADS搭载车辆的自动驾驶功能的启动和停止会使车内人员(包括驾驶员)困惑。在这种情况下，可在驾驶策略中考虑ADS操控车辆的自动驾驶功能的启动和停止与人机界面的依赖关系；
- 感知系统的局限：在恶劣天气条件下ADS搭载车辆的感知系统可能存在性能局限，或在特定用例下，存在性能局限。在驾驶策略设计中，可针对这些局限和不足，采取对策(例如，限制控制系统对执行器的权限，或将车辆安全地停下来)。

分析关注领域可以支持整车层面预期功能安全策略和驾驶策略的设计。可根据与运行环境、乘员、交通以及车载和车外系统的交互类型，对需要关注的领域进行分类：

- 从 ADS 搭载车辆的运行环境得到的关注领域。在该驾驶策略设计中，可采用以下示例性领先指标解决这类问题：
 - 设计运行范围；
 - 特定设计运行范围中的交通规则和公序良俗；
 - 道路基础设施(如交通灯、道路布局和交叉口类型)。
- 从驾驶自动化系统与 ADS 搭载车辆的乘员(驾驶员或乘客)之间的交互得出的关注领域。这些指标可在驾驶策略设计示例中，通过以下示例性领先和滞后指标解决：
 - 驾驶模式之间的切换，以及对驾驶策略范围之外的ADS搭载车辆子系统(如，人机接口)的行为的假设；
 - 切换至并维持降级模式；
- 从 ADS 操控车辆与交通场景中其他参与者之间的交互，以及与车载和车外系统的交互得出的关注领域。在该驾驶策略设计中，可采用以下示例性滞后指标解决这类问题：
 - 交通安全行驶所需的一套规则；及
 - ADS操控车辆在交互、预防或预测方面的不足或者局限(例如，感知系统传感器探测范围局限或覆盖范围不足)。

下文示例给出了如何利用关注领域分析来确保整车安全策略与驾驶策略设计的完整性。表D. 1至表D. 6分析了驾驶自动化系统的不同关注领域，以定义整车预期功能安全策略和驾驶策略要求，并描述了在偏离这些要求时，ADS操控车辆的预期行为。

注2：表D. 1至表D. 6给出了驾驶策略规范定义和设计的示例。本文使用了“应”这样的表述，但表格中的“应”仅表述示例要求，并不是要求遵守本文件的规定。

D. 1. 2. 2 ADS 操控车辆的运行环境引出的关注领域

通过设计、开发与确认驾驶策略来保证ADS操控车辆仅在给定的设计运行范围中运行，而设计运行范围有多个因素约束。车辆在设计运行范围之外运行时，危害行为的风险可能会增加。

设计运行范围可基于不同方面。示例包括：

- 地理限制：ADS 操控车辆只能在特定和有限区域(城市、国家等)不受监控地运行；
- 道路类型限制：ADS 操控车辆只能在特定一种(仅高速公路、仅城市道路等)或几种道路上不受监控地运行；
- 天气条件：ADS 操控车辆不允许在特定天气条件(大雨、大雪等)下运行；及
- 车速：ADS 操控车辆不允许超过特定速度运行。

表 D. 1 设计运行范围得出的考虑

目的	考虑设计运行范围之外的功能运行所引起的风险(包括驾驶员对设计运行范围的可能的误解或缺乏了解)。	
整车预期功能安全策略	ADS搭载车辆应确保“自动驾驶”模式在 ODD 之外不被启动。	
ADS操控车辆的考虑点	潜在后果	为降低风险而提出的驾驶策略功能要求(R)或假设(A)
ADS操控车辆在指定地理区域之外运行		R：驾驶策略应监督ADS操控车辆，确保其在指定区域内运行。

	由于ADS操控车辆在其行为未经确认的区域运行，造成危害事件(例如：任何类型的碰撞)。	R：在指定区域之外运行时，驾驶策略应监督ADS操控车辆，确保其切换到最小风险状态(MRC)。 A：驾驶策略应接收外部独立子系统(如定位子系统)发出的更新的位置信息。
ADS操控车辆行驶超过设计的车速	因任何潜在的局限(感知或执行局限)，ADS操控车辆造成危害事件(例如：任何类型的碰撞)。	R：驾驶策略应监督ADS操控车辆，确保车辆在最大设计车速以下运行。
ADS操控车辆的预期行为	驾驶策略应： <ul style="list-style-type: none"> — 监督 ADS 搭载车辆是在设计运行范围之内还是之外运行；及 — 如果 ADS 搭载车辆在设计运行范围之外运行，执行以下任一策略： — 禁止驾驶自动化功能启动(若尚未启动)； — 要求驾驶员收回控制权(如果驾驶自动化功能预见到 ODD 退出)；或 — 关闭自动驾驶功能，并切换到安全状态(如果驾驶员不在控制环路之内)。 	

表 D.2 从其他道路使用者行为的假设得出的考虑

目的	确保ADS操控车辆采用防御性驾驶技术作为减少潜在碰撞的一种手段。	
整车预期功能安全策略	ADS操控车辆符合相关驾驶规则、法律和公序良俗，除非违反其中的一项或多项才能避免事故。	
ADS操控车辆的考虑点	潜在后果	为降低风险而提出的驾驶策略功能要求(R)或假设(A)
其他道路使用者(ADS操控车辆除外)不遵守交通信号灯	ADS操控车辆在没有路权的情况下占据了一个路口，造成危害事件(例如：任何类型的碰撞)。	R：驾驶策略应根据其他道路使用者的行为假设，监测其他道路使用者是否有能力停车。
ADS操控车辆无视其他交通场景的参与者的路权	因任何潜在的局限(感知或执行局限)，ADS操控车辆造成危险事件(例如：任何类型的碰撞)。	R：驾驶策略应当利用已定义的对其他道路使用者的合理地最坏驾驶行为的假设来应对感知限制或者感知元件被遮挡造成的问题。
ADS操控车辆的预期行为	驾驶策略应： <ul style="list-style-type: none"> — 执行防御性驾驶技术(同时避免太过于保守可能引发的事故)； — 遵守当地的驾驶规则和公序良俗；及 — 只有在避免事故时，才能违反当地的驾驶规则和惯例。 	

示例：参考文献[43]和[44]通过定义了“路权是被赋予的，不能争夺”的原则，总结了实施防御性驾驶技术的必要性。

注：驾驶策略规范可以不同(或需要不同配置)，这取决于ADS操控车辆的目标市场(地域区别)和当地行为(例如，美国和欧洲对于十字路口的路权定义不同)。

表 D.3 从道路基础设施得出的考虑

目的	确保ADS操控车辆能够在设计运行范围中定义的所有道路条件下运行，而不违反预期功能安全要求。
整车预期功能安全策略整车安全策略	ADS操控车辆应确保不在设计运行范围之外启动驾驶自动化功能。

ADS操控车辆的考虑点	潜在后果	为降低风险而提出的驾驶策略功能要求 (R) 或假设 (A)
在道路施工区域，ADS操控车辆不能识别狭窄车道	ADS操控车辆失去车道轨迹，进入相邻车道，导致危害事件(例如：侧擦碰撞)。	R：驾驶策略应了解道路基础设施并监督ADS操控车辆的行为。
ADS操控车辆的预期行为	驾驶策略应实施防御性驾驶技术或监督ADS操控车辆，确保切换到降级运行模式。	

D. 1. 2. 3 ADS 操控车辆的运行环境引出的关注领域

设计驾驶策略可以根据ADS操控车辆的状态，允许多种降级运行模式。用户和ADS操控车辆之间的交互会影响驾驶策略监督驾驶任务完成情况的方式。

表 D. 4 从预测、预防或减轻 ADS 操控车辆性能局限的需求得出的考虑

目的	预测、预防或减轻驾驶自动化系统基础设施的已知限制(例如感知或执行子系统)。	
整车预期功能安全策略	驾驶自动化系统应确保ADS操控车辆不在设计运行范围之外启动。	
ADS操控车辆的考虑点	潜在后果	为降低风险而提出的驾驶策略功能要求 (R) 或假设 (A)
在恶劣的天气条件下，ADS操控车辆性能的会下降	恶劣天气条件下，ADS操控车辆无法处理感知和执行子系统的降级，导致危害事件(例如：任何类型的碰撞)。	R：驾驶策略应监督 ADS 操控车辆，确保根据天气条件调整性能。 A：驾驶策略能收到外部(驾驶策略外部)子系统发出的当前天气情况信息。
在特定类型的道路上或特定道路条件下，ADS操控车辆的性能会下降	在特定类型的道路上(例如，附着力低的砂石路面)，ADS操控车辆无法处理子系统降级，导致危害事件(例如：任何类型的碰撞)。	R：驾驶策略应监督 ADS 操控车辆的性能，确保适应道路类型。 A：驾驶策略能收到外部(驾驶策略外部)子系统发出的当前道路类型信息。
ADS操控车辆在接近有遮挡区域的路口时速度过快(遮挡区域是指感知子系统视野被建筑物或其他基础设施遮挡，无法提供可靠感知的区域)。	在接近遮挡区域时，ADS操控车辆无法及时感知到其他道路使用者，造成危害事件(例如：任何类型的碰撞)。	R：在因基础设施或道路设计造成遮挡时，驾驶策略应监督 ADS 操控车辆，确保减速(以示谨慎)。 A：驾驶策略能收到外部子系统发出的感知系统视野受阻信息。
ADS操控车辆的预期行为	驾驶策略应监督 ADS 操控车辆的基础设施的状态： — 在接近遮挡区域时，驾驶策略应调整 ADS 操控车辆的行为； — 在可能有损车辆行驶能力的恶劣天气条件下，驾驶策略应调整 ADS 操控车辆的行为； — 在执行子系统降级，可能有损车辆行驶能力的情况下，驾驶策略应调整 ADS 操控车辆的行为；及 — 在低附着力路面行驶时，驾驶策略应调整ADS操控车辆的行为。	

示例：参考文献[43]和[44]表明，在传感器感知因基础设施、道路设计和/或动态物体(如车辆、行人、自行车)被遮挡时，需要谨慎行驶。在接近路口时，如果感知系统探测其他道路使用者的能力受到其他建筑物或道路布局的影响，可能需要 ADS 操控车辆降速，才能保持谨慎行驶。

表 D. 5 从需要管理运行模式切换得到的考虑

目的	确保运行模式切换不会影响预期功能安全。需要考虑以下情况： — 驾驶员对启动/停用步骤的认识或理解不足； — 驾驶员对当前模式的误解； — 驾驶员在给定时间内以恢复充分的态势感知能力，以及控制当前运行状况的能力。	
整车预期功能安全策略	驾驶自动化系统应防止对驾驶自动化系统功能的可预见误用，并向驾驶员通报以下信息： — 预期系统行为及限制； — 对驾驶员的期望以及驾驶员取回车辆控制权的预期程序。	
ADS操控车辆的考虑点	潜在后果	为降低风险而提出的驾驶策略功能要求 (R) 或假设 (A)
信息不一致时，ADS操控车辆切换到降级运行模式(指在某些条件下驾驶员可以成为控制环的一部分的驾驶自动化等级)	驾驶员无法收回车辆控制权，ADS操控车辆可能导致危害事件(例如：任何类型的碰撞)。	R: 驾驶策略应监督车辆控制权从 ADS 操控车辆向驾驶员转移的过程。 A: 如果驾驶员在设计运行范围结束前未能收回控制权，驾驶策略应确保ADS操控车辆进入安全状态。
需要驾驶员接管时，ADS操控车辆未考虑人类驾驶员的局限(例如，接近设计运行范围边界时)。	驾驶员没有充足的时间收回车辆控制权，ADS操控车辆可能导致危害事件(例如：任何类型的碰撞)。	R: 驾驶策略应充分通知驾驶员，要求驾驶员收回车辆控制权。 A: 如果驾驶员在设计运行范围结束前未能取回控制权，驾驶策略应确保ADS操控车辆进入安全状态。
ADS操控车辆的预期行为	驾驶策略应通知驾驶员，需要取回车辆控制权。如果驾驶员未能取回车辆控制权，驾驶策略应： — 监督 ADS 操控车辆，确保切换到安全状态 (MRC)； — 禁用驾驶自动化系统。	

注：这组考虑点针对的是，在某些条件下，驾驶员可以成为控制环路的一部分的驾驶自动化等级。

D. 1. 2. 4 ADS 操控车辆与其他交通参与者之间的交互引出的关注领域

通过多种技术手段实现ADS操控车辆监督其他道路使用者，并据此调整自身的行为。RAND^[45]对这些技术进行了分类，并指出对ADS操控车辆周围的“安全包络”进行监督，是实现更高的驾驶自动化等级的最有效技术。对于其他水平的驾驶自动化系统，可以推荐其他方法。

“安全包络”是一个通用概念，可以用来制定驾驶策略可以遵循的所有原则。根据这一概念，ADS操控车辆可以在自车周围有一个或者多个边界。在某些场景下，违反其中一个或多个边界时，ADS操控车辆会作出不同响应。例如，驾驶策略适应这些场景，实施适当响应，维持预期功能安全。

由于适当响应的“整车”性质，D. 1. 2. 4假设：

- ADS 操控车辆不能控制交通场景中任何其他参与者的行为，因此其适当响应可定义为，ADS 操控车辆不会导致危险场景，并引发事故；
- 实现 ADS 操控车辆的预期功能安全仅与交通场景中的其他参与者(道路使用者)有关，如图 D. 2 所示。因此，可针对交通场景中的任何道路使用者，定义 ADS 操控车辆的适当响应。这些因素都会限制 ADS 操控车辆的动态行为(横向或纵向加速度、滚动和悬垂)。

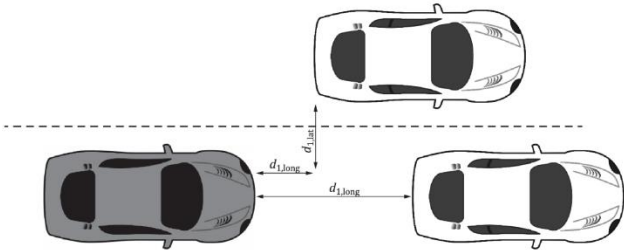


图 D.2 自车与其他道路使用者的相对位置的定义

表 D.6 从需要管理运行模式切换得到的考虑

目的	考虑以下因素，确保 ADS 操控车辆在预期设计运行范围内运行，同时不会造成事故： <ul style="list-style-type: none">— 其他道路使用者引起的风险；— ADS 操控车辆为确保其他道路使用者充分的可控性水平进行操控而引起的风险；— 其他道路使用者未按照可合理预见的假设采取行动而引起的风险；— 自车执行器性能引起的风险。	
整车预期功能安全策略	1) 对周围道路使用者而言，ADS 操控车辆的行为应尽可能可预测（例如，不得随意变道，在接近并道时的行为可预见等）。 2) 由 ADS 操控车辆应根据以下规则管理风险： <ul style="list-style-type: none">— 不能成为事故原因；— 尽可能合理的对其他交通参与者造成的风险保持鲁棒性；— 在能见度有限的地区保持谨慎；— 在遇到不明物体或其他交通参与者的意外行为时，保持谨慎；— 注意其他车辆（与前车保持安全距离，不得鲁莽超车）；及— 遵守适用的交通规则/法律，除非必须违反交通规则/法律才能避免事故；— 实现类人驾驶，驾驶策略既不过于激进，也不过于保守。 应遵守这些规则： <ul style="list-style-type: none">— 无论车辆在哪里行驶（如乡村、公路）；— 无论车辆在何时行驶（例如，尽管有动态车道分配、与时间有关的规则、引入新类型的交通标志、新规则等）。	
ADS操控车辆的考虑点	潜在后果	为降低风险而提出的驾驶策略功能要求 (R) 或假设 (A)
ADS操控车辆离前车太近/太紧	ADS操控车辆不遵守跟车距离，将导致危害事件（例如：追尾碰撞）。	R：驾驶策略应监督ADS操控车辆，确保与前车保持避免碰撞所需的最小距离。
ADS操控车辆变道时未考虑其他车辆	ADS操控车辆未能与其他道路使用者保持横向距离，导致危害事件（例如：侧擦碰撞）。	R：驾驶策略应监督ADS操控车辆，确保始终与任何道路使用者保持避免碰撞所需的最小横向距离。
ADS操控车辆在雪天行驶过快，未考虑执行器在湿滑路面上的性能	ADS操控车辆未减速且未谨慎驾驶，导致危害事件（例如：与物体/行人相撞）。	R：驾驶策略应监督ADS操控车辆的执行指令与给定的环境条件相匹配。

ADS操控车辆在无信号灯路口左转过于保守	ADS操控车辆被后车追尾	R：驾驶策略应监督ADS操控车辆，避免驾驶策略过于保守。
ADS操控车辆在无信道等路口左转过于激进	ADS操控车辆与直行车辆发生碰撞。	R：驾驶控制策略应监督ADS操控车辆，与直行车辆保持必要的安全时间间隔。
ADS操控车辆的预期行为	驾驶策略应监督ADS操控车辆，并根据其他道路使用者的行为，调整自车行为。	

示例：[43]和[44]解决了图 D.3 中所示驾驶情况，其中自车 c_1 以正接近车速跟随车 c_2 (c_1 比 c_2 快)。在该场景中，如果在某个时间点 t_d ，两车之间的距离正好是两车无法在给定响应时间避免碰撞的距离，则认为车 c_1 和 c_2 在时间点 t_d 有危险。 c_1 将对追尾碰撞负责。

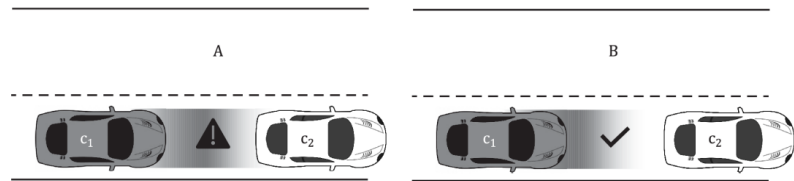


图 D.2 危险场景在正确实施响应前后的示例

c_1 的驾驶策略通过确保在任意时间点 t_d 两车的纵向距离足够大，使得 c_1 有足够的制动距离来避免与 c_2 的后撞。这涉及到 ADS 操控车辆利用他的制动和加速装置调节 c_1 与 c_2 的相对距离，以保持最小距离 d_{\min} 。

注：交通统计可被视为一种合适的方法，这种方法根据驾驶自动化的目标等级、ODD和市场，识别ADS操控车辆与其他参与者之间最常见和最严重的碰撞模式。这些模式可以代表ADS操控车辆在现场可能面临的最严重的危险情况，并可用于制定监督自车行为的驾驶策略规则。

D.1.3 整车级SOTIF策略和驾驶策略的验证和确认

驾驶策略是在目标ODD的上下文中设计的，用以实现 VLSS 的决策逻辑(D.1.2)。然而若驾驶策略不能充分反映现实生活中的情况，则它可能成为功能不足的潜在来源。验证和确认活动(第9、10 和 11 章)可以通过组合场景、结合相关参数等来揭示驾驶策略的弱点。然后通过进一步的SOTIF流程迭代(图7和8)解决已识别的弱点，包括进一步分析(第6和7章)和修改驾驶策略(第8章)。

通过定义衡量自动驾驶系统有效性的指标，驾驶策略还可以作为定义第9章中已选验证和确认方法的判定标准的法则。

例如 ADS 操控车辆性能可以通过监督如下违背驾驶策略的指标来衡量：

- 驾驶策略未能监督 ADS 操控车辆行为的次数；
- 驾驶策略无法检测到危险情况的次数；
- ADS 操控车辆造成的事故数量。

D.1.4 驾驶策略现场运行

在运行过程中，可以对驾驶策略的有效性进行评估。此活动可以通过比较系统使用情况(13.3)和观察到的驾驶策略有效性的统计信息来完成。

D.2 对机器学习的建议

D.2.1 总则

自动驾驶技术经常涉及一些类型的机器学习算法，特别是对于物体探测和分类。机器学习算法主要用于不可能完整描述问题的情况（例如，不可能通过数据完整地表征所有类型的行人，使得其总能被算法识别出来）。机器学习算法可以用来克服这一问题。机器学习算法通过提取数据中存在的相关性来学习，将输入映射到输出。因此与人类不同的是，机器学习算法不会学习上下文语义。虽然机器学习算法通常比非学习算法表现更好，但通常难以理解算法的预测过程。因此许多机器学习算法的局限是与直觉相悖的，也无法被规范定义。

为了最大限度地减少由于错误预测而导致的残余风险，一些方法可以用来减轻导致性能局限的机器学习组件的局限性（例如，物体识别率低于100%，非预期的偏差）。机器学习的训练，包括所使用的数据，有可能引入安全风险从而导致性能局限。例如，训练和验证数据可能引入源于数据偏差的特性离散和相关。而这些特征分布和相关，就系统的预期功能而言，可能是不恰当的甚至是不正确的。由于鲁棒且准确的机器学习组件对车辆的安全运行至关重要，需要开发学习系统和相应的数据收集流程以减缓机器学习组件的局限。

D. 2. 2 GB/T34590与SOTIF对机器学习的建议

D. 2. 2阐述了和比较了GB/T 34590与SOTIF在机器学习安全方面的职责：

- a) 离线训练中使用的工具需要符合 GB/T 34590. 8 第 11 章以及本文件附录 D. 2. 4 离线训练流程的部分。

注1：硬件和软件开发所包含的工具（例如离线服务器场和训练软件）应作为工具鉴定的一部分进行评估。

- b) 对于在车辆中实现机器学习的硬件（例如 GPU），考虑以下两方面：
 - 1) 随机硬件故障和硬件系统故障由 GB/T 34590. 5 覆盖；
 - 2) 硬件的性能局限可以既是 SOTIF 问题也是 GB/T 34590 的系统问题。
- c) 对于实现机器学习算法的软件，可以考虑以下几个方面：
 - 1) 机器学习软件通过对输入进行特定的运算（例如矩阵乘法、离散卷积、非线性函数）生成输出。因此，它本身与其他非学习算法没有什么不同，并且可以通过常规手段进行验证。特定运算的实现可以根据 GB/T 34590. 6 进行验证。

示例：在机器学习运算中需要特别关注浮点运算库，以及他在训练设备和嵌入式目标环境的区别。

- 2) 机器学习软件的功能（如物体检测）是 SOTIF 关注的另一方面。模型以及训练（数据驱动过程）获得的权重可能造成模型预测的不确定性，这种不确定性构成了本文件需要解决的功能不足。识别和缓解机器学习的局限（如由于内置偏差或不完整的训练集，错误的标注，数据缺失，过于关注罕见事件）是 SOTIF 过程中减少区域 2 和 3 的部分（根据图 5 和图 6 所示），同时也被 SOTIF 验证和确认。
- 3) 机器学习算法训练的权重可以被视为应用软件校准或配置数据，并可通 GB/T 34590. 6 附录 C 中的适当要求来解决。

注2：虽然机器学习的权重是数值，但它们可能对预期功能有定性的影响。GB/T 34590中的标定数据用于调整已知模型的行为，而模型的权重却用于定义模型本身。因此，更改权重需要对系统进行影响分析和重新确认。

注3：引入运行监视器，用于检查机器学习组件开发中假设条件，可同时适用于GB/T 34590和 SOTIF。监视器的检测结果可用于识别随机硬件故障和系统性硬件故障，系统性软件故障（GB/T 34590）以及系统局限（SOTIF），并最终进入安全状态。

D. 2. 3 采用机器学习算法的预期功能的安全

当采用机器学习技术实现安全相关系统时，定义相关功能十分重要。这意味者，由于很难在恰当的用例和场景下提供充分的安全论证，因而如果没有定义预期的功能，采用机器学习技术的功能会被视为是不安全的。例如，对于深度学习方法，由于固有的非线性和缺乏对这类算法的形式化确认，在离散的

用例和场景下的好的表现通常不足以证明他的安全性。对于这类算法，需要补充额外的确认步骤来支持安全论据。

复杂机器学习算法的行为很大程度上取决于训练集、机器学习模型架构，训练过程(训练算法，批处理大小，权重初始化，损失函数等)，他们显示了难以通过分析来理解的规范。因此，重要的是通过执行本文件中推荐的适当测试(第9、10和11章)并对机器学习局限性进行分析(第7章)来评估分配到机器学习算法的功能的安全性。

采用机器学习的要素可以检测影响其性能的条件(如在不利天气条件下传感器检测到物体的低置信度)。根据指导4.4，需要将这些已知条件和要素输出导致的结果共享给上层系统的开发人员。

尽管已知的残余场景包含触发条件(如导致假阳性或假阴性)，且这些触发条件在已标注的训练、验证、测试数据集中，使用机器学习的要素的离线训练过程还是可能会接受训练的参数。这些识别到的触发条件、对其风险评估和减轻SOTIF相关风险的潜在措施应与上层系统的开发人员共享，以便他们可以采取适当的行为。这些行为可以在系统层面进行，例如，机器学习要素的性能局限由处理链路中的后续组件解决。或者识别的触发条件可用于改进机器学习要素级别的训练过程。

将本文档应用于基于机器学习的要素时，可以考虑以下几点：

——功能和系统设计(第5章和第8章)。

- 用例的规范，包括相关的ODD，在SOTIF中具有重要作用，对于收集和创建用于训练、验证和测试的数据集也很重要。不太可能完全定义ODD的所有方面或所有情况下的机器学习的相关因素。训练集的质量对学习如何在已知非危险情况下正常运行至关重要(区域1)。
- 充分的测试数据集增加了基于机器学习组件的安全置信度(区域2和3)。系统设计(或架构设计)同时也需要明确阐述分配给基于机器学习算法的功能。

注1：机器学习主要有两种不确定度：认知不确定度和随机不确定度^{[46]，[47]}。当认知不足时，通常可以通过引入更多数据来减少认知不确定度。而随机不确定度与数据噪声的固有不确定性有关，因此不能随着更多数据而进一步减少。

示例1：一个物体检测子系统由基于机器学习的图像识别和后处理机制所组成。在这个概念中，机器学习算法的误分类不被视为故障，而是被视为性能相关事件，因为后处理机制通常能够从图像序列中过滤它，只有剩余的误分类率才有潜在的安全影响。

——分析(第6章和第7章)。

- 通过分析可以识别出测试用例和场景集，他们可以验证基于机器学习的组件的功能。

——V&V策略(第9章)。

- 识别用于测试的组件边界很重要。边界选择不仅会影响测试的准确性和完整性，而且会影响测试预言(例如仿真、测试数据和真值)的可用性和适用性。
- 测试可以在三个抽象层次上进行：
 - 1) 在基于机器学习的算法级别进行独立测试，可以有效地找到机器学习组件典型的未知性能局限(例如可视化)；
 - 2) 在组件级别，根据功能和需要测试的内容进行测试，是评估包含其他相关组件算法行为的更好方法(例如，在物体检测中的后处理过滤器)；
 - 3) 在整车级别，测试整车层面的危害行为。

与单独测试每个组件相比，测试完整的处理链可能需要更多的测试例和测试时间。

示例2：假设三个要素组成的处理链，且期望这三个要素一起实现0.1%的假阴性。在采用适当的测试步骤情况下，相比测试每个组件(预期的10%假阴性)，对处理链需要更多的测试。

——评估(第10章和第11章)。

- 基于机器学习的预期功能安全是通过评估、特别是通过测试(第10章和第11章)保证的。因此，确保这些评估方法的结果能够反映现实中的行为非常重要。

- 一旦为了改进功能对机器学习组件进行了进一步训练(第8章,如增加目标检测的类别),就需要重新测试该组件。重新测试是因为很难甚至不可能理解变化对内部算法行为的影响。因此以前的测试结果通常不再有效,并且不应被再次使用。

注2:应对已发布的算法或参数的任何更新实施变更管理。如果重新学习,无论是在线还是离线,开发都回到SOTIF过程的相关阶段。

——运行阶段(第13章)

根据第13章,功能在现场运行中受到监督。需要分析观察到的新风险以识别功能不足,包括使用机器学习组件的功能不足。如果基于分析的结果改进了机器学习组件,需要重新进行整个预期功能安全活动(从第5章到第12章),包括数据采集过程。

D.2.4 对机器学习算法离线训练部分的建议

机器学习通常涉及离线训练过程,其目的是确定机器学习算法的参数。离线机器学习训练可能涉及多个步骤和工具。由于内置偏差,不完整训练集或对模型的验证不足使得这些步骤和使用的工具会出现问题。

机器学习开发过程中的常见问题包括:

- 不完整的训练集或对训练参数的验证不足;
- 预测时与直觉相悖的原因(例如,对抗攻击)
- 训练数据选择会影响机器的性能(例如训练数据中的偏差会导致错误地学习映射关系);
- 学习过程不能人为控制(例如错误地学习映射关系);及
- 测试数据的选择决定了现场性能估计的准确性(例如不当地划分测试数据会导致过估计/欠估计)。

由于机器学习算法反直觉的性质,对于复杂的任务,特别是自动驾驶对于开放环境的识别无法实现100%的性能要求(即总会有算法输出错误预测的情况)。正确的训练能够减少错误结果的数量,但永远不会完全消除错误结果。SOTIF活动的任务之一是确保这些性能局限不会导致不合理的风险。

图 D.4 显示了一个训练过程。

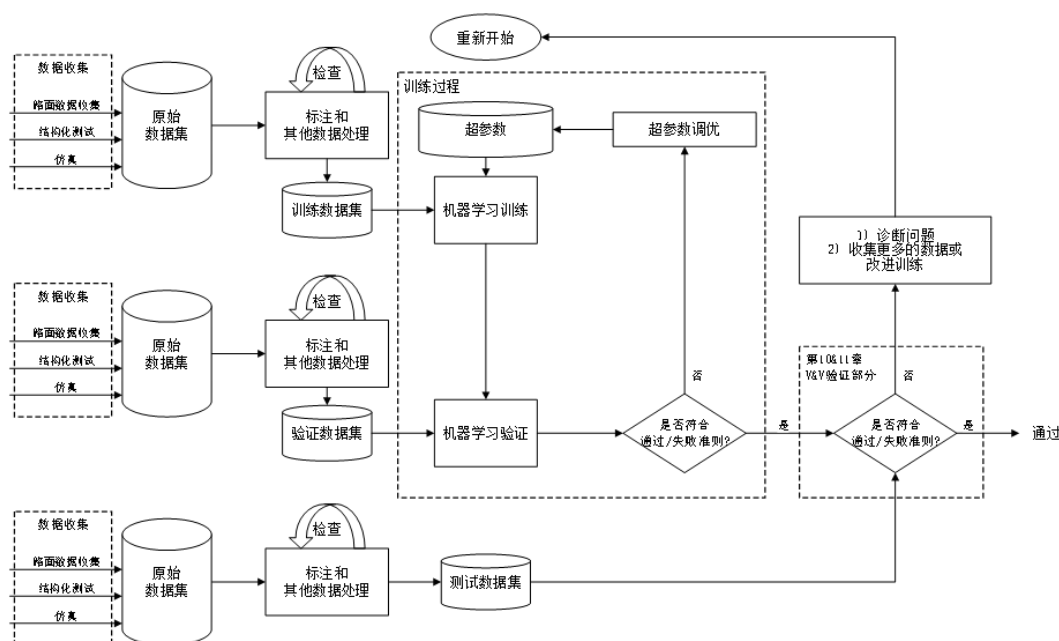


图 D.3 离线机器学习开发流程示例

图D. 4是从准备数据集开始的。数据集代表ODD中场景快照/场景，必须承认它是现实世界的一个有限近似。通过应用重要性抽样技术，经过训练的机器学习模型可以在罕见用例中表现的很好。数据集可以包含一些额外的考虑因素，比如产品生命周期的变化(如传感器老化)。数据可以从多个来源收集，例如，封闭道路上的测试、仿真、道路数据收集和标准基准数据集。

为了更好的模拟自动驾驶系统所处的现实世界，数据可能包括不同相关条件/因素的多种变化和组合。仿真和专项测试(例如设计并执行在封闭道路上的测试)可用于补充现场数据采集中很少自然发生的情况，并增加变化的多样性。当合成数据经过验证与现实世界数据相符后，可以用于扩充数据集。

然后，在用于ML训练、ML验证或ML模型测试之前，对数据进行预处理。预处理阶段可以根据类别(例如道路边界、汽车、摩托车、应急车辆)、特征(例如颜色、轮廓)或响应(例如要求的控制行为)对数据进行标记(注释)。数据标注可以使用经过培训的人员或通过自动化流程执行。通常，手动执行过程包括对标注的检查。要特别注意是数据的标注过程，以确保标签类别的正确性和机器学习过程中足够的边界精度。根据机器学习用例和数据集的类型，预处理阶段可能还涉及一些其他过程，例如过滤、数据增广和降维。预处理的数据通常使用数据清理技术(例如，删除重复或不相关的观察值等)进行增强。

下一步，数据被(充分地)分成独立的数据集包括训练、验证、测试，这些数据集用于不同的目的。为避免数据集之间的信息泄漏，尤其是训练和测试数据集之间的信息泄漏，通过测试数据评估机器学习模型的可靠性非常重要。训练集和验证集用于机器学习模型的训练过程。机器学习训练过程是一个超参数调优的循环过程，包括训练和验证过程。为了训练一个机器学习模型，训练数据会连续输入给模型，同时根据输出误差的斜率调整其参数(例如神经网络权重)。训练继续进行，直到达到预定的通过/失败标准，例如在物体检测或分类时的可接受的假阳性率和假阴性率。其他任务可能需要其他特定标准。

一旦完成机器学习模型的训练，就可以使用验证集根据通过/失败准则对其进行评估。如果结果不令人满意，将优化机器学习模型的超参数，并重复训练过程。完成机器学习训练后，作为V&V活动的一部分(第10章和第11章)，可通过测试数据集依据测试通过/失败准则对已训练的系统进行评估。

训练、验证和测试数据集的独立性可用于确保机器学习系统已学习到训练数据的基本特征，而不是其固有的巧合相关^[48]。

例如，我们获得了两个测试集：

- a) 利用从训练-验证-测试划分方法中得到的测试数据进行测试，有助于理解机器学习组件的泛化；
- b) 使用单独收集的数据集进行测试可以确保机器学习不会学习到巧合相关。

为了准确估算ML实际性能和探索未知的危险场景，测试数据仅用于整个流程中的测试验证部分，不能用于训练机器学习组件。理想情况下，可以用大的测试集来发现过拟合。这可以作为第10章和第11章中整车级V&V活动的一部分。

一旦满足作为规范定义和设计的一部分记录的测试通过/失败标准，则接受训练过的参数。如果验证失败，可以在收集更多数据和/或修改训练模型后重新启动该过程。一旦模型被接受，如果在进一步测试中观察到模型的不当行为或性能局限，则可以作为新信息被考虑，并重复整个过程。该循环与图7和8中SOTIF开发流程一致。

对于指定的ODD，标记过的数据集对场景的高覆盖率是训练鲁棒性的关键参数。在开发阶段SOTIF迭代过程中和运行阶段发现的新场景可以导致训练、验证和测试数据集的更新。

通过分析机器学习软件的可解释性，可以提高其可信度^{[49], [50]}。可解释性分析可以证明机器学习的决策(例如分类)是基于相关数据而非人为^[51]。可解释性分析可以是机器学习确认和/或测试验证期间的通过/失败准则的一部分(图D. 4)。

D. 2.5 机器学习算法的离线训练过程分析

通过分析可以发现许多训练局限问题，包括验证和确认活动。图D. 5描述了训练过程中分析问题的
一般流程。

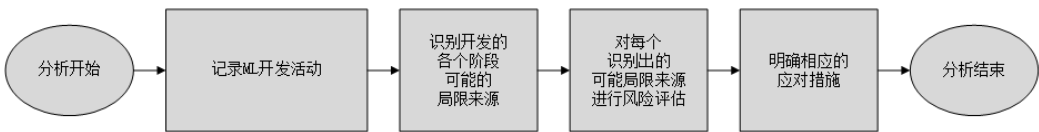


图 D. 4 机器学习算法离线训练过程的分析步骤

该分析方法可以用于分析所有的过程，例如：

- 数据收集路线策划；
- 数据收集；
- 数据上传和接收；
- 数据标记和评审；
- 数据整理和检索；
- 元数据标记和评审；
- 训练、验证和确认测试数据集的创建；
- ML 训练；
- ML 验证；
- ML 配置管理；
- ML 部署及集成到软件。

对于指定的任务和ODD数据收集旨在提供多样性和完整性：

- 车辆和驾驶员；
- 路线和驾驶条件；
- 结构化数据收集(例如，使用基于轨迹场景的数据收集)。

注：一个类似FMEA(失效模式和影响分析)的分析可用于分析和消除离线训练过程中可能存在的偏差和局限。

分析结果不仅可用于改进训练过程(D. 2. 5)，还可对系统开发产生影响，包括：

- 规范定义和设计(第 5 章)：通过识别潜在的系统性问题(第 7 章，A. 2. 8)和之后的改进活动(图 10)；
- 软件工具使用的信心(A. 2. 9)。

D. 3 地图预期功能安全的考虑

D. 3. 1 地图预期功能安全考虑简介

地图用于支持或实现ADAS及自动驾驶所需的功能，例如定位、路径跟踪、物体的车道分配和地标识别(如交叉路口和车道合流)。此外，地图还可以与感知传感器融合，以增加感知系统的置信度和/或增强探测故障的能力。D. 3详细说明了使用地图支持或实施安全相关功能时应考虑的地图使用的一些方面。

D. 3. 2 地图的规范定义和设计

地图属性作为规范定义和设计的一部分进行说明(见第5章)。设计考虑因素可以包括：

- 车辆功能：
 - 地图的使用；
- 对地图的依赖性；
- 在使用地图的情况下，整车层面的行为或功能：
 - 地图不可用(例如，当与地图服务器的连接丢失时或车辆嵌入式设备上的地图丢失时)；

- 地图不准确或过时；
- ADS操控车辆感知系统与地图发生冲突的解决；
- 定义地图特征：
 - 地图系统需求；
 - 地图数据需求；
 - 地图正确性要求；
 - 地图精度级别(即高精度地图还是普通精度级别的地图)；
 - 地图信息流描述；
 - 地图中记录物体的位置精度要求；
 - 地图的有效区域；
 - 保证和维持地图正确性的机制(即保证一定水平的地图质量)。
- 地图更新的技术手段：
 - 地图更新机制；
 - 地图更新频率；
 - 云端和车载地图的存储及地图更新的解决方案。
- 已知的地图局限：
 - 数据采集的局限；
 - 数据处理的局限；
 - 地图融合的局限(包括地图更新、多地图融合、多驱动融合)。

注：随着地图的使用年限增加，数据的不确定性会降低。

D.3.3 地图预期功能安全的建议

由于临时(例如临时车道关闭)或者永久(例如新路标)的环境变化导致地图不准确或过时，就是与SOTIF相关的地图问题示例。

其目的是确保地图系统的不足不会影响SOTIF。规范定义和设计可以指导如何通过各种方法处理地图的局限，例如：

- 限制无地图区域的功能；
- 限制地图精度较低的区域的功能；
- 更新地图或偏差。

SOTIF要求可以指定地图更新的频率。虽然可以通过地图和自车感知系统结果之间的比较来检查地图是否过时，但是仍然有可能由于感知系统的局限导致感知的结果与地图不匹配，仅凭感知可能无法保证检测到所有地图不足。

虽然地图不能始终提供实时道路状况，如因事故、施工和天气(如洪水)造成的车道封闭，但预期功能安全要求可以规定，基于安全分析，地图提供的永久道路基础设施标识应在一定精度公差范围内。系统可以通过实时监督一些指标来预测地图更新的需求(“过时地图”问题)。此外，规范定义和设计旨在记录地图系统的所有已知局限；在开发任何使用地图及其服务的功能时，都应考虑应对这些局限的措施。

注1：由于系统失效导致地图损坏的相关问题被GB/T 34590涵盖，不在 SOTIF 的范围内。损坏示例包括内存损坏、地图访问错误和地图下载错误等。

注2：在构建地图的过程中，可能会产生系统性错误或功能不足。可以通过分析或审核流程来发现这些问题。

D.4 V2X 功能安全的考虑

D.4.1 V2X预期功能安全考虑简介

V2X(车联万物)是一种允许车辆与其他车辆、道路基础设施、道路行人和云平台进行通信的机制。V2X可用于加强道路安全、提高效率和减少污染^{[52], [53]}。V2X可以用来满足不同车辆的互联需求,比如远程维护、交通运输管理以及车内信息娱乐等领域的应用。

V2X有能力告知自车其周边环境情况,尤其在恶劣天气环境和复杂交通场景下^[54]。例如,车辆很容易通过V2X得到交通灯的状态,转换相位和具体计时信息等。V2X能够为自动驾驶车辆提供一些额外信息,比如天气状况、道路交通事故,道路施工及道路使用者。

许多先进的应用或者用例依赖于V2X通信,例如:车辆编队行驶、远程驾驶^[55]。在这些用例下,一些车辆依赖V2X消息进行控制。应考虑V2X的预期功能安全(SOTIF)问题。

可以在扩展的整车层面上分析应用V2X的系统的预期功能安全,包括车外要素(信息源和通信资源)。在这种情况下,V2X系统被视为系统的一个复杂传感器(示例见图D.6)。

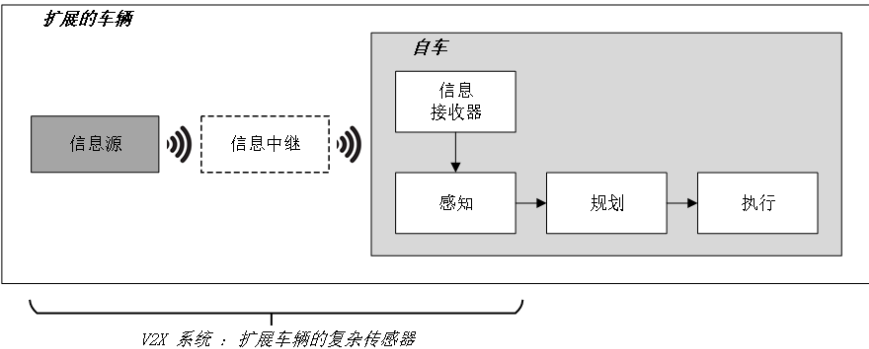


图 D.5 作为扩展车辆一部分的 V2X 系统的示例

如果V2X被用作安全关键功能的一部分，D.4详细介绍了V2X使用应该考虑的几个方面。

D.4.2 V2X通信的规范定义和设计

- V2X 系统要求：
 - 延时要求；
 - 可靠性要求；
 - 互操作性要求。
- V2X 数据要求：
 - V2X消息中包含的物体或事件的准确率要求：数据被认为正确或真实的程度(包括位置和时间精度)；
 - V2X消息中包含的物体或事件的完整性要求：确保数据元素没有损坏；
 - 精确度要求：平均值的标准差；
 - 分辨率要求：两个相邻值的最小差值；
 - 可追溯性要求：跟踪质量满足情况的能力；
 - 一致性要求：符合互操作性标准，以及基本通信要求、基本系统要求、防护要求、安全要求、信任和保证等级要求；
- V2X 消息在整车层面的功能应用；
- 已知 V2X 的局限(例如，超出道路设施的覆盖范围，其他设备的干扰)。

D.4.3 V2X的预期功能安全实施

V2X的预期功能安全(SOTIF)主要关注由于V2X性能局限而导致V2X消息不正确的问题,例如数据的不准确或过时。可以通过V2X消息和自车感知的校验来检测问题,以判断V2X消息的实时性和准确性。

不同应用中使用的不同类型的V2X消息。这些不同类型的V2X消息可能会有对延迟/可靠性/更新频率等有不同的要求。

V2X的类型可以根据信息内容变更频率来分类：

- 更新内容很少的静态 V2X 消息消息(例如交通标志每周更新)；
- 在数小时内更新的半动态信息(例如事故、天气状况)；
- 实时更新的动态信息(例如交通信号灯状态、车辆编队中车辆的动态行为)。

因此，系统规范可以规定不同类型V2X消息的延迟、可靠性和/或更新频率要求，以满足功能/系统安全分析规定的公差。

D.5 感知系统性能目标量化与常见传感器性能局限举例

态势感知对于实现先进的辅助驾驶或者自动驾驶功能至关重要。态势感知主要依赖于感知系统。感知系统通常由一系列传感器构成。充分考虑不同类型传感器的性能局限，定义合理的感知系统性能目标对于保障驾驶自动化系统安全很有必要。

感知系统性能目标量化应该考虑以下基本要求：

- 感知系统应具备与系统 ODD 下最大运行速度相匹配的感知范围(包含距离)；
- 感知系统应具备全向视野范围；
- 对于 ODD 内合理可预见的物体，尤其是关键物体(例如，行人，车辆，两轮车，摩托车等常见交通参与者)，感知系统应能进行识别和区分；
- 当感知系统对物体识别的置信度下降时(例如，受天气等影响或不同类型传感器输入冲突)，应具备整车级别的 SOTIF 安全策略(例如，请求接管、降速、降级等)。
- 在开放道路上，感知系统(可包括地图)应具备识别交通标识、标线和信号的能力，以遵守交通规则。

驾驶自动化系统常见的传感器类型包括视觉传感器、毫米波雷达、激光雷达、超声波传感器等。其中视觉传感器也可用于对车内驾驶员状态进行监督。其中不同类型传感器的典型性能局限(非详尽)如下表所示：

表 D.7 不同类型传感器性能局限举例

传感器类型	功能局限
视觉传感器	容易受到光照条件影响(暗光、眩光等)； 图像模糊、失真或曝光过度(例如，由于曝光不当、车辆高速运动等原因)等； 位置估计或者深度估计可能不准确； 目标检测和分割可能不准确(例如，AI算法)。
毫米波雷达	容易受到多径效应(例如在隧道中)或路面金属(例如，易拉罐、交通标识牌)影响； 角度分辨率低； 对某些波形，可能无法区分在相同距离上的多个物体的速度； 垂直向无分辨率或分辨率低； 速度只能在径向上直接观察到； 容易受到其他毫米波雷达的干扰； 对静止物体物检测能力弱。
激光雷达	容易受到雨、雪、灰尘、雾天或水雾等影响； 扫描样式密度不足或者覆盖不足； 道路上的轮胎等黑暗物体返回的强度信号低； 容易受到其他激光雷达脉冲的干扰。

传感器类型	功能局限
超声波传感器	容易受到气流、空气湿度、温度影响； 垂直方向分辨率低； 角分辨率低、精度低且非常有限感知范围。

D.6 OTA 更新的 SOTIF 考虑

由于验证和确认过程几乎不可能遍历所有场景，另外，运行环境也在不断发生变化。因此在运行阶段可能会识别到新的触发条件和功能不足。如果这些触发条件和功能不足导致的风险不可接受，则需要快速响应从而消除或减少这些风险。例如，可通过更新地图数据、更新机器学习算法和ODD的变更等措施来应对。这些都依赖于空中更新(OTA)技术，OTA提供了一种灵活、方便的方式来快速的修正新识别功能不足。但是需要关注的是，在OTA过程中，需要合适的更新策略和流程，否则可能会引入额外的风险。

OTA更新需要在第5章规范定义和设计章节进行必要定义，可能的方面包括：

- OTA 更新策略；
 - 对新识别到的触发条件和功能不足进行分类；
 - 对新识别的触发条件和功能不足的影响进行分析；
 - OTA更新过程和更新后对其他车辆功能的影响分析；
 - OTA更新后功能的安全运行。
- OTA 更新的条件；
 - 车辆状态；
 - 功能状态。

示例：在 OTA 更新过程中，部分车辆功能(例如，驾驶功能)可能不可用，因此，车辆应该停在不妨碍其他交通流正常通行的地方进行升级。

- OTA 更新中的 HMI 考虑；
 - OTA更新需要获得用户同意；
 - OTA更新前需要提供充分信息给用户，从而避免用户误用。例如，OTA更新的目的、可能受影响的车辆功能、更新进度和预期结束时间等。

附录 E

（资料性）

风险接受准则示例

E.1 概述

据GB/T 34590标准，不合理风险被定义为按照现行的安全观念，被判断为在某种环境下不可接受的风险。一种可能的对“现行安全观念”的解释来自联合国157号法规。其中“不合理风险”被定义为：与有经验且专注的人类驾驶员相比，对驾驶员，乘客和其他道路参与者的整体风险水平升高了。

根据该定义，一种可行的方法是从ADS系统在ODD内安全性能的角度来建立风险接受准则。即：ADS系统（包括后援用户，如必要）在其ODD内合理可预见的场景下需要具备与有经验且专注的驾驶员相同或者更好的安全性能。

风险接受准则可以在以下至少一个层面进行量化：

一整车行为层面：通过与有经验且专注的驾驶员比较，论证ADS系统（包括后援用户，如必要）可以应对ODD内合理可预见的场景，从而确认残余风险可接受。

示例1：ADS系统操作车辆制动系统（例如，0.45g）减速，从而避免与前车的追尾碰撞。

注1：对有经验且专注的驾驶员（例如，有能力以0.45g的合理制动避免与前车的追尾碰撞）的刻画是需要的，从而用于判定风险是否可接受。

示例2：当ADS系统发出接管请求时，后援用户在合理的时间内获取驾驶权，从而继续执行动态驾驶任务。

注2：对于后援用户能力的举证可以被用于论证风险接受准则是否达成。

注3：安全策略（例如，最小风险制动，警告和降级、接管请求、远程辅助等）可以被用于论证风险接受准则是否达成。

注4：ODD边界的风险也应该被考虑。

一要素层面：ADS系统（后援接管用户，如必要）在其ODD内合理可预见的场景下，需要具备与有经验且专注的驾驶员相同或者更好的性能（包括态势感知性能、决策规划性能和车辆操作性能），从而确认参与风险可接受。

示例3：论证决策系统风险可接受的一个论据可以是ADS与其他道路参与者保持了合理的安全距离或安全时间间隔。其中与有经验且专注的驾驶员相比，即不保守，也不激进。

示例4：论证感知系统风险可接受的一个论据可以是在ODD内ADS感知系统对于交通信号灯识别的准确率至少与有经验且专注的驾驶员相同。

注5：安全策略（例如，最小风险制动，警告和降级、接管请求、远程辅助等）可以被用于论证风险接受准则是否达成。

注6：ODD边界的风险也应该被考虑。

E.2 风险接受准则示例

本节提供了如何建立风险接受准则基准的示例方法。具体的，提供了一个在交叉路口推导合理安全时间间隔的示例。通过与有经验且专注的驾驶员相比，避免ADS系统决策不合理。

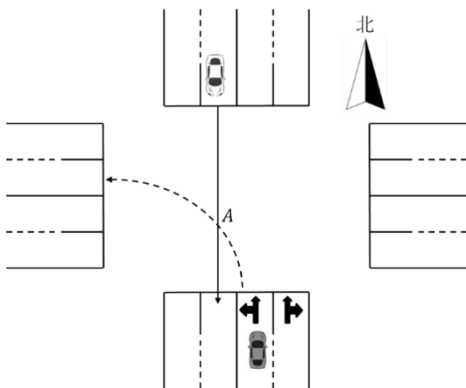


图 E. 1 无保护左转示意图

考虑如图E. 1所示的无保护左转场景，其中自车（灰色车）在左转车道向北运动，期望在无信号灯的路口左转。自车需要考虑由北向南行驶的对向来车（称为他车，白色车）的风险。自车和他车到达轨迹交汇点A的时间分别记为T1和T2，则此时，有经验且专注的驾驶员需要避免如下不安全情形(见图E. 2)：

- 情形 1：自车未礼让，自车先到达交汇点，没有为他车预留充足的时间间隔。即， $T2 - T1 \leq X1$ ；
- 情形 2：自车礼让，使得他车首先经过交汇点，但是距离他车时间间隔过近。即， $T1 - T2 \leq X2$ ；



图 E. 2 不同时间间隔下的安全性

假定T1和T2都在5s以内，对人类驾驶员通过十字路口的时间间隔统计结果见图E. 3和图E. 4。

当自车先通过交汇点时，此时，参数X1的累积分布函数如图E. 3所示，即给出了当自车首先通过交汇点时，为他车预留的时间间隔的分布。其中该时间间隔取决于当自车通过交汇点时刻点时他车的速度以及他车与交汇点的距离。

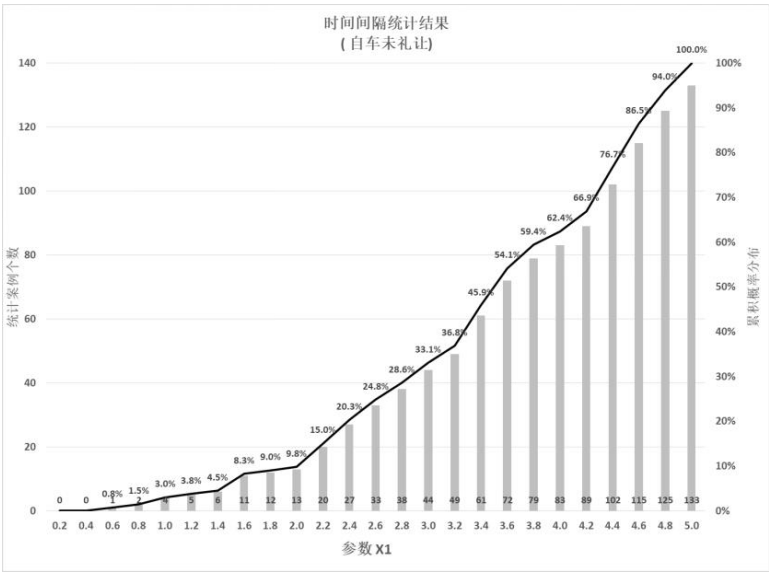


图 E. 3 参数 X1 的分布

基于对有经验且专注驾驶员的定义，可以获得一个合理的参数 x_1 ，即自车应当避免通过十字路口当为他车预留的时间间隔 $\leq x_1$ 。

当他车先通过交汇点时，此时，参数 x_2 的累计分布函数如图E.4所示。即给出了当自车礼让他车，使得他车先经过交汇点时刻，自车与他车保持的时间间隔分布。同样的，也可以获得一个合理的参数 x_2 ，即当自车应避免与他车保持时间间隔 $\leq x_2$ 。

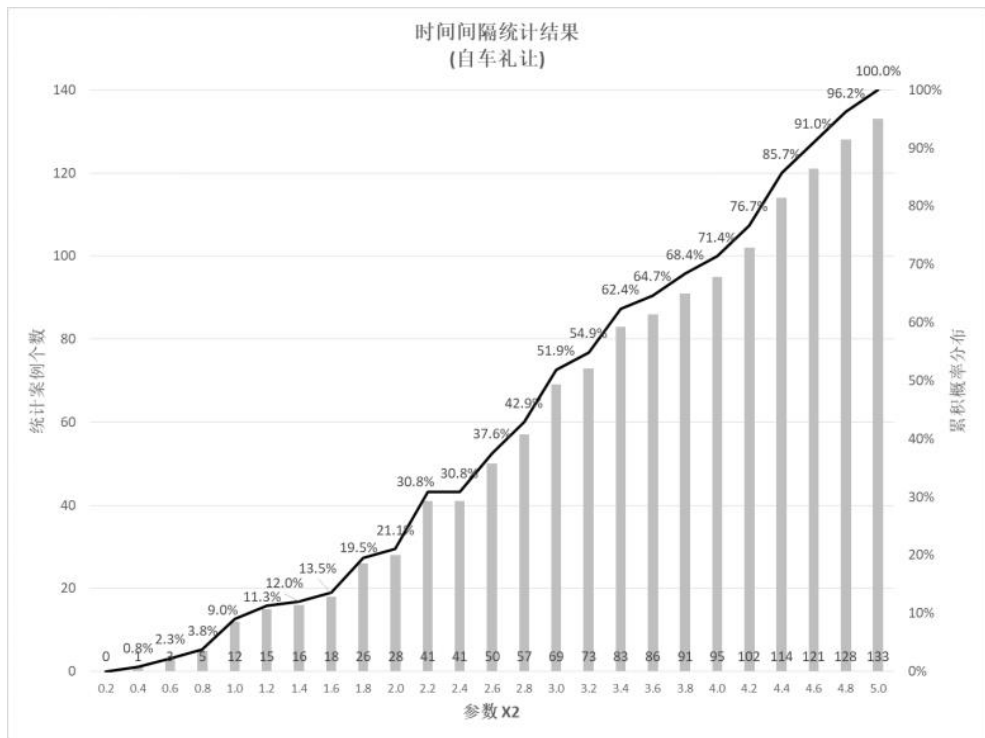


图 E.4 参数 x_2 的分布

通过合理的定义参数 x_1 和 x_2 ，可以帮助ADS系统设计，避免ADS系统决策不合理。因此，可以论证：由于与有经验且专注的驾驶员具有类似的决策逻辑，从而不存在不合理风险。

考虑到现实驾驶场景自车与他车并不总是匀速运动，因此两车到达交汇点的时间间隔是动态变化的，并且自车与他车可能存在一定博弈，但是驾驶决策的逻辑是类似的。

需要说明的是本节为了说明相应的概念，进行了充分的简化。在实际决策中，还需要考虑其他影响因素。例如，他车是否按照预测的速度、轨迹进行运动、自车的执行系统是否会按照ADS系统要求进行运动控制、场景中是否存在弱势交通参与者、天气、路面状况等其他环境因素等。

参 考 文 献

- [1] COMMISSION RECOMMENDATION of 22 December 2006 on safe and efficient in-vehicle information and communication systems: update of the European Statement of Principles on human machine interface (2007/78/EC): <http://data.europa.eu/eli/reco/2007/78/oj>
- [2] Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, SAE Recommended Practice J3016_201806, https://www.sae.org/standards/content/j3016_201806
- [3] ULBRICH, S., MENZEL, T., RESCHKA, A., SCHULDT, F. and MAUER, M. "Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving", 2015 IEEE 18th International Conference on Intelligent Transportation Systems (ITSC), <https://doi.org/10.1109/ITSC.2015.164>
- [4] CENELEC-Standard EN 50126-2:2017 Clause A.1 (RAMS)
- [5] ISO 34502 ("Road vehicles – Engineering framework and process of scenario-based safety evaluation")
- [6] Statistics and data about reported accidents and casualties on public roads in Great Britain (STATS19), UK Department for Transport, <https://www.gov.uk/government/collections/road-accidents-and-safety-statistics>
- [7] German In-Depth Accident Study (GIDAS), accident data collection project in Germany, <https://www.gidas.org/en/willkommen/>
- [8] NASS General Estimates System (GES), US Department of Transportation, <https://www.nhtsa.gov/national-automotive-sampling-system/nass-general-estimates-system>
- [9] CARE database (Community database on Accidents on the Roads in Europe), https://ec.europa.eu/transport/road_safety/specialist/observatory/methodology_tools/about_care_en
- [10] IGLAD (Europe), <http://www.iglad.net/>
- [11] Code of Practice for the design and evaluation of ADAS, EU Project RESPONSE 3; http://www.acea.be/uploads/publications/20090831_Code_of_Practice_ADAS.pdf
- [12] DIN SAE SPEC 91381:2019, "Terms and Definitions Related to Testing of Automated Vehicle Technologies"
- [13] KUHN, D. S., KACKER, R. N. and LEI, Y., "Combinatorial testing", NIST report, June 25, 2012, <https://www.nist.gov/publications/combinatorial-testing>
- [14] KELLY, T. AND ROB WEAVER, R. "The Goal Structuring Notation - A Safety Argument Notation", <https://wwwusers.cs.york.ac.uk/tpk/dsn2004.pdf>
- [15] Stellet, J.E., Brade T., Poddey A., Jesenski S. and Branz, W. "Formalisation and algorithmic approach to the automated driving validation problem", 2019 IEEE Intelligent Vehicles Symposium (IV), <https://doi.org/10.1109/IVS.2019.8813894>
- [16] SHAPPELL, S.A. and WIEGMANN, D.A., The Human Factors Analysis and Classification-System - HFACS, February 2000 Final Report. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

[17]HARTJEN, L., PHILIPP, R., SCHULDT, F., HOWAR F. AND FRIEDRICH B. " Classification of Driving Maneuvers in Urban Traffic for Parametrization of Test Scenarios " in: 9. Tagung Automatisiertes Fahren, Lehrstuhl für Fahrzeugtechnik mit TÜV SÜD Akademie: <https://mediatum.ub.tum.de/1535131>.

[18]BSI PAS 1883:2020, AVSC Best Practice for Describing an Operational Design Domain

[19] LEVESON, N. *Engineering a Safer World - Systems Thinking Applied to Safety*. MIT Press, Cambridge, Massachusetts, USA 2011

[20] LEVESON, N., AND THOMAS, J. *STPA-Handbook*. 2018. Available for download at psas.scripts.mit.edu/home/get_file.php?name=STPA_handbook.pdf

[21]ABDULKHALEQ, A. ET AL A Systematic Approach Based on STPA for Developing a Dependable Architecture for Fully Automated Driving Vehicles, 4th European STAMP Workshop 2016, Procedia Engineering, 179, 41-51, 2017
<https://www.sciencedirect.com/science/article/pii/S1877705817312109>

[22]ABDULKHALEQ, A. ET AL Using STPA in Compliance with ISO 26262 for Developing a Safe Architecture for Fully Automated Vehicles. arXiv preprint arXiv:1703.03657, 2017.

[23]ABDULKHALEQ, A AND WAGNER, S. AND LEVESON, N. *A Comprehensive Safety Engineering approach for Software-Intensive Systems Based on STPA*. Procedia Engineering, 128:2 - 11, 2015, https://www.researchgate.net/publication/265508075_Experiences_with_Applying_STPA_to_SoftwareIntensive_Systems_in_the_Automotive_Domain

[24]SABALIAUSKAITE, G. AND SHEN LIEW, L. AND CUI, J. *Integrating Autonomous Vehicle Safety and Security Analysis Using STPA Method and the Six-Step Model*. International Journal on Advances in Security, 11(1&2):160 - 169, 2018.

[25]ABDULKHALEQ, A. ET AL A Systematic Approach Based on STPA for Developing a Dependable Architecture for Fully Automated Driving Vehicles

[26]FABRIS, S., PRIDDY, J. and HARRIS, F., "Method for Hazard Severity Assessment for the Case of Unintended Deceleration", presented at 2012 VDA Auto SYS conference in Berlin.

[27]PIAO, J., and MCDONALD, M. 'Low speed car following behaviour from floating vehicle data'. IEEE IV2003 Intelligent Vehicles Symposium.

[28]ALLEN, R., MAGDALENO, R., SERAFIN, C., ECKERT, S. ET AL.. "Driver Car Following Behavior Under Test Track and Open Road Driving Condition," SAE Technical Paper 970170, 1997, <https://doi.org/10.4271/970170>

[29]NHTSSA Traffic Safety Facts
2015, <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812384>

[30]FABRIS, S., PRIDDY, J. and HARRIS, F., "Method for hazard severity assessment for the case of undemanded deceleration.", Presented at VDA Automotive SYS Conference, Berlin, June 19/20, 2012, https://www.researchgate.net/publication/344452155_Method_for_hazard_severity_assessment_for_Method_for_hazard_severity_assessment_for_the_case_of_undemanded_deceleration_-_Simone_Fabris.

[31] LITTLEWOOD B. and WRIGHT, D., "Some Conservative Stopping Rules for the Operational Testing of SafetyCritical Software", IEEE Trans. SW Engng., 23(11), 673-683, Nov. 1997

- [32]SIPOC - Wikipedia, <https://en.wikipedia.org/wiki/SIPOC>
- [33]HIRSENKORN, N., KOLSI, H., SELMI, M., SCHÄRMANN, A., HANKE, T., RAUCH, A., RASSHOFER, R., BIEBL, E.: Learning Sensor Models for Virtual Test and Development. 11. Workshop Fahrerassistenzsysteme und automatisiertes Fahren, UniDAS, Walting, 2017
- [34]E. DE GELDER AND J. P. PAARDEKOOPER, “Assessment of Automated Driving Systems using real-life scenarios,” IEEE Intell. Veh. Symp. Proc., no. IV, pp. 589–594, 2017.
- [35]Functional Mockup Interface, <http://functional-mockup-interface.org/>
- [36]ASAM OpenDRIVE, <http://www.asam.net/standards/detail/opendrive/>
- [37]ASAM OpenCRG, <http://www.asam.net/standards/detail/opencrg/>
- [38]ASAM OpenSCENARIO, <http://www.asam.net/standards/detail/openscenario/>
- [39]Open Simulation Interface (OSI), <https://github.com/OpenSimulationInterface>
- [40]Navigation Data Standard, <https://www.nds-association.org/>
- [41]CityGML, <http://www.opengeospatial.org/standards/citygml>
- [42]VAICENAVICIUS, J., WIKLUND, T., GRIGATE, A., VYSNIAUSKAS, I., and KEEN, S. D. ‘Self-driving car safety quantification via component-level analysis’. SAE International Journal of Connected and Automated Vehicles, Volume 4, Issue 1, 2021.
- [43]SHALEV-SCHWARZ S., SHAMMAH S., SHASHUA A., On a Formal Model of Safe and Scalable Self-driving Cars <https://arxiv.org/abs/1708.06374v6>
- [44]NISTÉR D., LEE H.-L., NG J., WANG Y., An Introduction to the Safety Force Field, <https://www.nvidia.com/content/dam/en-zz/Solutions/self-driving-cars/safety-force-field/an-introduction-to-the-safety-force-field-v2.pdf>
- [45]FRAADE-BLANDAR L, BLUMENTHAL M. S., ANDERSON J. M. KALRA N. - RAND: Measuring Automated Vehicle Safety - https://www.rand.org/content/dam/rand/pubs/research_reports/RR2600/RR2662/RAND_RR2662.pdf
- [46]KENDALL, A. and GAL, Y., “What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?”, NIPS 2017.
- [47]PHAN, B., KHAN, S., SALAY, R. and CZARNECKI, K., “Bayesian Uncertainty Quantification with Synthetic Data”. WAISE 2019.
- [48]KOOPMAN, P. and WAGNER, M., “Autonomous Vehicle Safety: An Interdisciplinary Challenge,” IEEE Intelligent Transportation Systems Magazine, Special Issue on SSIV, 2017, in press Vol. 9 #1, Spring 2017, pp. 90–96
- [49]MOLNAR, C., “A Guide for Making Black Box Models Explainable, 2021, <https://christophm.github.io/interpretable-ml-book/>
- [50]ZHANG, Q. AND ZHU, S.-C., “Visual Interpretability for Deep Learning: a Survey”, 2018, <https://arxiv.org/abs/1802.00614>
- [51]LAPUSCHKIN, S., WÄLDCHEN, S., BINDER, A., MONTAVON, G., SAMEK, W., and MÜLLER, K.-R., “Unmasking Clever Hans predictors and assessing what machines really learn”, 2019, In: Nature Communications 1096 (2019), <https://www.nature.com/articles/s41467-019-08987-4>
- [52]D.4 U.S. Department of Transportation. (Jul.2017). vehicle-to-vehicle communication technology. [Online]. Available:https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/v2v_fact_sheet_101414_v2a.pdf

[53] S. TSUGAWA, S. JESCHKE, and S. E. SHLADOVER, “A Review of Truck Platooning Projects for Energy Savings”, IEEE Transactions on Intelligent Vehicles, vol. 1, no. 1, 2016

[54] J. WANG, J. LIU, and N. KATO, “Networking and communications in autonomous driving: A survey”, IEEE Communications Surveys & Tutorials, vol. 21, no. 2, Q2, 2019.

[55] 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Enhancement of 3GPP support for V2X scenarios; Stage 1 (Release 16) 3GPP TS 22.186 V16.2.0 (2019-06).

[56] ISO 21448 (FDIS) Road vehicles — Safety of the intended functionality
