



HAL
open science

Evaluation des cas d'usages des véhicules automatisés et connectés : Vers une approche basée sur les scénarios visant à réduire la quantité de tests en simulation ou environnement réel

Hugues Blache, Pierre-Antoine Laharotte, Nour-Eddin El Faouzi

► To cite this version:

Hugues Blache, Pierre-Antoine Laharotte, Nour-Eddin El Faouzi. Evaluation des cas d'usages des véhicules automatisés et connectés : Vers une approche basée sur les scénarios visant à réduire la quantité de tests en simulation ou environnement réel. 20èmes Rencontres des Jeunes Chercheurs en Intelligence Artificielle, Jun 2022, Saint-Etienne, France. hal-03765371

HAL Id: hal-03765371

<https://hal.science/hal-03765371v1>

Submitted on 31 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Evaluation des cas d'usages des véhicules automatisés et connectés : Vers une approche basée sur les scénarios visant à réduire la quantité de tests en simulation ou environnement réel

Hugues Blache¹, Pierre-Antoine Laharotte¹, Nour-Eddin El Faouzi¹

¹Univ Lyon, Univ Gustave Eiffel, ENTPE, LICIT-Eco7, F-69675 Lyon, France

Résumé

Avec le développement des Véhicules Automatisés et Connectés (VAC), se pose la question de la validation et de l'homologation de ces nouveaux outils. La quantité de tests à réaliser pour évaluer les cas d'usages relatifs aux VACs est conséquente. Dans cet article, nous esquissons une démarche de réduction du nombre de cas d'usage visant à rassembler ceux faisant appel à des scénarios similaires. Pour ce faire, en s'appuyant sur les définitions issues de la littérature, nous identifions les cas d'usages existants, les caractérisons par des variables, puis exploitons ces variables pour rattacher les cas d'usage à un ensemble de scénarios.

Mots-clés

Cas d'Usage, Scénario, Véhicule Automatisé et Connecté, Réduction, Classification.

Abstract

With the development of Automated and Connected Vehicles (CAVs), the question of validation and approval of these new tools arises. The quantity of tests to be carried out to evaluate the use cases relating to CAVs is considerable. In this article, we outline an approach to reduce the number of use cases by gathering those involving similar scenarios. To do this, based on definitions from the literature, we identify existing use cases, characterise them by variables, and then use these variables to link the use cases to a set of scenarios.

Keywords

Use case, Scenario, Automated and Connected Vehicles, Clustering.

1 Introduction

Afin de couvrir un ensemble de conditions environnantes suffisamment vastes pour inclure les principales situations à risque, il est estimé qu'il faudrait environ 275 millions de miles [19] pour valider la fiabilité d'une voiture autonome (ou Véhicule Automatisé et Connecté - VAC). Concrètement, ceci nécessiterait la circulation en continu pendant 12,5 ans de 100 véhicules autonomes à une vitesse de 25mph (soit environ 40 km/h). **Une telle durée d'évaluation, excédant la décennie, serait-elle acceptable pour**

les pouvoirs publics ou les constructeurs automobiles?

Cette latence aurait en plus des répercussions sur l'intégration de toute nouvelle fonctionnalité aux VAC. En effet, chacune des nouvelles fonctionnalités devrait justifier du même niveau d'évaluation. Ces évaluations, qui ont pour but de déterminer le nombre de kilomètres à parcourir jusqu'à la validation ou l'émergence du prochain accident, sont communément appelées "approches basées sur la distance" (*Distance-based Approach* en anglais) [27]. Néanmoins, une autre approche, plus prometteuse et parcimonieuse, émerge dans la littérature relative à l'évaluation des VACs. Cette approche est "basée sur les scénarios" (*Scenario-based Approach* en anglais). [27]. Le but de cette approche est de drastiquement réduire la quantité de tests à réaliser sur le terrain ou en simulation. Elle se base, par exemple, sur le regroupement des scénarios s'appuyant sur des paramètres similaires, puis elle cible les scénarios nécessaires à la validation des systèmes de VAC [7]. C'est cette approche que nous avons décidé de déployer pour nos recherches.

2 Positionnement

La réduction des scénarios et des tests n'est pas propre aux transports, mais est un sujet d'étude multidisciplinaire [5, 8, 22]. La réduction est notamment pleinement ancrée dans ces disciplines dans le but d'économiser le temps et le coût tout en s'appuyant sur des analyses fines d'un système. Plusieurs approches existent dans la littérature relative aux VACs :

- **Les plans d'expériences** : Cette approche a pour but de prédire les réponses d'un système depuis un minimum d'expériences réalisées. Pour cela, des modèles mathématiques sont mis en place selon le nombre de facteurs (variables) pris en compte et selon leurs niveaux (valeurs) [16]. Dans le domaine du transport, cette approche est employée par [4].
- **Les plans d'échantillonnages** : Cette approche permet de tirer, à partir d'une population, un sous-ensemble représentatif de cette population. Ces plans d'échantillonnage parcimonieux visent à définir combien d'observations sont nécessaires pour décrire un phénomène [2]. On retrouve cette pro-

blématique dans le domaine du transport [3].

- **La recherche de cas nominaux** : Cette approche vise à cibler les cas nominaux et permet, dans un premier temps, de visualiser les performances du système dans des situations "usuelles". D'après le domaine de l'ingénierie du logiciel [5], il existe plusieurs techniques pour éliminer les tests. L'une d'elles, la classification non-supervisée [18, 20] est intéressante à prendre en compte pour l'évaluation des VACs.
- **La recherche des scénarios critiques** : Cette approche vise à cibler les scénarios considérés comme critiques, *i.e.* les scénarios ayant une probabilité élevée d'aboutir à un accident [21, 27]. Il existe différentes manières et indicateurs pour identifier les cas critiques. Au sein du projet Pegasus¹, les indicateurs s'appuient sur les valeurs du time-to-collision (TTC) ou du time-to-collision-speed (TTC_VCOL).
- **La décomposition fonctionnelle du système** : Cette démarche a pour but de décomposer des fonctions ou des systèmes complexes en "sous-systèmes" moins complexes. Pour les VAC, d'après le Graads [6], ces systèmes peuvent être décomposés en 6 couches. Cependant, distinguer problématique des transports et problématiques des télécommunications/conception du véhicule est difficile dans ce cas.

À ce jour, il est à noter l'existence de quelques travaux exploratoires visant à l'identification et/ou la réduction des scénarios [28, 34]. Cependant, d'après nos connaissances, il n'existe pas, dans la littérature, de procédé d'analyse et de réduction drastique d'une multitude de cas d'usage des VACs en vue de leur évaluation. Le but est donc d'identifier et regrouper les cas d'usages sollicitant les mêmes scénarios. La démarche innovante de notre approche propose, à la fois, de réduire les scénarios à tester, mais également de réduire en amont le cas d'usage à analyser pour proposer des plans d'expériences exhaustifs. Ces plans d'expériences tenteront de prendre un maximum de cas d'usage existant. Avant d'aborder la démarche adoptée dans la section 4, il est important de poser des définitions pour la compréhension du document. Ces définitions sont expliquées dans la section 3.

3 Définitions élémentaires

La figure 1 illustre les définitions élémentaires de nos recherches. On remarque alors que le cas d'usage englobe le tout, dont le système à évaluer (*i.e.* le véhicule automatisé, ou véhicule *ego*). Le cas d'usage constitue la donnée d'entrée alimentant la démarche de réduction. Ainsi, la notion de **cas d'usage** englobe :

1. Le **système d'étude** de par la définition de la **plage fonctionnelle**.
2. Un **comportement désiré** du cas d'usage.

3. Les **limites de fonctionnalités**.

4. Un ensemble de **scénarios**, eux-mêmes, composés de **scènes**.

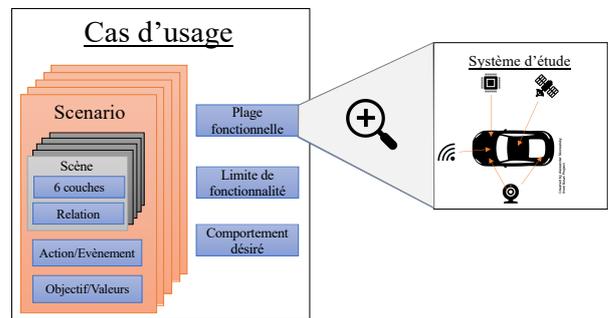


FIGURE 1 – Définition générale, figure inspirée de [13, 32]

3.1 Scène

Une **scène** décrit un "instantané de l'environnement, y compris le paysage et les éléments dynamiques, ainsi que les représentations de soi de tous les acteurs et observateurs (compétences et aptitudes (*c.-à-d.* champs de vision ou occlusion)), et les relations entre ces entités. Seule, une représentation scénique dans un monde simulé peut être globale (scène objective, vérité, terrain). Dans le monde réel, la scène est incomplète, incorrecte, incertaine, et d'un ou plusieurs points de vue des observateurs (scène subjective)" [32].

Une scène est composée de "scenario" et d'éléments dynamiques. Pour les représenter, nous adoptons la décomposition en 6 couches, sans priorisation, explicitée par [10, 26, 29] et détaillée sur la figure 2. Pour illustrer la décomposition en 6 couches des scènes, on peut faire une allégorie à l'hamburger. Schématiquement, un hamburger est composé de différentes couches : du pain, de la sauce, une salade, un steak, du fromage, de la tomate puis du pain. Il est tout à fait possible de remplacer le steak de viande par une galette végétale ou de remplacer la tomate Roma par de la cœur de bœuf, cela ne changera pas la définition même de l'hamburger. Cette notion de changement est ainsi applicable dans la notion de scène, on peut fixer les valeurs des couches et changer une valeur de chaque couche pour faire varier les scènes.

3.2 Scénario

Un **scénario** est une "séquence de scènes ordonnées temporellement. Chaque scénario commence par une scène initiale et finit par une scène finale. Des actions et des événements, ainsi que des objectifs et des valeurs peuvent être spécifiés pour caractériser le déroulement temporel d'un scénario. Contrairement à une scène, un scénario s'étend sur un certain laps de temps." [32]. Les éléments du scénario sont décrits sur la Tab 2.

sp Pour se raccrocher au concept de validation d'un système de VAC et être en adéquation avec le modèle en V^2 ,

1. <https://www.pegasusprojekt.de/en/pegasus-method>

2. https://fr.wikipedia.org/wiki/Cycle_en_V

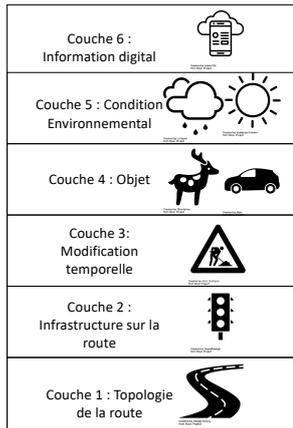


FIGURE 2 – Répartition en couches dans une scène, inspirée de [10,29]

TABLE 1 – Définition des éléments des scènes

Nom	Définition
Véhicule <i>ego</i>	le Véhicule, lui-même, le conducteur et les options d'automatisation ou connectivité.
Objet dynamique et acteur	Objet qui est en mouvement dans le scénario. Un acteur est un objet dynamique de la scène.
Objet statique et paysage	Élément stationnaire et environnement statique

il possible de distinguer trois types de scénarios [25].

- **Scénarios fonctionnels** : *Les scénarios fonctionnels incluent des scénarios d'exploitation au niveau sémantique. Les entités du domaine et les relations de ces entités sont décrites via une notation de scénario linguistique. Les scénarios sont cohérents et un vocabulaire est utilisé pour la description.*
- **Scénarios logiques** : *Les scénarios logiques incluent des scénarios de fonctionnement au niveau de l'espace d'état. Les scénarios logiques représentent les entités et les relations de ces entités à l'aide de plages de paramètres dans l'espace d'état. Les plages de paramètres peuvent éventuellement être spécifiées :*
 - *par des distributions de probabilité ;*
 - *à l'aide de corrélations, ou ;*
 - *de conditions numériques.**. Un scénario logique comprend une notation formelle du scénario.*
- **Scénarios concrets** : *Des scénarios concrets décrivent distinctement des scénarios de fonctionnement au niveau de l'espace d'état. Les scénarios concrets représentent des entités et les relations de ces entités à l'aide de valeurs concrètes pour chaque paramètre dans l'espace d'état*

TABLE 2 – Définition des éléments des scénarios

Nom	Définition
Action & Evènement	Actions effectuées par le véhicules <i>ego</i> Evènement : Action faites par l'environnement de la voiture [12, 13]
Objectif et valeurs	Tâche données aux "conducteur ou à l'automatisme". Cette objectif est présent le long du scénario pour le conducteur ou l'automate est influe sur "les actions de navigation, guidage et de stabilisation" [15]

3.3 Cas d'usage

Le **cas d'usage** est une "*combinaison/description fonctionnelle du système, comprenant la limite du système, le comportement désiré et la plage fonctionnelle du système, ainsi que des scénarios satisfaits par ce système. Ainsi, un cas d'usage est une manière d'utiliser le système pour un usage de l'utilisateur.*" [32].éléments des cas d'usage sont décrits sur la Tab 3

4 Démarche globale adoptée

La démarche globale vise à proposer un procédé de réduction de la quantité de cas d'usage et de scénarios à explorer en vue de valider les nouvelles fonctionnalités embarquées dans les Véhicules Automatisés et Connectés (VACs). Dans un premier temps, nous nous focalisons sur la compréhension du lien entre cas d'usage et la réalisation d'une scène de trafic, que ce soit sur le terrain ou en environnement simulé. L'objectif est de comprendre l'enchaînement des liens logiques conduisant du cas d'usage à une instance de l'expérimentation. Pour ce faire, nous proposons de décrire ce cheminement sous la forme d'un graphe (ou arbre) dépeint en Figure 4. La lecture descendante peut se faire de la manière suivante :

Pour un Cas d'usage (U_c) donné, la validation peut-être opérée sur un set de j scénarios fonctionnels (S_f) avec $j \leq i$ (i étant le nombre de U_c totaux), communément appelé relation 1 à n . Chaque scénario fonctionnel, c'est-à-dire adoptant une forme verbale, peut-être traduite par un set de paramètres se présentant sous la forme d'un scénario logique (S_l) (relation 1 à 1). Outre ces scénarios fonctionnels, on peut identifier d scénarios concrets (S_c) avec $k \leq d$ (relation 1 à n), k étant le nombre total de scénarios logiques. Un scénario concret correspond dès lors à la réalisation d'un

3. Le Green Light Optimal Speed Advice (GLOSA) est un service basé sur un système de communication Infrastructure-to-Vehicles (I2V). Il permet d'envoyer des informations relatives à l'état du feu de circulation (appelées SPATEM en Europe) vers les véhicules connectés en approche. L'objectif de cette communication est de permettre au véhicule connecté de tirer parti des informations relatives au statut des feux de circulation pour l'aider à optimiser sa vitesse d'approche de l'intersection [30].

Couche 6 Technologie sans fil : Moyenne portée Communication: V2X	Couche 6 Portée : [100,1000]m Communication: V2X	Couche 6 Portée 1000 m Communication: V2X
Couche 5 Temps : Ensoleillé Perturbation : Aucune	Couche 5 Temps : Pourcentage de nuage [0,15]% Perturbation : Aucune	Couche 5 Temps : Pourcentage de nuage 0 % Perturbation : Aucune
Couche 4 Véhicules sujets : Véhicules particuliers Autres Agent : 3 véhicules particuliers	Couche 4 Véhicules ego : Dimension [2.5;5]m Autres Agents : Dimension [2.5;5]m	Couche 4 Véhicules ego : Dimension 4 m Autres Agents : Dimension 4 m
Couche 3 Aucun travaux sur la route	Couche 3 Aucun travaux sur la route	Couche 3 Aucun travaux sur la route
Couche 2 Infrastructure: Feux de circulation Limite de vitesse : Urbain	Couche 2 Feux de circulation : Cycle [90;120] s Limite de vitesse : [30,50] km/h	Couche 2 Feux de circulation : Cycle 90 s Limite de vitesse : 50 km/h
Couche 1 Jonction : Intersection en T Route 1 à 3 : Géométrie plane sans courbure	Couche 1 Intersection en T Route 1 à 3 : courbure [0;2]°, largeur : [2;3,5]m	Couche 1 Intersection en T Route 1 à 3 : courbure 0°, largeur : 3,5 m

(a) Scénarios fonctionnels

(b) Scénarios logiques

(c) Scénarios concrets

FIGURE 3 – Description des 3 types de scénarios

scénario logique. Cette distinction entre scénario concret et scénario logique est proche à celle adoptée en statistiques pour distinguer une variable aléatoire et sa réalisation. Finalement, l'ensemble des scénarios concrets sont constitués de p différentes scènes (dernière couche) avec $m \leq p$ (relation 1 à n), m étant le nombre total de scénarios concrets. La démarche inverse, ascendante, est possible en adoptant le même cheminement de pensées, mais en remplaçant les relations 1 à n en n à 1.

Par la suite, les deux cheminements, ascendants et descendants, seront adoptés en vue de s'assurer de l'exhaustivité des scénarios considérés.

4. le développement de plans d'expériences parcimonieux en scénarios à tester (*i.e.* faisant appel à un nombre réduit de scénarios à tester).

4.1 Caractérisation des nœuds du modèle - Identification des cas d'usages, des scénarios fonctionnels, logiques et concrets (étape 1)

Dans cette première étape du procédé, l'objectif est de déterminer et caractériser les nœuds du graphe illustré en Figure 4.

4.1.1 Identification des cas d'usage (U_c)

Cette étape se basera principalement sur la littérature et sur les rapports déjà existants, que ce soit du côté étatique (Scoop [1] , C-Roads [17] , Pegasus ...) que dans les recherches effectuées par des institutions (ETSI, etc.). Il est également envisagé de prendre en considération la génération de cas d'usage à l'aide de mots-clefs, en s'appuyant sur les principes des ontologies [13]. Cette identification se doit d'être la plus exhaustive possible, d'abord en rassemblant un maximum de cas d'usage (même similaires), puis, en identifiant les plages fonctionnelles des systèmes étudiés, les limites des systèmes et les comportements désirés pour lequel le système a été conçu. Une attention particulière sera portée sur les lieux géographiques pris en compte. Nous considérerons une architecture permettant d'intégrer l'ensemble des études, mais les travaux (ex : études de réduction et sélection des cas d'usage) se focaliseront sur les situations européennes, voire françaises. En complément, la couverture adéquate de l'ensemble des scénarios existants par l'approche adoptée reste un point à étudier. De même, il est pertinent de savoir si tous les U_c existants que l'on identifiera recouvrent bien les aspects de la mobilité intelligente. Finalement, au vu du nombre conséquent de U_c s existant, l'étude de la validation lors de la réduction des scénarios ne se focalisera que sur un petit nombre d' U_c .

4.1.2 Identification des scénarios fonctionnels (S_f)

Cette partie dépendra des cas d'usages identifiés et des limites de fonctionnalité prises en compte. Des vocabulaires spécifiques peuvent et doivent être utilisés pour décrire

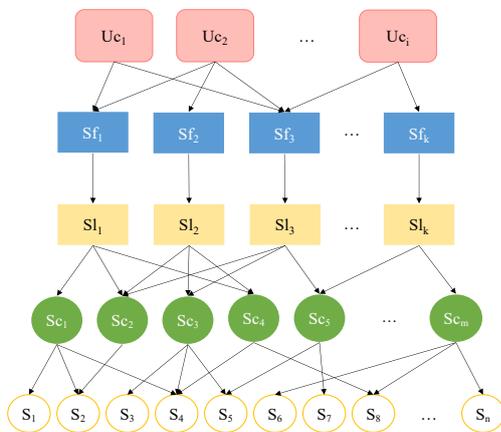


FIGURE 4 – Cheminement hiérarchique de la démarche d'identification des scénarios et des cas d'usage.

La démarche adoptée se décline alors en 4 étapes visant à peupler/renseigner cet arbre :

1. l'identification et la caractérisation des nœuds de l'arbre ;
2. la détermination des connexions entre les nœuds, c'est-à-dire l'établissement des arcs et de leur pondération ;
3. la réduction du graphe de l'arbre en vue de regrouper les cas d'usage similaires ;

TABLE 3 – Définition des éléments des cas d’usages

Nom	Définition
Plage de fonctionnalité du système	Correspond à la plage/gamme de fonctionnalités du système. Par exemple, le système GLOSA ³ évalué ne peut pas communiquer à plus de 1000m.
Limite d’un système	Correspond à la difficulté de réalisation d’une tâche par un système pour lequel il est conçu. Cette notion peut être raliée à celle de Operational Design Domain (ODD). Par exemple, le système de GLOSA pris en compte ne s’applique que sur les intersections munies de feux de circulation.
Comportement désiré	Correspond à l’attente en termes de comportement que l’on a fixé pour le système. Par exemple, pour le système GLOSA, il s’agit de l’adaptation de la vitesse du conducteur selon les phases.

les différentes situations des VACs, plusieurs techniques sont présentées dans la littérature, comme l’utilisation des Langages de Description des Scénarios (SDL) [37], de langage d’ontologie web⁴(OWL) [23] ou via des bases comme OpenScenario. Nous adopterons le OWL à l’aide du logiciel *protégé* et en prenons des termes émis par des normes ODD, comme le PAS 1883 :2020.

4.1.3 Identification des scénarios logiques (S_l)

Il existe plusieurs façons de générer les scénarios logiques :

- **Traduction des scénarios fonctionnels (sémantiques) en scénarios logiques (plages de valeurs de plusieurs indicateurs)**. Cette approche est étudiée dans [24]. En pratique, les mots clefs utilisés dans l’ontologie sont traduits par des plages de paramètres.
- **Rassemblement des scénarios concrets en cluster pour la création de scénarios logiques**. Plusieurs pistes sont à envisager pour le regroupement des scénarios concrets (S_c), mais nous tenterons au maximum de limiter les biais de perception mentale explicités dans la littérature [18, 20]. Une approche de clustering, qui a attiré notre attention, est celle du fuzzy-c-means (ou, de façon générale, les approches de clustering "flou"). Ce type de clustering est en mesure d’affecter à un objet (individu) plusieurs

clusters, contrairement au k-means, qui affecte une unique classe à un objet (individu). Comme illustré par la figure 4, les scénarios concrets sont susceptibles d’appartenir à plusieurs scénarios logiques. En effet, plusieurs scénarios logiques sont susceptibles de couvrir une même plage de paramètres, c’est-à-dire que les scénarios logiques ne sont pas disjoints. Dès lors, les approches basées sur les regroupements "flou" semblent une piste pertinente à explorer.

De plus en s’inspirant de l’article [9], un des points clef dans le choix des données et de leur sélection (détaillée par la suite) est le lieu géographique. En effet, que nous soyons en Amérique du Nord, en Europe (selon les pays) ou même en Chine les lois et les règles routières seront totalement différentes. Pour peu que le sens de circulation change, un système doit être appliqué selon un lieu géographique donné pour éviter tout problème de sécurité et de validation. Un autre élément qui nécessitera une grande attention est le milieu dans lequel les données sont prélevées et observées. Par exemple, le cas des milieux autoroutiers ou urbains répond à des codes et des comportements du trafic différents. Une grande attention doit être portée à ce type de cas par la suite afin d’identifier les limites des cas d’usage et de sélectionner les scénarios qui y répondent. Un autre paramètre à intégrer consiste en l’état du trafic : par exemple, pour un trafic dense, il faut voir le nombre de scénarios qui y répondent pour au final identifier si un pour un cas d’usage donné, la réalisation est possible. Finalement, il est possible que certaines informations ne soient pas disponibles selon les couches des scènes (figure 2). Par exemple, lors des données, les informations relatives aux données météo peuvent venir à manquer. Cependant, quelques articles s’intéressent à la prise en considération des aspects relatifs aux conditions environnementales [14].

4.1.4 Identification / qualification des scénarios concrets (S_c)

Les scénarios concrets dépendront de plusieurs facteurs :

- **Données récoltées** : Comme nous avons vu dans le document précédent, il existe différents types de données disponibles. Nous en avons ciblé quelques-unes :
 - Données Geostationnaire : Ces données récoltées par drone permettraient de définir sur des lieux fixes les occurrences de plusieurs scénarios.
 - Nuscene [11] : Ces données, libres d’accès et recueillies à partir de CAVs, permettraient d’identifier des scénarios inhabituels et non rencontrés dans les bases de données précédemment citées.
- **Début et fin d’un scénario** : Une attention doit-être portée sur ce point pour définir quand commence un scénario et quand il s’achève. Nous supposons que l’attention peut se porter sur les manœuvres des véhicules pour délimiter les scénarios. Cependant, une manœuvre ne se suffit pas à elle-même, dans le sens où la dangerosité de la manœuvre dépend

4. https://fr.wikipedia.org/wiki/Web_Ontology_Language

des conditions initiales : on peut faire la même manœuvre du point de vue du véhicule *ego* dans l'espace et le temps, pour autant, les conséquences ne seront pas les mêmes suivant la position et la densité de véhicules obstacles dans son environnement. D'autre part, il y a également la notion d'enchaînement de plusieurs manœuvres, qui peut conduire à des situations plus ou moins complexes. Dans l'absolu, le partage en manœuvres est une option envisageable car simplificatrice du problème. Il est toujours nécessaire de simplifier le problème, c'est ce qui fait les différences entre les modèles." Ainsi on peut définir le début et la fin d'un scénario comme l'accomplissement de l'objectif du véhicule *ego*. Par exemple, sur la figure 5, le début du scénario serait marquée par le véhicule *ego* sur la route 1 et la fin par le véhicule sur la route 2.

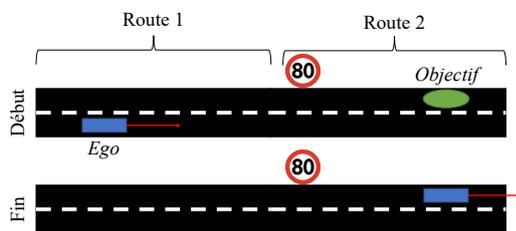


FIGURE 5 – Représentation du début et de la fin d'un scénario

4.2 Connexions entre les nœuds (étapes 2)

Les arcs du graphe (Figure 4) et leur pondération seront déterminés de deux manières distinctes :

- **Pour les arcs localisés entre les nœuds "Cas d'usage" et les nœuds "Scénarios fonctionnels"**, on s'appuiera sur les approches par le haut (Top-Down), développées dans la littérature. Il s'agit pour l'essentiel de se baser sur l'ontologie construite, puis d'établir des correspondances entre le vocabulaire employé pour caractériser le cas d'usage et le vocabulaire à l'œuvre dans le scénario fonctionnel. Pour ce faire, l'établissement d'un arc (lien/connexion) entre un cas d'usage (U_c) et un scénario fonctionnel (S_f) est opéré entre la limite des fonctionnalités du U_c et les couches du scénario fonctionnel.
- **Pour les arcs localisés entre les nœuds "Scénarios concrets" et "Scènes"**, on s'appuiera sur les approches par le bas (Bottom-Up) consistant à regrouper des scènes à partir de données observées ou simulées. Les connexions entre les scènes et les scénarios concrets (Sc) seront établies par des probabilités d'occurrence résultant des observations.

4.3 Réduction des cas d'usage (étape 3)

La démarche envisagée part du constat suivant : Si deux cas d'usages font ressortir un même set de scénarios ou des scénarios similaires, alors la démarche de réduction des

scénarios à étudier devrait être similaire. Il est donc inutile de construire deux plans d'expérience pour chacun. L'effort serait à la fois double et « inutile » car le plan d'expérience développé pour le premier cas d'usage est réutilisable pour le second cas d'usage.

Un procédé de regroupement des cas d'usage s'appuyant sur les approches issues de l'apprentissage automatique et nécessitant le développement d'une métrique particulière devra être mis en oeuvre pour répondre à cette problématique. Cette étape suivra la sélection des scénarios fonctionnels des différents cas d'usage. On identifiera les similarités des différents cas d'usages selon s'ils regroupent un pourcentage de scénarios fonctionnels similaires avec des techniques d'apprentissage non-supervisé. Une piste additionnelle sera d'introduire des métriques entre les mots ou les branches des S_f pour déterminer la similarité entre chaque S_f . Pour ce faire, on déterminera des taux d'importance entre chaque critère moins discriminant que d'autres. Dans le domaine de la sémantique, le regroupement de textes [36] existe. On se penchera sur les notions que l'on peut nommer *clustering-semantic* [35] ou bien sur le *Natural Language Processing Clustering*.

En guise d'illustration, si nous reprenons les cas d'usage : GLOSA et Red Light Violation Warning⁵. Il est fort à parier que les scénarios fonctionnels seront par nature similaires, car les deux cas d'usage font référence à des situations contenant des feux de circulation en milieu urbain. Identifier les scénarios pertinents pour évaluer le cas d'usage GLOSA, permettra d'identifier les scénarios nécessaires d'explorer pour la validation du cas d'usage Red Light Violation Warning.

4.4 Réduction des scénarios : vers des plans d'expérience parcimonieux en scénarios à tester (étape 4)

Au vu d'un nombre élevé de scénario générer par les étapes précédentes, tester tous les scénarios reviendrait à de l'approche par distance, ainsi la nécessité de réduire les scénarios est une voie indispensable. Nous pourrions adopter la technique proposée par Sumaili et al [31], dans le domaine de l'éolien :

- Ils déterminent une distance entre chaque scénario. Cette distance peut être par exemple euclidienne.
- Ils déterminent une valeur seuil (valeurs de tolérance) pour déterminer si deux scénarios sont voisins
- Ils déterminent dans chaque groupe de scénarios, le scénario ayant le plus de voisins comme le scénario représentatif (scénario attractif) de ces voisins
- Ils réitérent la démarche en ne gardant que le scénario représentatif et ils le comparent avec les autres scénarios représentatifs.

L'algorithme s'achève quand l'ensemble des distances interclasses sont inférieures au seuil de tolérance. Cette démarche nécessite une fois encore de développer une mé-

5. Cas d'usage "qui diffuse la phase et la synchronisation du signal (SPat) et d'autres données sur l'appareil embarqué, permettant des avertissements pour les violations imminentes des feux rouges" [33]

trique qualifiant la distance entre deux scénarios, puis de faire appel à des méthodes tirées de l'Intelligence Artificielle pour réduire la quantité de scénarios à explorer.

5 Conclusion

Dans cet article, nous avons défini les notions et développé les jalons d'une démarche visant à aboutir à la production de plans d'expériences parcimonieux en scénarios à explorer, mais couvrant un large spectre de cas d'usages en vue de l'évaluation de nouveaux services VACs. En particulier, le procédé développe une architecture basée sur un graphe logique assurant la génération et qualification de l'ensemble des scénarios à explorer, avant de proposer des procédés de réduction et d'élagage du graphe. La prochaine étape de nos recherches est double.

- Générer les scénarios fonctionnels à l'aide d'une ontologie.
- Identifier les cas d'usage et leurs scénarios fonctionnels.

Une limite à notre démarche est la possibilité d'engendrer de nombreux biais mentaux lors de l'approche descendante s'appuyant sur des ontologies. Pour réduire ce risque, nous tenterons au maximum d'interagir avec différents chercheurs et experts du domaine. Nous nous appuyerons également sur les ressources de données disponibles pour compléter le graphe et assurer l'exploration la plus exhaustive possible des citations.

Références

- [1] C-ITS French Use Cases Catalogue – Functional Description. Technical report, Ministry for an Ecological and Solidary Transition – Directorate General for Infrastructure, Transport and the Sea (DGITM), 2020.
- [2] Guide d'élaboration de plans d'échantillonnage temporel et de reconstitution de données. Technical report, ADEME, 2021.
- [3] Yasuhiro Akagi, Ryosuke Kato, Sou Kitajima, Jacobo Antona-Makoshi, and Nobuyuki Uchida. A risk-index based sampling method to generate scenarios for the evaluation of automated driving vehicle safety. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pages 667–672. IEEE, 2019.
- [4] Ali P Akgüngör and Osman Yıldız. Sensitivity analysis of an accident prediction model by the fractional factorial method. *Accident Analysis & Prevention*, 39(1) :63–68, 2007.
- [5] Marwah Alian, Dima Suleiman, and Adnan Shaout. Test Case Reduction Techniques - Survey. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 7(5), June 2016. Number : 5 Publisher : The Science and Information (SAI) Organization Limited.
- [6] Christian Amersbach and Hermann Winner. Functional decomposition : An approach to reduce the approval effort for highly automated driving. In *8. Tagung Fahrerassistenz*, 2017.
- [7] Christian Amersbach and Hermann Winner. Defining Required and Feasible Test Coverage for Scenario-Based Validation of Highly Automated Vehicles*. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pages 425–430, October 2019.
- [8] Margaret Armstrong, Aziz Ndiaye, Rija Razanatsimba, and Alain Galli. Scenario reduction applied to geostatistical simulations. *Mathematical Geosciences*, 45(2) :165–182, 2013.
- [9] Johannes Bach, Jacob Langner, Stefan Otten, Eric Sax, and Marc Holzäpfel. Test scenario selection for system-level verification and validation of geolocation-dependent automotive control systems. In *2017 International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, pages 203–210, 2017.
- [10] Gerrit Bagschik, Till Menzel, and Markus Maurer. Ontology based Scene Creation for the Development of Automated Vehicles. In *2018 IEEE Intelligent Vehicles Symposium (IV)*, pages 1813–1820, June 2018. ISSN : 1931-0587.
- [11] Holger Caesar, Varun Bankiti, Alex H. Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. nuscenes : A multimodal dataset for autonomous driving. In *CVPR*, 2020.
- [12] Wei Chen. *Formal Modeling and Automatic Generation of Test Cases for the Autonomous Vehicle*. phdthesis, Université Paris-Saclay, September 2020.
- [13] Wei Chen and Leïla Kloul. An Ontology-based Approach to Generate the Advanced Driver Assistance Use Cases of Highway Traffic :. In *Proceedings of the 10th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, pages 75–83, Seville, Spain, 2018. SCITEPRESS - Science and Technology Publications.
- [14] Ying Chen, Jiwon Kim, and Hani S. Mahmassani. Pattern recognition using clustering algorithm for scenario definition in traffic simulation-based decision support systems. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 798–803, October 2014. ISSN : 2153-0017.
- [15] Sebastian Geyer, Marcel Baltzer, Benjamin Franz, Stephan Hakuli, Michaela Kauer, Martin Kienle, Sonja Meier, Thomas Weißgerber, Klaus Bengler, Ralph Bruder, Frank Flemisch, and Hermann Winner. Concept and development of a unified ontology for generating test and use-case catalogues for assisted and automated vehicle guidance. *IET Intelligent Transport Systems*, 8(3) :183–189, 2014. eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1049/iet-its.2012.0188>.
- [16] Jacques Goupy. *Introduction aux plans d'expériences : avec applications*. Dunod, 2013.

- [17] Damaris Anna Gruber and Wolfgang Kernstock. Milestone 30 - Definition use cases and location including a deployment plan for C-ITS elements for all pilot sites. Technical report, C-ROADS, 2020.
- [18] Florian Hauer, Ilias Gerostathopoulos, Tabea Schmidt, and Alexander Pretschner. Clustering Traffic Scenarios Using Mental Models as Little as Possible. In *2020 IEEE Intelligent Vehicles Symposium (IV)*, pages 1007–1012, October 2020. ISSN : 2642-7214.
- [19] Nidhi Kalra and Susan M. Paddock. Driving to safety : How many miles of driving would it take to demonstrate autonomous vehicle reliability? *Transportation Research Part A : Policy and Practice*, 94 :182–193, December 2016.
- [20] Jonas Kerber, Sebastian Wagner, Korbinian Groh, Dominik Notz, Thomas Kühbeck, Daniel Watzenig, and Alois Knoll. Clustering of the Scenario Space for the Assessment of Automated Driving. In *2020 IEEE Intelligent Vehicles Symposium (IV)*, pages 578–583, October 2020. ISSN : 2642-7214.
- [21] Arne Lamm and Axel Hahn. Towards Critical-Scenario Based Testing With Maritime Observation Data. In *2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO)*, pages 1–10, May 2018.
- [22] Zukui Li and Christodoulos A Floudas. Optimal scenario reduction framework based on distance of uncertainty distribution and output performance : I. single reduction via mixed integer linear optimization. *Computers & Chemical Engineering*, 70 :50–66, 2014.
- [23] Till Menzel, Gerrit Bagschik, Leon Isensee, Andre Schomburg, and Markus Maurer. From Functional to Logical Scenarios : Detailing a Keyword-Based Scenario Description for Execution in a Simulation Environment. In *2019 IEEE Intelligent Vehicles Symposium (IV)*, pages 2383–2390, June 2019. ISSN : 2642-7214.
- [24] Till Menzel, Gerrit Bagschik, Leon Isensee, Andre Schomburg, and Markus Maurer. From functional to logical scenarios : Detailing a keyword-based scenario description for execution in a simulation environment. In *2019 IEEE Intelligent Vehicles Symposium (IV)*, pages 2383–2390. IEEE, 2019.
- [25] Till Menzel, Gerrit Bagschik, and Markus Maurer. Scenarios for development, test and validation of automated vehicles. In *2018 IEEE Intelligent Vehicles Symposium (IV)*, pages 1821–1827. IEEE, 2018.
- [26] Demin Nalic, Tomislav Mihalj, Maximilian Bäumler, Matthias Lehmann, and Stefan Bernsteiner. scenario-based testing of automated driving systems : a literature survey. page 11, 2020.
- [27] Stefan Riedmaier, Thomas Ponn, Dieter Ludwig, Bernhard Schick, and Frank Diermeyer. Survey on Scenario-Based Safety Assessment of Automated Vehicles. *IEEE Access*, 8 :87456–87477, 2020. Conference Name : IEEE Access.
- [28] Stefan Riedmaier, Daniel Schneider, Daniel Watzenig, Frank Diermeyer, and Bernhard Schick. Model Validation and Scenario Selection for Virtual-Based Homologation of Automated Vehicles. *Applied Sciences*, 11(1) :35, January 2021. Number : 1 Publisher : Multidisciplinary Digital Publishing Institute.
- [29] Maike Scholtes, Lukas Westhofen, Lara Ruth Turner, Katrin Lotto, Michael Schuldes, Hendrik Weber, Nicolas Wagener, Christian Neurohr, Martin Herbert Bollmann, Franziska Körtke, Johannes Hiller, Michael Hoss, Julian Bock, and Lutz Eckstein. 6-Layer Model for a Structured Description and Categorization of Urban Traffic and Environment. *IEEE Access*, 9 :59131–59147, 2021. Conference Name : IEEE Access.
- [30] Marcin Seredynski, Bernabe Dorronsoro, and Djamel Khadraoui. Comparison of green light optimal speed advisory approaches. In *16th international IEEE conference on intelligent transportation systems (ITSC 2013)*, pages 2187–2192. IEEE, 2013.
- [31] Jean Sumaili, Hrvoje Keko, Vladimiro Miranda, Zhi Zhou, Audun Botterud, and Jianhui Wang. Finding representative wind power scenarios and their probabilities for stochastic models. In *2011 16th International Conference on Intelligent System Applications to Power Systems*, pages 1–6, September 2011.
- [32] Simon Ulbrich, Till Menzel, Andreas Reschka, Fabian Schuldt, and Markus Maurer. Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, pages 982–988, September 2015. ISSN : 2153-0017.
- [33] USDT. Intelligent Transportation Systems - CV Pilot Deployment Program, 2022.
- [34] Lennart Vater, Andreas Pütz, Levasseur Tellis, and Lutz Eckstein. Test Case Selection Method for the Verification of Automated Driving Systems. *ATZelectronics worldwide*, 16(11) :40–45, November 2021.
- [35] Wei Wang, Romaric Besançon, Olivier Ferret, and Brigitte Grau. Regroupement sémantique de relations pour l'extraction d'information non supervisée. In *20eme Conférence sur le Traitement Automatique des Langues Naturelles*, 2013.
- [36] Tingting Wei, Yonghe Lu, Huiyou Chang, Qiang Zhou, and Xianyu Bao. A semantic approach for text clustering using wordnet and lexical chains. *Expert Systems with Applications*, 42(4) :2264–2275, 2015.
- [37] Xizhe Zhang, Siddhartha Khastgir, and Paul Jennings. Scenario Description Language for Automated Driving Systems : A Two Level Abstraction Approach. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 973–980, October 2020. ISSN : 2577-1655.