



HAL
open science

Systèmes multi-agents, auto-organisation et contrôle par apprentissage constructiviste pour la modélisation et la régulation dans les systèmes coopératifs de trafic

Maxime Guériau

► **To cite this version:**

Maxime Guériau. Systèmes multi-agents, auto-organisation et contrôle par apprentissage constructiviste pour la modélisation et la régulation dans les systèmes coopératifs de trafic. Intelligence artificielle [cs.AI]. Université de Lyon I Claude Bernard, 2016. Français. NNT : 2016LYSE1318 . tel-01504421

HAL Id: tel-01504421

<https://hal.science/tel-01504421>

Submitted on 10 Apr 2017

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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



N° d'ordre NNT : 2016LYSE1318

THÈSE DE DOCTORAT DE L'UNIVERSITÉ DE LYON

opérée au sein de
l'Université Claude Bernard Lyon 1

École Doctorale ED512
Informatique et Mathématiques

Spécialité de doctorat : Informatique

Soutenue publiquement le 12/12/2016, par :

Maxime GUÉRIAU

Systemes multi-agents, auto-organisation et contrôle par apprentissage constructiviste pour la modélisation et la régulation dans les systèmes coopératifs de trafic

Devant le jury composé de :

Amal EL FALLAH SEGHRUCHNI, Professeur, LIP6-CNRS – UPMC Présidente du jury
Hani MAHMASSANI, Professeur, Northwestern University – USA Rapporteur
Patrick REIGNIER, Professeur, INRIA-LIG – Université Grenoble-Alpes Rapporteur
Giovanna DI MARZO SERUGENDO, Professeur, ISS – Université de Genève Rapporteuse

Salima HASSAS, Professeur, LIRIS-CNRS – Université de Lyon 1 Directrice de thèse
Romain BILLOT, Maître de Conférences, Lab-STICC-CNRS – Télécom Bretagne .. Co-directeur
Frédéric ARMETTA, Maître de Conférences, LIRIS-CNRS – Université de Lyon 1 . Co-encadrant
Nour-Eddin EL FAOUZI, Directeur de Recherche, IFSTTAR-ENTPE – Univ. Lyon . Co-encadrant



Thèse préparée au sein du laboratoire

LIRIS-CNRS

Laboratoire d'InfoRmatique en Image et Systèmes d'information
UMR 5205 CNRS
Université Claude Bernard Lyon 1
Bâtiment Nautibus
Campus de la Doua
25 avenue Pierre de Coubertin
69622 Villeurbanne Cedex, France

Et au laboratoire

LICIT

Laboratoire Ingénierie Circulation Transports
UMR T_E IFSTTAR-ENTPE
Université de Lyon



IFSTTAR
COSYS/LICIT
25 Avenue François Mitterrand Case 24
Cité des mobilités
69675 Bron Cedex, France

ENTPE
LICIT
Rue Maurice Audin
69518 Vaulx-en-Verin Cedex, France



Cette thèse a été financée par une Allocation Doctorale de Recherche de

La Région Rhône-Alpes

*Dans le cadre du dispositif Communautés de Recherche Académique
ARC 7 : Innovations, Mobilités, Territoires et Dynamiques Urbaines*

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université

Président du Conseil Académique

Vice-président du Conseil d'Administration

Vice-président du Conseil Formation et Vie Universitaire

Vice-président de la Commission Recherche

Directrice Générale des Services

M. le Professeur Frédéric FLEURY

M. le Professeur Hamda BEN HADID

M. le Professeur Didier REVEL

M. le Professeur Philippe CHEVALIER

M. Fabrice VALLÉE

Mme Dominique MARCHAND

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard

Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux

Faculté d'Odontologie

Institut des Sciences Pharmaceutiques et Biologiques

Institut des Sciences et Techniques de la Réadaptation

Département de formation et Centre de Recherche en Biologie Humaine

Directeur : M. le Professeur G.RODE

Directeur : Mme la Professeure C. BURILLON

Directeur : M. le Professeur D. BOURGEOIS

Directeur : Mme la Professeure C. VINCIGUERRA

Directeur : M. X. PERROT

Directeur : Mme la Professeure A-M. SCHOTT

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies

Département Biologie

Département Chimie Biochimie

Département GEP

Département Informatique

Département Mathématiques

Département Mécanique

Département Physique

UFR Sciences et Techniques des Activités Physiques et Sportives

Observatoire des Sciences de l'Univers de Lyon

Polytech Lyon

Ecole Supérieure de Chimie Physique Electronique

Institut Universitaire de Technologie de Lyon 1

Ecole Supérieure du Professorat et de l'Education

Institut de Science Financière et d'Assurances

Directeur : M. F. DE MARCHI

Directeur : M. le Professeur F. THEVENARD

Directeur : Mme C. FELIX

Directeur : M. Hassan HAMMOURI

Directeur : M. le Professeur S. AKKOUCHE

Directeur : M. le Professeur G. TOMANOV

Directeur : M. le Professeur H. BEN HADID

Directeur : M. le Professeur J-C PLENET

Directeur : M. Y. VANPOULLE

Directeur : M. B. GUIDERDONI

Directeur : M. le Professeur E.PERRIN

Directeur : M. G. PIGNAULT

Directeur : M. le Professeur C. VITON

Directeur : M. le Professeur A. MOUGNIOTTE

Directeur : M. N. LEBOISNE

Titre

Systemes multi-agents, auto-organisation et contrôle par apprentissage constructiviste pour la modélisation et la régulation dans les systèmes coopératifs de trafic.

Résumé

Dans un proche futur, les véhicules connectés et autonomes remplaceront nos véhicules actuels, et il sera nécessaire de repenser intégralement la mobilité. Le conducteur, avec ses lacunes, sera de plus en plus assisté, et un jour détrôné par un système embarqué, capable d'agir plus rapidement, tout en ayant une représentation plus précise et fiable de son environnement de navigation. Pourtant, il reste encore du chemin à parcourir avant d'arriver à un tel stade de maturité : l'environnement du véhicule est complexe, imprévisible et conflictuel, car partagé avec d'autres acteurs de la mobilité.

Ce travail de thèse vise à anticiper l'arrivée de ces nouveaux véhicules afin de proposer des comportements coopératifs au niveau des véhicules et de l'infrastructure tout en permettant un contrôle décentralisé de ce type de système complexe. Dans le cadre d'une approche multi-agents et d'une architecture distribuée, nous proposons d'abord une modélisation par couplage des dynamiques physique et communicationnelle, auxquelles s'ajoute une intégration de la fiabilité de l'information (confiance).

L'étape suivante a été de développer un cadre de simulation propice à l'implémentation de nos modèles dans le cas des véhicules connectés. Nous introduisons un nouveau simulateur de trafic, construit comme une extension d'une plateforme existante, au sein duquel les flux d'informations entre les véhicules et avec l'infrastructure prennent la forme d'échanges de messages. Toutes les informations du système proviennent de capteurs, et toutes les entités, modélisées comme des agents, sont autonomes dans leur prise de décision. De nouvelles formes de contrôle sont désormais envisageables en utilisant des consignes transmises par l'infrastructure communicante. Le couplage des dynamiques assure la cohérence et l'inter-dépendance des différents modèles dans le simulateur. Nous montrons en simulation que, grâce à l'intégration d'informations supplémentaires via la communication, les véhicules, modélisés par un modèle microscopique multi-anticipatif bilatéral, sont capables de réduire l'effet de perturbations propagées au sein d'un flux.

En termes de stratégies de contrôle, une des problématiques principales est de garantir une forme de contrôle qui s'adaptera aux différentes phases de déploiement des systèmes coopératifs. L'analogie avec des problèmes de l'IA (problème de cognition) nous a mené à traiter le problème de manière plus abstraite : comment permettre à un système autonome de contrôler son environnement. Les approches constructivistes, que nous avons retenues, modélisent le processus de cognition comme un phénomène de construction itératif. Pour le trafic coopératif, l'avantage est de disposer d'un système capable de générer ses propres stratégies, en utilisant ou non des connaissances expertes, et de les faire évoluer au cours du temps pour s'adapter aux véhicules composant le flux. Les résultats de notre approche sont présentés dans deux cadres de simulation. Le premier est un prototype visant à illustrer les comportements de bas niveau dans un environnement simplifié. Nous montrons que le modèle est capable dans ce cadre de combiner différentes représentations individuelles pour construire une représentation et de s'adapter à différents contextes en les recombinaison dynamiquement. Puis, dans le cadre de simulation du trafic coopératif, les résultats laissent entrevoir le potentiel de notre approche dans des applications réelles.

Mots-clés

Systèmes Coopératifs, Systèmes Multi-Agents, Couplage, Simulation, Apprentissage, Constructivisme

Title

Multi-agent systems, self-organization and constructivist learning for Cooperative Intelligent Transportation Systems modeling and control.

Abstract

In a near future, connected and automated vehicles will progressively replace current vehicles, leading to deep changes in transportation. The driver will be soon assisted and then replaced by an embedded system, able to act quicker, relying on a more robust and precise representation of its surrounding environment. However, some steps are still needed before coming up with such a level of automation since the vehicle environment is complex and unpredictable.

This work intends to anticipate the introduction of these new kinds of vehicles by providing cooperative behaviors at both infrastructure and vehicle levels, at the same time allowing a decentralized control of these systems. We propose a distributed modeling framework, using multi-agent systems, relying on the coupling of the system dynamics: information, communication and reliability (modeled through the concept of trust).

The next step was to develop a simulation framework enabling the implementation of our models for connected vehicles applications. We present a new microscopic traffic simulator, built as an extension of an existing platform, and able to model information exchanges using messages between vehicles and with the infrastructure. All data are provided by sensors and all entities, modeled as agents, are autonomous regarding their decision process. Thanks to the simulator, it is possible to imagine new control strategies relying on recommendations disseminated by the connected infrastructure. Consistency and interdependence of the simulator components are ensured by the dynamic coupling. As for the vehicles' dynamics, we propose a bilateral multi-anticipative model that integrates additional information from communications in the vehicle decision process. Results in simulation confirm that the model is able to reduce the propagation of perturbation through the flow, leading to a more homogeneous and stable traffic.

One of the major issues regarding traffic control strategies will be to dynamically adapt the action policy to the several deployment stages of cooperative transportation systems. The similarities with Artificial Intelligence problems like cognition motivate a more abstract study: how to model an

autonomous system able to control its environment. We choose the constructivist approaches, that propose to model the cognition process as an iterative building process. For cooperative traffic, the benefits lie in the ability of the system to generate its own strategies, relying or not on domain specific knowledge, and then make them evolve to be adapted to vehicles in the flow. The results from our approach are presented in two distinct simulation frameworks. The first one is an experimentation prototype aiming at highlighting the low-level behaviors in a simplified environment. In this context, we show that the model is able to combine efficiently several individual concurrent representations in order to build a high-level representation that can be adapted to several contexts. The second framework is the traffic simulator where the results lead to some insights about the potential of our approach for such realistic applications.

Keywords

Cooperative Intelligent Transportation Systems, Multi-Agent Systems, Coupling, Simulation, Learning, Constructivism

Remerciements

Penser au contenu de cette page a toujours fait partie des motivations principales qui m'ont mené à me lancer et à persévérer dans l'aventure de ma thèse de doctorat. Cette aventure a pris forme lorsque j'ai quitté ma campagne franc-comtoise natale poussé par un soutien sans failles dont j'ai la chance de profiter encore aujourd'hui.

Mes premiers remerciements vont aux lecteurs assidus du présent document dont font partie les membres du jury. Merci à M^{mes} Amal El Fallah Seghrouchni et Giovanna Di Marzo Serugendo et MM. Hani Mahmassani et Patrick Reignier d'avoir accepté d'étudier et d'évaluer le fruit de mon travail. Je sais combien votre temps est précieux, et je vous exprime toute ma gratitude pour vos critiques et commentaires à la fois précis et profitables.

Durant ces quelques années, j'ai pu profiter de la garde partagée de mes (nombreux) encadrants bienveillants. Je me souviens encore de notre rencontre, le périple pour arriver à Lyon, et l'envie de partager votre environnement. Merci à vous d'avoir cru en moi à ce moment là et de ne pas avoir (trop) changé d'avis depuis. Selon moi, vous méritez largement un paragraphe personnalisé :

- Salima, j'ai toujours été fier de me présenter comme un de tes thésards. Merci pour ta disponibilité, le partage de tes idées, et la grande liberté accordée dans mon travail. Par la confiance que tu m'as accordé, j'ai vraiment eu l'impression que l'on avançait ensemble, et que cette thèse est l'aboutissement d'un travail commun et complémentaire. Il reste encore du chemin à parcourir, et j'espère que l'on poursuivra nos efforts rapidement.
- Frédéric, tu as partagé tes idées et contribué à faire évoluer le sujet vers des thématiques innovantes et stimulantes. Merci d'avoir pris le temps de réfléchir avec moi lors des phases de simulation qui ne donnaient pas grand chose, et de m'avoir incité à chercher sans cesse à prendre du recul sur mes travaux.
- Nour Eddin, merci d'avoir été aussi présent, et surtout aux bons moments, pour me permettre de profiter de ta prise de recul et de tes critiques extrêmement pertinentes. J'ai aussi beaucoup apprécié la personne cachée derrière le directeur de labo. Un grand merci pour l'opportunité de continuer un bout de chemin dans ton équipe, ce qui ne m'a pas demandé de profonde réflexion tellement le cadre de travail que tu as construit au LICIT me semble unique.
- Romain, tu m'as accompagné dans mon quotidien, et j'ai eu la chance de compter parmi tes tout premiers doctorants. Avant tout, je souhaiterais souligner à quel point tu as mené cette aventure avec succès. Je n'arrive toujours pas à cerner ton sens de l'humour mais cela ne m'empêche pas d'admirer ton caractère admirablement équilibré. Le moment est aussi venu de m'excuser pour toutes les heures passées à concevoir des figures en Tikz au lieu de réfléchir au fond des choses. Merci d'avoir, malgré cela, assuré la continuité après ton exil en Bretagne.

L'une des joies du co-encadrement est de rencontrer deux fois plus de collègues (et même plus encore). Sans distinguer les labos, merci en particulier à :

- Anne-Christine, qui a contribué à maintenir la glycémie du LICIT au plus haut tous les lundis. Merci d'avoir ruiné les pauses café en y apportant des mots-croisés, heureusement que tu participes encore en tant que public privilégié à nos bêtises en tout genre. Un immense merci pour la relecture attentive du présent manuscrit, tu es une véritable grammar nazi, ce qui est un compliment inestimable de ma part.

- Jean-Luc, qui a bien failli me motiver à monter un douzaine de "smart" startups au lieu d'avancer sur ma thèse. Merci pour les discussions à des années lumières du quotidien, et félicitations pour t'être débarrassé de ce satané carré blanc et de Martin.
- Bernard, qui incarne le mélange subtil entre le flemmard et le scientifique averti : l'informaticien. Au plaisir de travailler ensemble, ce n'est plus un SCOOP, pour le meilleur et pour le pire.
- Arthur, qui sera aussi une des prochaines victimes en charge de me supporter.
- Sonia, qui a toujours fait preuve d'une compréhension et d'une patience infinie.
- Anahid, Fathia, Matthieu, Sergey, qui ne m'ont supporté que pendant quelque temps.
- Aurélien, Nicolas, Ludovic, Mathieu, Laetitia, Christine, Cécile, Colette, Delphine, Véronique, Angelo, qui ont eu beaucoup moins de chance.
- Clélia, Fairouze, Kamel, Julien, Johana, Matthis, Maxime, Nicole, Vincent, Guilhem, Sergio, qui seront tous (très) bientôt des docteurs accomplis. Courage à vous tous.
- Étienne, Fouad, Baudouin, qui sont passés par la case doctorat et qui ont survécu. Merci de m'avoir donné un peu d'espoir de ce côté.
- André, à qui je souhaite toute la réussite qu'il mérite. Merci d'avoir échangé ton expérience et ta bonne humeur communicative aux moments où j'en avais vraiment besoin.
- Aurélien *a.k.a.* Junior, qui est un puits d'idées sans fin lorsqu'il s'agit de s'aérer l'esprit. Je suis certain du succès de notre très prochain studio de jeux vidéos indépendants (quand on aura enfin une idée vraiment originale).
- Merci à Raphaël, qui a contribué à rendre l'expérience de la thèse complètement différente. Nous avons souffert ensemble, et je ne t'ai pas ménagé. En contrepartie, avec Adèle, vous m'avez souvent ridiculisé au cours de parties de jeux de société s'achevant irrémédiablement par un échec collectif. Merci pour tout et le meilleur dans votre nouvelle vie.
- Pierre-Antoine, avec qui j'ai partagé le même bureau, et qui a supporté ma proximité et mon aversion quotidienne pour les capteurs Bluetooth. Merci pour les moments de défoulement sur le court de l'IFSTTAR. Merci d'avoir contribué à mon échec financier lors de nos multiples escapades à Décathlon. Anaïs, désolé de l'avoir converti lui-aussi à des visites trop fréquentes chez Décathlon, et merci d'avoir bien voulu me le prêter de temps en temps. Merci, et à très bientôt les copains.

J'aimerais profiter des prochaines lignes pour retourner un court moment en Franche-Comté, comme j'aurais voulu le faire beaucoup plus souvent depuis trois ans. Merci Papa, merci Maman, depuis tout petit j'ai acquis votre confiance, vous m'avez laissé choisir ma voie et avez accompagné toutes mes décisions. Merci pour votre soutien. Merci les frères et sœurs, Pierre, Noémie et Quentin. Merci à Jean et Mina pour les moments partagés en toute simplicité. Enfin, gros bisous à Mamie.

Il me reste encore quelques lignes sur cette page, et celles-ci sont pour toi. Désormais, tu aspiras à l'avenir à te rapprocher du soleil. Tu as su éclairer mes pas, dissiper mes doutes, réchauffer mon cœur et par dessus tout, supporter mes râleries (ne t'en fais pas, j'en ai partagé une bonne partie avec Raphaël et Pierre-Antoine). Tu as pu partager une partie de mon quotidien, et pas la plus désagréable lors de notre voyage (100% professionnel, bien entendu) à San Fransisco. Depuis, j'ai encore du travail pour te convaincre que je ne fais pas "*que de glander toute la journée*". Sara, tu l'as sans doute deviné, c'est vers toi que je me tourne pour conclure ces quelques pages, de la même façon que je les ai commencées : tu peux aussi compter sur moi pour être ton "soutien sans failles". Je nous souhaite de poursuivre notre propre aventure, ici comme ailleurs, mais avant tout ensemble.

Table des matières

Introduction	1
1 Contexte et positionnement	3
1.1 Systèmes de Transports Intelligents Coopératifs – C-ITS	4
1.1.1 Systèmes de Transports Intelligents – ITS	4
1.1.2 Systèmes Coopératifs	6
1.1.3 État des lieux des C-ITS	10
1.1.4 Enjeux des C-ITS	13
1.2 Défis et hypothèses	14
1.2.1 Verrous scientifiques	14
1.2.2 Objectifs, contributions et hypothèses	15
1.3 Approche méthodologique	16
1.3.1 Aperçu	16
1.3.2 Organisation du manuscrit	16
2 Couplage de dynamiques pour la modélisation du trafic coopératif	19
2.1 Approches distribuées et modélisation du trafic	20
2.1.1 Modélisation du trafic	20
2.1.2 Modèles pour le trafic coopératif	28
2.1.3 Les Systèmes Multi-Agents – SMA	31
2.1.4 Les SMA appliqués aux transports	35
2.2 Réputation, confiance et coopération	37
2.2.1 Définitions	37
2.2.2 Coopération dans les SMA	38
2.2.3 Systèmes fondés sur la confiance et la réputation	38
2.3 Modélisation multi-agent des C-ITS par couplage de dynamiques	41
2.4 Application du couplage : un modèle multi-anticipatif bilatéral	43
2.4.1 Objectifs	43
2.4.2 Hypothèses	44
2.4.3 Interactions physiques	45
2.4.4 Communication	46
2.4.5 Fiabilité des capteurs	48
2.5 Bilan et perspectives	49
3 Simulateur multi-agent du trafic coopératif	51
3.1 Approches distribuées et simulation : analyse de l'existant	52

3.2	Modèles et implémentation	54
3.2.1	Extension de MovSim : MASCAT	54
3.2.2	Architecture	55
3.2.3	Perception	57
3.2.4	Décision	60
3.2.5	Action	63
3.2.6	Autres améliorations du simulateur et perspectives	65
3.3	Indicateurs de performance et d'évaluation	69
3.3.1	Écoulement du flux	69
3.3.2	Sécurité	70
3.3.3	Consommation et émissions	71
3.4	Applications : stratégie locale d'homogénéisation	71
3.4.1	Scénario	72
3.4.2	Effets de l'homogénéisation	73
3.4.3	Impact de la fiabilité des capteurs	76
3.5	Vers l'élaboration de stratégies de contrôle I2V des véhicules connectés	80
3.5.1	Insertion	81
3.5.2	Fermeture de voie	83
3.5.3	Discussion	83
3.6	Bilan et perspectives	84
4	Vers un contrôle décentralisé et adaptatif du trafic coopératif	87
4.1	Contrôle et régulation du trafic dans l'environnement du véhicule connecté	88
4.1.1	Approches classiques de contrôle et régulation du trafic	88
4.1.2	Contrôle et régulation des C-ITS	89
4.1.3	Vers une approche constructiviste pour le contrôle des C-ITS	90
4.2	Motivations	91
4.2.1	Caractéristiques d'un système constructiviste	91
4.2.2	Une architecture de référence	93
4.2.3	Problématique et verrous scientifiques	95
4.3	Modèle conceptuel	103
4.3.1	Vue d'ensemble	104
4.3.2	Initialisation	105
4.3.3	Construction d'une perception de haut niveau	107
4.3.4	Contrôle de l'environnement et renforcement de la représentation	109
4.3.5	Processus d'amélioration de la discrétisation par associations	112
4.3.6	Perspectives	114
4.4	Étude du modèle	116
4.4.1	Un prototype expérimental	116
4.4.2	L'apport des représentations concurrentes	120
4.4.3	Capacité d'adaptation	123
4.4.4	Affinage de la représentation	127
4.5	Application : stratégie I2V de contrôle des C-ITS	129
4.5.1	Cadre expérimental	129
4.5.2	Implémentation	131
4.5.3	Résultats	134
4.5.4	Discussion	136

4.6 Bilan et perspectives	136
Conclusion et perspectives	139
Synthèse des contributions	139
Perspectives	141
Bibliographie	143
Articles de l’auteur et valorisation	157
Table des figures	159
Liste des tableaux	163
Liste des sigles	165

Introduction

La mobilité des biens et des personnes constitue un enjeu majeur des sociétés modernes dont le développement va de pair avec l'essor des systèmes de transports. Par le passé, le développement de l'infrastructure, en particulier routière, permettait de répondre à la demande croissante. Plus récemment, les contraintes budgétaires amènent plutôt à s'orienter vers un meilleur usage de l'infrastructure. Depuis quelques dizaines d'années, des efforts ont été entrepris pour proposer de nouveaux systèmes intelligents, qui permettront, par l'intermédiaire de stratégies de contrôle ou de régulation, d'améliorer l'usage des ressources dédiées aux transports. Ces systèmes de transport intelligents tentent d'apporter des solutions aux problèmes de congestion, qui accompagnent la saturation des réseaux. Le trafic routier est également un contributeur aux émissions de polluants dont l'ampleur ne cesse de croître (la part attribuée au trafic routier atteint plus de 50% dans Paris depuis 2012 [5]). En particulier, les zones péri-urbaines concentrent de fortes demandes et des contraintes fortes en termes d'aménagement des territoires. Dans ce contexte, les stratégies déployées n'ont que peu d'impact sur la gestion du trafic. Plus récemment, nous assistons à l'apparition d'équipements basés sur les nouvelles technologies de l'information et de la communication dans les véhicules. Ce développement est accompagné par la récente tendance vers l'automatisation, qui laisse dessiner les contours du futur de la mobilité. La communication inter-véhiculaire autorise des échanges d'informations au sein des flux allant dans le sens d'un meilleur usage de l'infrastructure. Ces nouveaux véhicules, évoluant sur une infrastructure connectée, font partie des systèmes coopératifs. Aujourd'hui, la coopération entre les acteurs de la mobilité (usagers, gestionnaires, infrastructure et véhicules) n'est encore qu'un objectif. Les premiers systèmes coopératifs atteignent la phase de déploiement, qui sera progressive. Dans le futur, le véhicule autonome remplacera à terme les véhicules que nous connaissons. Notre travail vise à étudier la phase de transition entre un flux de véhicules actuels, jusqu'à un flux de véhicules entièrement autonomes, en s'intéressant plus particulièrement à la phase de transition entre des flux mixtes partiellement composés de véhicules connectés et/ou autonomes. Notre travail de thèse propose des stratégies de contrôle et de régulation adaptées aux différentes phases de déploiement des systèmes coopératifs.

Cependant, notre démarche se heurte à un manque de disponibilité des données. En effet, très peu de véhicules connectés et encore moins autonomes circulent sur les routes (la plupart étant utilisés à des fins expérimentales). Le cadre législatif, notamment lié à l'automatisation des véhicules ou à l'anonymisation des données échangées, reste encore à définir, même si des efforts de standardisation assurent l'interopérabilité des systèmes. Pourtant, notre objectif est d'anticiper l'arrivée de ce type de véhicule. La solution passe donc par une étude en simulation. Les interactions complexes de ces systèmes posent dans un premier temps la question de l'échelle et du formalisme de modélisation. Contrairement aux approches classiques de modélisation du trafic, les systèmes coopératifs préfigurent un usage différent de l'infrastructure, possiblement communicante. Les

véhicules seront aussi équipés de capteurs et de dispositifs leur permettant de se représenter leur environnement de navigation. La masse et la complexité de ces informations, couplées aux échanges entre les entités par communication, contribuent à ce système complexe, difficile à modéliser. En effet, la nature distribuée des échanges et de la prise de décision (embarquée au niveau des véhicules ou de l'infrastructure) impose un choix de modélisation particulier : quelle échelle de modélisation choisir ? Comment modéliser les échanges d'information ? Comment modéliser l'infrastructure ? Comment assurer l'autonomie des systèmes ?

Grâce à un cadre de modélisation suffisamment réaliste, il est possible de concevoir de nouvelles stratégies mettant à profit les technologies qui sous-tendent les systèmes coopératifs. Le problème de congestion est particulièrement intéressant puisque ce phénomène est difficile à prévoir et à combattre. Les systèmes coopératifs laissent entrevoir des possibilités d'anticipation de ce type de phénomène en s'appuyant sur l'échange d'informations. Les véhicules devront alors être dotés d'un mécanisme de prise de décision plus complexe et aussi plus autonome pour traiter ces données rapidement. L'infrastructure intelligente a aussi un rôle important à jouer puisqu'elle pourrait relayer des informations plus globales, et intégrer des objectifs des gestionnaires. Quels pourraient être les bénéfices de ces systèmes coopératifs ? Quelles en seront leurs limitations ? Ces questions étant très générales, notre travail se concentre en priorité sur les sections péri-urbaines, où le problème de congestion est récurrent. Dans ce cadre, est-ce que les systèmes coopératifs permettront un meilleur écoulement des flux ? Quel sera le rôle de l'infrastructure ?

L'introduction sur le réseau de nouveaux types de véhicules (connectés et/ou autonomes) entraînera l'apparition de nouveaux usages et de nouveaux comportements (anticipation, coopération ...) qui pourraient être mal interprétés ou mal acceptés par les usagers. Les approches de contrôle et de régulation actuelles, développées en observant les flux de trafic en l'absence de ces systèmes, ne seront plus adaptées. Des stratégies innovantes doivent donc être proposées pour prendre en compte les nouveaux échanges d'information entre les véhicules et avec l'infrastructure dits coopératifs. Le déploiement progressif de ces systèmes intelligents imposera de concevoir des stratégies de contrôle capables de s'y adapter. Que deviendront les stratégies actuelles ? Comment rendre ces stratégies plus adaptatives ? Comment les gestionnaires pourront-ils intégrer leurs objectifs tout en garantissant les attentes de mobilité des usagers ?

Notre travail débute avec l'étude du domaine des systèmes coopératifs de transport intelligents. Les questions portant sur le développement et le déploiement futur de ces systèmes font apparaître des problèmes sous-jacents plus généraux. Ces verrous scientifiques nous orientent vers le domaine de l'intelligence artificielle et notamment vers l'étude des systèmes distribués. La réflexion autour de ces problématiques pourra donc dépasser le cadre des systèmes coopératifs pour englober la thématique de la cognition. Néanmoins, ils constitueront l'environnement de validation de nos contributions.

Contexte et positionnement

” *My dream is to achieve persistent autonomy of lifelong map learning. Say, a robot car that, as it drove around [...], it would get better and better. It would learn about the world.*

— **John Leonard**

Professeur et chercheur, MIT

Le développement des technologies de l'information et de la communication gagne du terrain au profit des applications dédiées à la mobilité. Les récentes avancées en termes de collecte de données permettent aux gestionnaires d'observer en temps réel les conditions de trafic sur leur réseau. Depuis quelques années, des systèmes permettent de diffuser ces informations aux usagers, dans le but d'améliorer la qualité des flux et donc l'usage de l'infrastructure. Ces systèmes fondés sur la télématique et les technologies de l'information et de la communication (TIC) sont connus sous le nom de Systèmes de Transport Intelligents (ITS). Plus récemment, la tendance se porte vers l'automatisation mais aussi vers la connectivité des véhicules ; entre eux et avec l'infrastructure. Les interactions entre les entités mobiles (véhicules) et fixes (infrastructure) se complexifient et il est désormais possible d'imaginer l'apparition de nouveaux comportements collectifs. Parmi ces comportements, l'automatisation, la connectivité et le partage de ressources (co-voiturage, véhicules libre service) constituent des tendances lourdes pour optimiser l'usage et le fonctionnement de l'infrastructure routière par les différents acteurs de la mobilité. Pour ce qui est de la mobilité connectée, les systèmes dits coopératifs n'en sont aujourd'hui qu'à leur phase de prototype, et leur déploiement progressif sera dépendant du compromis attendu entre les avantages qu'ils pourront proposer face à leurs limitations possibles. En effet, il s'agit de convaincre tous les acteurs : les gestionnaires qui devront investir dans l'infrastructure intelligente, les constructeurs qui souhaiteront équiper leurs véhicules, et les usagers qui s'adapteront à de nouveaux usages. Dans ce chapitre, nous abordons cette problématique sans limiter notre étude à une phase précise du déploiement des systèmes coopératifs.

La section 1.1 présente le contexte général de notre travail : les Systèmes de Transport Intelligents Coopératifs (C-ITS). Pour cela, nous rappelons l'historique de leur développement en les replaçant dans le cadre des ITS actuels (section 1.1.1), avant de décrire plus précisément les spécificités et particularités apportées par les C-ITS (section 1.1.2). Le stade actuel du déploiement est décrit au travers d'une revue des projets pilotes en la matière (section 1.1.3). Le déploiement de ces systèmes est confronté à des enjeux multiples (flux, usagers, et environnement) qui sont présentés en section 1.1.4. La problématique de notre travail est illustrée à travers une discussion des verrous scientifiques (section 1.2) qui nous ont mené à élaborer le cadre méthodologique de ce travail et à orienter l'organisation du présent manuscrit (section 1.3).

1.1 Systèmes de Transports Intelligents Coopératifs – C-ITS

La technologie et plus particulièrement la télématique occupe une place de plus en plus prépondérante dans les véhicules et leur environnement de navigation. Cette intelligence digitale est directement intégrée ou déportée dans le véhicule pour donner naissance à des systèmes innovants qui transforment depuis peu la mobilité en une expérience ultra-connectée. Aussi, le domaine des transports profite déjà depuis quelques décennies de l'apport des technologies de l'information et de la communication. Ce développement a donné naissance aux ITS. Le développement industriel passe aussi par l'automatisation des outils. L'automobile ne déroge pas à la règle puisque les systèmes autonomes embarqués se développent très rapidement.

Cette tendance du développement technologique du secteur automobile (et des transports en général) fait face à un des enjeux actuels les plus importants. Les flux de personnes et de biens se développent de façon exponentielle alors que l'espace sur les territoires est de plus en plus limité. Comment envisager de faire cohabiter un nombre d'acteurs de la mobilité toujours croissant dans ce contexte ? Le trafic routier est l'un des modes de transports privilégié, en contexte urbain comme à la campagne. La question qui se pose alors est : peut-on impliquer et faire coopérer les acteurs des ITS ? L'évolution logique de ces systèmes complexes passe par le développement de nouveaux types de véhicules. Seront-ils coopératifs ? Connectés ? Automatiques ou autonomes ? Quoiqu'il en soit, le succès de ces Systèmes Coopératifs passera par une phase d'étude et de déploiement qui devra montrer leur potentiel.

1.1.1 Systèmes de Transports Intelligents – ITS

Le début du développement des Systèmes de Transports Intelligents (ITS) remonte à une quarantaine d'années. Les premières briques ont porté sur l'informations aux usagers. Plus récemment, avec l'implication de plus en plus forte de l'industrie automobile et la démocratisation des technologies de la communication et de l'internet, des systèmes avancés d'aide à la conduite (ADAS) et des capteurs embarqués ont vu le jour. Le point central reste la donnée, qu'il faut recueillir et partager dans une forme adaptée aux usagers.

1.1.1.1 Information dynamique aux usagers

En France, la première initiative notable de système d'information dédié aux transports date des années 70 avec Bison Futé. Les données sont alors recueillies par des opérateurs, sur le bord de la chaussée, en charge du comptage manuel des véhicules. L'information partagée, très générale, était donnée aux usagers par voie indirecte (radio). L'équipement des routes se développe rapidement et les premiers feux de tricolores adaptatifs apparaissent. Ils utilisent un capteur incrusté dans la chaussée (une boucle électromagnétique qui détecte les véhicules par inductance) pour activer les plans de feux ou modifier les cycles de feux.

Les premiers Panneaux à Message Variable (PMV) apparaissent dans les années 80. Cet équipement, couplé à des méthodes de diffusion, permet une dissémination d'information plus précise. Les

automobilistes peuvent accéder à des informations relatives à la sécurité (alertes) ou obtiennent une estimation du temps de parcours sur leur itinéraire.

Les premiers systèmes de régulation et de gestion du trafic apparaissent dans les années qui suivent. Ils se focalisent d'abord sur les transports en commun (information et gestion) avant de s'intéresser aux véhicules particuliers (VP). Des systèmes comme la régulation d'accès [159] se mettent en place pour réduire les congestions.

Les années 2000 sont marquées par le développement des technologies de l'information et de la communication. Pourtant, ces technologies mettent du temps avant de faire partie du quotidien des automobilistes. Cette période est consacrée aux problèmes de sécurité routière et de transition énergétique. Les premiers systèmes d'auto-partage et de covoiturage en sont les exemples.

1.1.1.2 Systèmes embarqués

La suite de la modernisation des véhicules et de leur infrastructure passe par l'introduction de systèmes embarqués, allant à petit pas dans la direction de l'automatisation. Cette catégorie regroupe les systèmes d'aide à la conduite ou systèmes avancés d'assistance au conducteur (ADAS : *Advanced Driver Assistance Systems*). Ce sont des dispositifs automatiques intégrés aux véhicules destinés à assister le conducteur dans ses tâches de conduite et le cas échéant corriger d'éventuelles erreurs, apporter un bénéfice en termes d'anticipation ou de perception, ou soulager le conducteur d'une partie de ses tâches de conduite. Il existe des systèmes actifs qui agissent directement sur la dynamique du véhicule (freins, direction, accélération) comme le maintien de trajectoire (LKS : *Lane Keeping System*), et des systèmes passifs qui informent le conducteur, comme la détection d'angle mort, ou les autres usagers comme le système d'avertissement de franchissement de ligne (LDWS : *Lane Departure Warning System*).

La tendance vers la sécurisation des flux a mené la recherche à proposer des solutions permettant d'éviter les collisions ou de les prévenir (*collision avoidance / warning*). Ces systèmes qui visent avant tout une amélioration de la sécurité reposent sur l'automatisation (partielle) de l'accélération du véhicule [198]. Le système le plus étudié est le régulateur de vitesse à contrôle de distance (ACC : *Adaptive Cruise Control*). Ce système se compose d'un capteur de proximité permettant au véhicule de percevoir la distance et la vitesse relative avec son véhicule prédécesseur (*leader*). L'objectif est de proposer un système dont le temps de réaction est plus faible que celui du conducteur. Hoogendoorn *et al.* [93] ont montré l'impact positif en termes de vitesse et de débit, quel que soit le taux de pénétration, apporté par un AICC (*Autonomous Intelligent Cruise Control*) fondé sur un modèle contrôlant entièrement les commandes longitudinales du véhicule.

Tous ces travaux prennent en compte un taux d'équipement (*i.e.* un pourcentage de véhicules équipés) partiel. Les résultats expérimentaux obtenus montrent un impact sur la globalité du flux, ce qui corrobore la possibilité de fluidifier un flux par une action localisée et distribuée. L'étape suivante consiste à ne plus considérer l'action uniquement sur un seul véhicule, mais à introduire une forme de communication qui ouvre la voie de la coopération.

1.1.2 Systèmes Coopératifs

Les systèmes coopératifs (C-ITS) constituent l'évolution logique des systèmes de transports intelligents dont ils forment la seconde génération. Ils sont accompagnés par des changements induits par le déploiement de nouveaux types de véhicules : autonomes, connectés, équipés et qui deviendront coopératifs. L'infrastructure joue aussi un rôle primordial dans la régulation de ces systèmes complexes, qui reposent sur les nouvelles technologies de la communication.

1.1.2.1 Véhicules autonomes, connectés, équipés et coopératifs

Le développement rapide des systèmes embarqués a entraîné l'apparition de nouveaux types de véhicules qui ont depuis longtemps dépassé le stade de prototype. Tous ces véhicules font partie de la catégorie des véhicules intelligents dans la mesure où ils manipulent des flux de données de manière avancée. Ces données permettent au véhicule de percevoir son environnement immédiat. La figure 1.1 représente les différentes interactions qu'un véhicule intelligent peut avoir avec son environnement local.

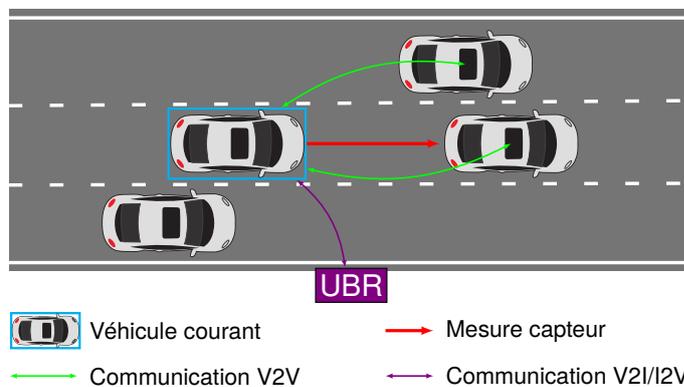


Fig. 1.1: Les interactions d'un véhicule intelligent et son environnement.

Le véhicule considéré est dit **équipé** s'il embarque un ou plusieurs capteurs. Ces capteurs permettent au véhicule équipé de mesurer certains paramètres dynamiquement et peuvent être utilisés pour différentes applications. Ces applications peuvent être limitées à des systèmes automatiques (activation automatique des feux, système anti-collision, système anti-franchissement de ligne blanche) ou peuvent contribuer à rendre le véhicule partiellement autonome (assistance à la manœuvre de parking, régulateur de vitesse intelligent, ou plus tard conduite autonome).

Les véhicules communicants, ou **connectés**, sont capables d'échanger des informations grâce à une forme de communication sans fil. La communication peut se faire entre les véhicules (on parle alors de communication Véhicule à Véhicule – *Vehicle to Vehicle* – V2V) et avec l'infrastructure (Infrastructure à Véhicule – I2V – et Véhicule à Infrastructure – V2I). L'infrastructure communicante est nommée de manière générique : Unité de Bord de Route (UBR).

Les véhicules **autonomes** se démocratisent depuis l'avènement de la Google Car. Les premiers véhicules à être équipés de modules de conduite autonome de série sont déjà en vente (notamment chez Tesla). Ces véhicules sont capables, grâce à leurs capteurs embarqués et une représentation

de leur environnement de navigation, de se mouvoir sur le réseau sans aucune action du conducteur. Pour le moment, il s'agit de systèmes déconnectables ou ne fonctionnant que sur une partie bien connue du réseau. Néanmoins, leur développement suscite un intérêt grandissant, et leur déploiement intégral n'est sans doute plus qu'une question de temps. Le défi est tout de même bien plus difficile qu'il n'y paraît (en comparaison des avions qui sont autonomes depuis longtemps) car leur environnement est très complexe et dynamique.

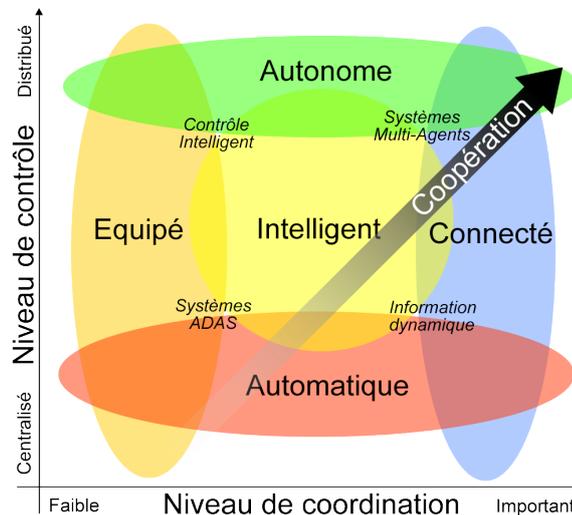


Fig. 1.2: Représentation simplifiée des véhicules différents types de véhicules dans le contexte des Systèmes Coopératifs [20].

Cette complexité est d'ailleurs une des motivations de ces véhicules "coopératifs" qui ne limitent plus leur actions à la réalisation de leurs propres objectifs. Les informations échangées profitent au flux, et les objectifs se situent désormais à l'échelle du réseau. La coopération entre les véhicules peut être vue comme la solution permettant de garantir une décentralisation allant vers l'autonomie du processus de décision du véhicule, dépendant des informations de la communication pour compléter sa perception partielle de l'environnement et issue de ses capteurs (figure 1.2). Un système autonome se distingue d'un système automatique par sa capacité d'adaptation et sa faculté à évoluer dans un environnement dynamique.

1.1.2.2 Rôle de l'infrastructure

Les C-ITS introduisent une nouvelle forme d'intelligence dans les infrastructures existantes. Sous l'appellation générique "Unités de Bord de Route" (UBR, en anglais *Road Side Unit*), ces entités peuvent avoir des rôles très variés : du simple recueil d'informations à la régulation du trafic.

Les C-ITS autorisent l'envoi de messages entre les véhicules (V2V), qui permettront de mieux anticiper les problèmes survenant sur le réseau. Ces échanges restent cependant locaux, et la portée des messages reste limitée par le rayon de communication entre véhicules. Pour étendre le champ de perception des usagers, il est envisagé d'utiliser les UBR comme des relais d'information. Les informations échangées (I2V) sont alors des notifications concernant des événements dangereux [66] ou des conditions de trafic dense [157]. L'idée est simplement de propager des informations plus précises, directement dans l'habitacle, même si le principe reste proche des PMV.

Face au coût important d'installation de ces bornes, le déploiement des UBR repose sur une optimisation de leur placement [15]. Une solution pour améliorer la couverture de l'infrastructure est de proposer une sorte d'Unité de Bord de Route mobile. Cette idée est analogue avec les techniques d'agrégation dans les réseaux *ad hoc* mobiles [101] dont le but est d'améliorer la diffusion de l'information en définissant des schémas de propagation (via certains véhicules par exemple). Il est aussi possible d'embarquer le système d'une UBR dans un véhicule (c'est par exemple le cas dans le projet SCOOP – section 1.1.3 – où des véhicules de patrouilleurs autoroutiers se comportent comme des UBR mobiles). En pratique, il y a fort à parier que les UBR soient installées en premier lieu sur des sites pilotes choisis pour leur intérêt expérimental (réseau souvent congestionné) ou sur des sites facilitant leur installation (présence de capteurs, électricité, accès au réseau d'information).

En plus du rôle de relais, les UBR se distingueront probablement des véhicules coopératifs par leur connectivité avec les systèmes des gestionnaires d'infrastructure. Cela ouvre la possibilité d'utiliser des informations provenant des capteurs côté infrastructure (boucles électromagnétiques, stations météorologiques, etc.) ou même d'autres UBR. La connectivité avec un centre de contrôle du trafic permet aussi d'envisager la problématique de la régulation dynamique du trafic grâce aux UBR. L'infrastructure devient alors une véritable entité de contrôle du trafic coopératif (chapitre 4).

1.1.2.3 Architecture du réseau de communication

Perçus comme une évolution des réseaux de capteurs mobiles (*Vehicular Ad-hoc NETWORK* – VANET), les C-ITS reposent avant tout sur la capacité de communication entre les entités de ces systèmes. Les véhicules proviennent de différents constructeurs, et les gestionnaires d'infrastructure diffèrent d'un pays (voire d'une région) à l'autre. Pour parvenir à faire communiquer tous ces acteurs, il est nécessaire que ces différents systèmes inter-opèrent entre eux à défaut d'une standardisation du protocole de communication, à la fois entre les véhicules (V2V) mais aussi avec les unités d'infrastructure (V2I/I2V). Les dernières années ont donc été consacrées à l'émergence de l'interopérabilité et de nouveaux standards dans le monde des communications inter-véhiculaires.

L'évolution des technologies de communication mobiles (GSM) avec l'avènement récent de la 4G a pendant un temps provoqué un regain d'intérêt pour leur utilisation dans les ITS. Le besoin d'un réseau dédié et intégrant davantage la problématique liée à la sécurité a cependant pris l'ascendant lors de l'effort de standardisation notamment opérée par le consortium *Car2Car Communication*. Une nouvelle norme baptisée WiFi 802.11p et aussi connue sous le nom de WAVE (*Wireless Access in Vehicular Environments*) a été préférée. Ce standard assure des communications à courte portée (moins de 300 mètres) entre des entités fixes ou mobiles sur des canaux dédiés. Le schéma d'échange retenu est la diffusion (*broadcast*). Un système comme les C-ITS se décompose alors en UEV (Unité Embarquée Véhicule) embarquées dans les véhicules et en UBR (Unités de Bord de Route). Les véhicules intègrent aussi une UA (Unité Application) qui permet de convertir les données échangées dans une interface pour l'utilisateur ou de les intégrer dans le contrôle du véhicule. Des tests grandeur nature ont montré la robustesse du protocole retenu avec des valeurs de latences observées [153] compatibles avec des applications dédiées à la sécurité (de l'ordre de 100 ms). Les communications sur le réseau mobile sont envisagées pour des applications moins critiques (e.g le recueil de données).

La figure 1.3 illustre les différents types échanges d'information dans les systèmes coopératifs. Les véhicules connectés peuvent échanger de l'information entre eux (V2V) et avec l'infrastructure

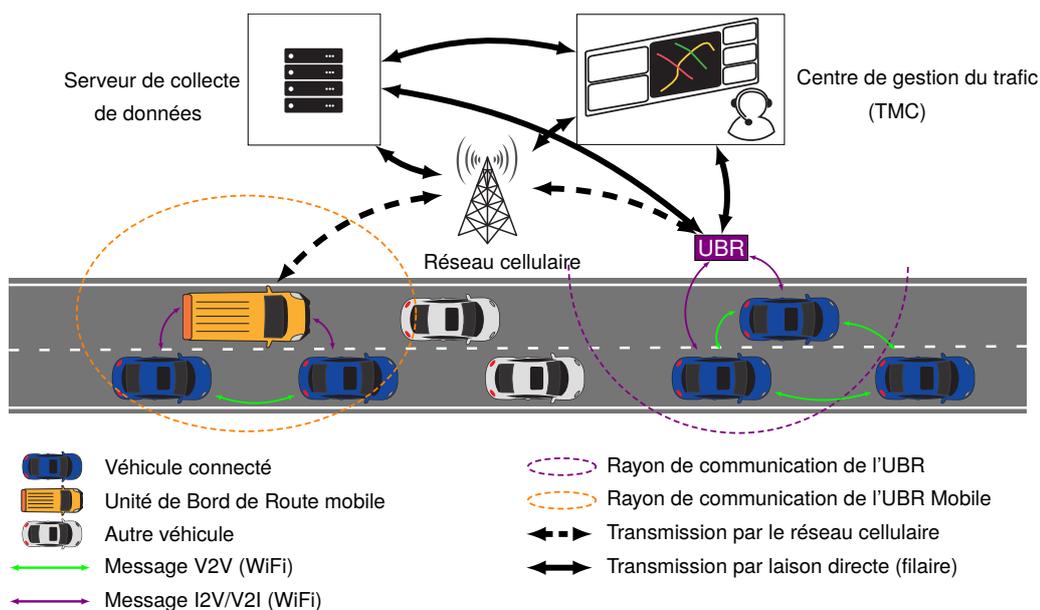


Fig. 1.3: Échanges d'information dans les systèmes coopératifs

(V2I). Ces données peuvent être relayées à un centre de gestion du trafic (*Traffic Management Center* – TMC). Ces échanges reposent sur une liaison filaire lorsqu'elle est prévue ou peuvent s'effectuer sur le réseau cellulaire (à la discrétion du gestionnaire d'infrastructure). Le TMC peut alors disposer d'une remontée d'information pertinente pour la mise en place des stratégies de gestion et de régulation du trafic qui seront propagées par le biais des UBR. Les UBR, fixes ou mobiles peuvent aussi fonctionner en autonomie pour certaines applications (rediffusion ou diffusion d'information locale). Dans un premier temps, les informations proviennent des conducteurs (ou des passagers), mais de plus en plus de véhicules sont équipés de capteurs qui rendent possible l'automatisation du recueil de certaines données. Ces systèmes embarqués profitent des nouvelles technologies en matière de capteurs. Toutes les données récoltées ouvrent la voie à la proposition de nouveaux cas d'usage des systèmes coopératifs.

1.1.2.4 Premiers cas d'usage

Les C-ITS préfigurent une nouvelle façon de gérer les flux de trafic grâce à l'inter-connexion de tous les acteurs, de l'utilisateur jusqu'au gestionnaire. Pourtant, les usages des informations échangées sont encore pour beaucoup à proposer. Afin de laisser entrevoir le potentiel de ces systèmes, nous proposons d'étudier quelques exemples de cas d'usages envisagés. Les années futures verront sans doute l'apparition de nouveaux usages facilitant toujours plus la mobilité.

La figure 1.4 regroupe certains des cas d'usage d'ores et déjà identifiés. Des catégories se distinguent immédiatement :

- **Les stratégies de régulation multi-modales.** Il s'agit d'une interopérabilité entre les acteurs de la mobilité. Un usager peut être orienté vers un service de transports en commun par le TMC qui dispose, grâce au recueil de données, d'une perception élargie de l'offre de déplacement et des conditions associées. Ces informations agrégées peuvent être partagées telles quelles par le gestionnaire où être intégrées à une stratégie de régulation.

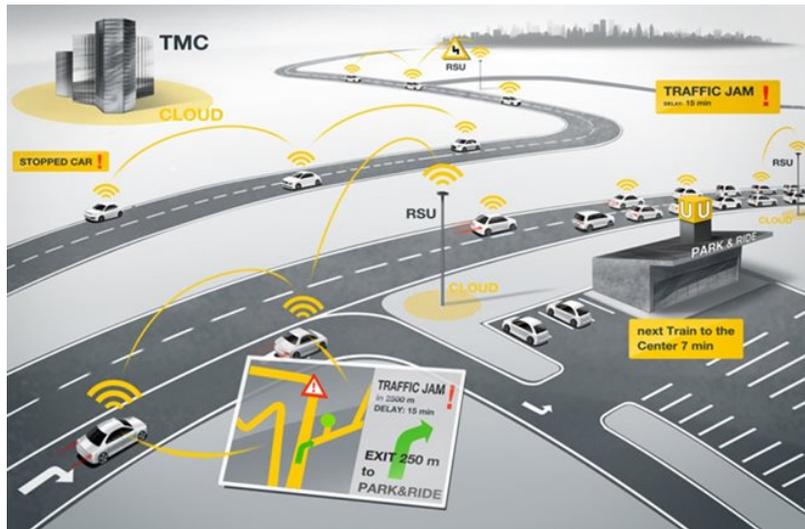


Fig. 1.4: Illustration de certains cas d'usages envisagés dans les C-ITS.
(Source : <http://www.kapsch.net>)

- **Les stratégies de routage.** Sur le même principe, même s'il est possible de s'affranchir localement de l'infrastructure, l'objectif est d'adapter dynamiquement l'itinéraire des véhicules. La mise en place d'itinéraires de délestage permet de mieux répartir la demande sur le réseau. Les technologies C-ITS permettent une information plus précise et adaptative.
- **Les systèmes d'alerte et de sécurité.** Ces stratégies ne passent pas forcément par l'usage de l'infrastructure. L'idée est de propager des informations critiques (incident, véhicule en panne, piéton sur la chaussée, accident) en amont du flux (V2V ou V2V et V2I/I2V). Le maître mot est l'anticipation des phénomènes perturbant l'écoulement du flux. Ce sont les applications phare des C-ITS, qui seront probablement décisives dans leur déploiement.
- **Les stratégies de régulation locales.** Ces stratégies sont sans doute les plus prometteuses dans la mesure où elles permettent un contrôle ciblé et adaptatif des flux. L'apport des systèmes automatisés permet de s'affranchir des contraintes liées aux conducteurs pour atteindre des objectifs à l'échelle du flux. Ces stratégies locales peuvent être gouvernées par le TMC ou reposer sur un algorithme complètement décentralisé (e.g. le système C-ACC [199] – *Cooperative Adaptive Cruise Control*).

Ces applications sont des cas concrets d'utilisation des systèmes coopératifs, qui s'inscrivent dans les objectifs des premiers projets visant à déployer les C-ITS.

1.1.3 État des lieux des C-ITS

Le développement des Systèmes Coopératifs, bien que récent, a déjà franchi la barre du prototypage dans plusieurs pays. On recense plusieurs projets pilotes ou des démonstrations qui visent à montrer l'impact de l'introduction de la communication et de l'équipement de l'infrastructure. Même si le déploiement à grande échelle n'est prévu que pour dans quelques années, plusieurs initiatives confirment la tendance lourde vers plus de connectivité et la promotion des C-ITS.

1.1.3.1 Projets démonstratifs emblématiques

Trois initiatives principales sont remarquables autour du globe. Il s'agit de projets européens, japonais et américains. Souvent précurseurs en matière de transport, les États-Unis comptent sur la dotation de leurs *Departments of Transportation* pour développer des projets innovants. En 2012 la NHTSA (*National Highway Transports Safety Administration*) a lancé un projet pilote dans le Michigan prévoyant le déploiement de 3000 véhicules connectés. Il s'agit de la première initiative à grande échelle du genre, et les objectifs étaient avant tout d'étudier une baisse des accidents sur ces véhicules connectés. Porté par l'Université du Michigan, ce projet nommé *Connected Vehicle Safety Pilot*¹ a mené à une réflexion portant sur les questions de la propriété des données dans les ITS. En plus des problématiques liées à la sécurité routière, la sécurité des données (anonymisation) est un des freins possibles au déploiement des systèmes coopératifs.

Le Japon a été le site des premiers essais portés sur l'infrastructure connectée. En 2009, une expérience nommée *ITS Spot Service* a permis l'installation de 1600 bornes de recueil d'informations avec un espacement important (de 5 à 10 km), favorisant les échanges en V2V. L'architecture prévoyait donc un enregistrement des données sur le trajet et une restitution asynchrone aux UBR. Lors de l'expérimentation en 2015, ces données, utilisées dans le cadre d'alertes dans les véhicules, ont montré le potentiel d'amélioration des conditions de circulation sur les routes équipées.



Fig. 1.5: Le C-ITS Corridor européen
(Source : <http://www.c-its-korridor.de>)

Du côté du vieux continent, l'initiative la plus remarquable est le Corridor C-ITS lancé en 2013 qui relie Rotterdam à Vienne, via Francfort (figure 1.5). Le projet prévoit deux types de service principaux : les alertes chantiers et les *probe services*, i.e. les services de gestion du trafic qui reposent sur des données individuelles (position et vitesse du véhicule partagées) et liées au réseau (stations météorologiques). L'effort est porté sur le développement d'une interface unifiée pour les applications C-ITS dans les trois pays, c'est-à-dire allant vers une harmonisation des standards V2X. Parmi les enjeux visés par le projet, on retrouve les thématiques de la sécurité, de l'amélioration de l'écoulement du flux, ainsi que la réduction des émissions de polluants. Des tests d'interopérabilité sont prévus avec SCOOP, le projet pilote de déploiement des C-ITS en France.

1.1.3.2 Le cas français : le projet SCOOP@F

En France, l'intérêt porté au déploiement des C-ITS s'est exprimé à travers une continuité dans des projets de recherche financés par le Ministère de l'Écologie, du Développement Durable et de l'Énergie (MEDDE). Les premiers tests à petite échelle ont été menés dans le cadre



Fig. 1.6: Logo du projet SCOOP@F

1. Site internet du projet Safety Pilot : <http://safetypilot.umtri.umich.edu/>

du projet SCORE@F (Système Coopératif Routier Expérimental Français), entre 2011 et 2013. Sur les bases encourageantes apportées par cette première étape, le MEDDE a ensuite présenté en mai 2014 un projet pilote plus ambitieux nommé SCOOP@F². Il prévoit l'équipement et le déploiement de 3000 véhicules sur 2000 km de routes et rues. La phase de déploiement est planifiée à partir de l'année 2016 lors d'une expérimentation à grande échelle sur des sites pilotes. Les études d'impact menées sur les données récoltées permettront d'évaluer les bénéfices du système SCOOP dès 2018.

Ce projet compte beaucoup de partenaires privés et publics, comme des collectivités locales, des gestionnaires du réseau routier, des constructeurs automobiles (PSA et Renault) et des académiques et instituts de recherche (e.g. IFSTTAR³, Université de Reims Champagne-Ardenne, Institut Mines-Télécom). Depuis janvier 2016, de nouveaux partenaires ont rejoint le projet : un opérateur télécom (Orange), un fournisseur de services de sécurité (Opentrust) et aussi des partenaires étrangers d'Autriche, d'Espagne et du Portugal.

Afin de proposer une grande diversité de types de routes (autoroutes, voies rapides urbaines, bidirectionnelles, etc.), cinq sites pilotes ont été retenus :

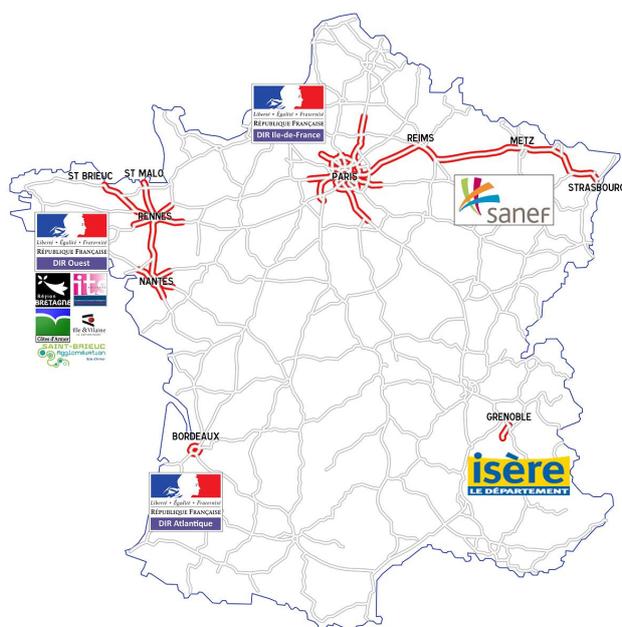


Fig. 1.7: Carte des sites pilotes du projet SCOOP@F

- Le réseau routier national non concédé de la région Île-de-France (partenaire : DIR⁴ Île-de-France), avec une grande partie des équipements déployés sur l'A86 et la Francilienne (2^{ème} et 3^{ème} rocades de Paris) ;
- Plusieurs routes en Bretagne (DIR Ouest, conseils départementaux des Côtes-d'Armor et d'Ille-et-Vilaine, conseil régional de Bretagne, Saint-Brieuc Agglomération, ITS Bretagne) ;
- Le "Corridor Est" sur l'A4 Paris-Strasbourg (SANEF⁵) ;
- La rocade de Bordeaux (DIR Atlantique) ;
- Des routes départementales en Isère (conseil départemental).

2. Site internet du projet SCOOP : <http://www.scoop.developpement-durable.gouv.fr>

3. Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux <http://www.ifsttar.fr>

4. Direction des routes

5. Société des Autoroutes du Nord et de l'Est de la France

Les objectifs de SCOOP reposent sur les 4 grandes catégories suivantes :

- **Amélioration de la sécurité routière** pour les usagers et les personnels des gestionnaires d'infrastructure grâce à l'échange d'informations de véhicule à véhicule et entre véhicules et infrastructure ;
- **Optimisation de la gestion de trafic et l'information routière**, notamment en temps réel, de façon à réduire la congestion, et contribuer ainsi à la réduction des émissions de polluants ;
- **Optimisation des coûts de gestion de l'infrastructure** tout en offrant de nouveaux services à l'usager ;
- **Développement du véhicule de demain**, le secteur des ITS coopératifs offrant un potentiel significatif de croissance et de création d'emplois au service de la sécurité du conducteur.

Enfin, ce projet s'intéresse à la définition des cas d'usages possibles des C-ITS, qui s'inscrivent directement dans la réalité des technologies aujourd'hui matures et disponibles. L'objectif des expériences conduites en 2016 et 2017 est de tester et d'évaluer les services déployés dans le cadre du projet. Six groupes de services ont été proposés dans SCOOP :

- *Collecte de données*. Il s'agit de données transmises automatiquement par les véhicules, ou sur initiative du conducteur, telles que position, vitesse de déclenchement d'airbags, d'ABS, d'ESP, de feux de détresse, ou encore un signalement d'animaux ou de personnes sur la route, brouillard, fortes pluies, véhicule à contre-sens, etc.
- *Alertes relatives aux chantiers*. Il s'agit de données transmises par le gestionnaire d'infrastructure : début et fin de zone de chantier, intervention sur les voies, véhicules de viabilité hivernale, etc.
- *Signalisation embarquée (indications de conduite)*. Il s'agit d'afficher sur le tableau de bord la signalisation de bord de route, les indications des Panneaux à Message Variables (PMV), etc.
- *Signalisation embarquée (événements inopinés)*. Les informations suivantes peuvent ainsi être transmises par les gestionnaires d'infrastructure ou par d'autres véhicules : route glissante, animaux, individus ou obstacles sur la route, zone d'accident non sécurisée, conducteurs à contre-sens, freinage d'urgence, etc.
- *Information sur le trafic routier*. Ces informations transmises par les gestionnaires sont par exemple l'indication du niveau de trafic, le temps de parcours, les re-routages conseillés, etc.
- *Parcs-relais et multi-modalité*. Ces informations concernent l'occupation des parcs-relais, les informations sur les horaires des transports en commun, etc.

Ces groupes de services s'attaquent majoritairement à l'enjeu de la sécurité routière. Cette tendance est au cœur de la problématique de l'acceptabilité des systèmes comme celui déployé dans le cadre de SCOOP. La phase de déploiement à grande échelle, qui se déroulera dans la deuxième phase du projet (entre 2017 et 2018) ne se focalisera pas forcément sur tous ces services, mais on peut s'attendre à ce que la sécurité fasse partie des enjeux majeurs étudiés.

1.1.4 Enjeux des C-ITS

Les projets de déploiement des C-ITS sont motivés par des enjeux sociétaux majeurs qui influenceront (positivement ou négativement) l'avenir des Systèmes Coopératifs. Les projets actuels se limitent à des projets de déploiement localisés sur des sites pilotes. L'étude des impacts des systèmes

développés se limitera donc à des cadres restreints et à des expérimentations ciblées. Néanmoins, plusieurs tendances sont à intégrer dans les objectifs de nos travaux.

Du côté du réseau étudié, on constate que les autoroutes, axes souvent sujets à la congestion, sont privilégiées. Cela est à mettre en rapport avec l'investissement pour le moment très élevé que représente l'implantation d'unités de bord de route. L'impact attendu du système est aussi très dépendant du taux d'équipement des véhicules (on parle de taux de pénétration), et malgré une augmentation significative, le nombre de véhicules déployés reste encore faible. Il est donc important de proposer une étude de plus grande envergure, pour confirmer le potentiel pressenti des C-ITS. Cette étude, telle que nous la proposons dans le cadre de notre travail, aura besoin d'un outil de simulation pour reproduire les comportements des véhicules équipés et leurs interactions avec l'infrastructure. Dans un premier temps, le scénario du trafic autoroutier semble être un point de départ favorable à l'apparition d'impacts positifs sur le trafic si l'on se fie aux premiers résultats obtenus dans les expérimentations.

Pour observer et analyser l'impact du déploiement des C-ITS, nous serons amenés à choisir des indicateurs qui seront liés aux enjeux principaux de tels systèmes. Nous proposons de nous concentrer sur les enjeux majeurs qui motivent le déploiement des C-ITS : la **sécurité**, la **gestion du trafic** et l'**environnement**.

1.2 Défis et hypothèses

Le domaine d'étude des Systèmes Coopératifs propose un cadre de recherche exceptionnellement riche, de par sa nouveauté, sa complexité, et ses enjeux. Derrière l'étude de ces systèmes se cache en fait la volonté d'accélérer leur déploiement en dévoilant leur potentiel futur. Les études actuelles se limitent, à cause de la faible disponibilité des véhicules connectés et autonomes, à des expérimentations à plus petite échelle. Bien que prometteurs, les résultats ne garantissent pas que les effets observés perdureront lorsque le taux de pénétration sur le marché augmentera. Le défi est alors d'anticiper cette phase en proposant des modèles capables, par leur réalisme, de convaincre tous les acteurs des C-ITS.

1.2.1 Verrous scientifiques

Plusieurs problématiques restent encore aujourd'hui non résolues. Les véhicules connectés et autonomes (VCA) ne sont pas encore suffisamment déployés pour permettre une étude à grande échelle. Aussi, la modélisation et la simulation restent une solution pour l'évaluation et l'extrapolation des effets à large échelle (montée en échelle). Un élément critique est que dans cette période de déploiement progressif qui prendra plusieurs années, les VCA devront cohabiter avec des véhicules standards. Cette cohabitation constitue un des défis au regard de l'efficacité des stratégies de régulation fondées sur la coopération et l'automatisation.

1.2.1.1 De la modélisation à l'aide à la décision

L'objectif des approches classiques en simulation est de reproduire le plus fidèlement possible un phénomène physique grâce à des modèles mathématiques. La finalité de ces simulations n'est pas la performance pure, mais plutôt la vérification des modèles qui peuvent par la suite servir de base pour l'action et prévoir le problème modélisé. Cela implique que le concepteur doit disposer à la fois d'un grand nombre de données du problème mais aussi confronter ses modèles au plus de situations possibles. Pour certains problèmes, les états atteints par le système réel étudié ne sont pas maîtrisables ou atteignables, ce qui peut compliquer la phase de validation des modèles. Dans ce contexte, comment proposer un cadre de simulation suffisamment réaliste pour étudier le comportement de futurs systèmes complexes ? La résolution de ce verrou permet de transformer un outil de simulation en un outil d'aide à la décision.

1.2.1.2 Stratégies de gestion du trafic : contrôle et régulation

Les phases successives de déploiement des Systèmes Coopératifs s'accompagneront de l'émergence de nouveaux comportements. Ces comportements pourront être introduits par les nouveaux véhicules (connectés et/ou autonomes), ou résulter de la réaction des usagers face aux nouvelles technologies. L'apport des technologies de la communication ne se limitera pas à l'adaptation des systèmes existants mais permettra aussi d'imaginer de nouvelles solutions pour les acteurs des C-ITS. Les enjeux majeurs de la mobilité de demain sont liés aux problèmes de saturation des réseaux. Les Systèmes Coopératifs doivent apporter des réponses à ces défis pour que leur déploiement soit un réel succès. Les approches de contrôle et de régulation doivent être repensées ou adaptées. L'une des possibilités est d'utiliser une partie des Systèmes Coopératifs qui est souvent négligée dans les approches existantes : l'infrastructure. Quelle forme de contrôle doit-on alors envisager ? Comment intégrer les problématiques de fiabilité de l'information ? Les stratégies de contrôle seront-elles identiques pour toutes les phases du déploiement des C-ITS ?

1.2.2 Objectifs, contributions et hypothèses

L'objectif final de notre travail est de proposer un outil d'aide à la décision qui permettra de pointer les forces et les faiblesses des Systèmes Coopératifs. Difficile dans ce cas de limiter l'étude à certains scénarios ou cas d'usage. Notre travail propose de poser les fondations d'un outil complet dédié à la modélisation des Systèmes Coopératifs. Cette modélisation, qui repose sur des concepts tirés de l'intelligence artificielle distribuée, est intégrée dans un outil de simulation. Ce simulateur constitue un outil d'aide à la décision permettant de concevoir et d'implémenter de nouvelles stratégies dédiées aux C-ITS. Ces stratégies, qui seront déployées progressivement, peuvent mener à l'apparition de nouveaux comportements observés à l'échelle des flux. Les approches classiques de contrôle risquent alors de ne pas s'adapter en contexte de flux mixtes (véhicules connectés et autonomes). Cela nous oriente vers l'élaboration de stratégies de contrôle, qui trouvent des analogies fortes avec les problématiques d'apprentissage en IA.

En résumé, les contributions saillantes de notre travail sont :

- un cadre innovant de **modélisation fine du trafic coopératif**, intégrant la **communication**, la **coopération** et les problématiques de **contrôle** et de **régulation** grâce aux approches **multi-agents**.
- un **outil de simulation** robuste et performant permettant d'**étudier le comportement** des nouveaux véhicules et **des flux**.
- une étude des **bénéfices** et des **limitations attendues des Systèmes Coopératifs** durant leurs phases de déploiement successives : l'outil de simulation devient alors un **outil d'aide à la décision**.
- la **proposition de nouvelles stratégies de contrôle** qui s'adaptent à l'émergence de nouveaux comportements induits par le déploiement des C-ITS.

Le cadre de l'étude proposée dans notre travail est en relation avec les enjeux identifiés (section 1.1.4). Nous visons les sections d'autoroutes, et notamment péri-urbaines, qui sont préférées dans les projets de déploiement actuels (section 1.1.3). Les véhicules les plus représentatifs du trafic (voitures) sont modélisés sans pour autant limiter la portée de l'étude aux autres véhicules (notamment les poids-lourds, mais pas les deux-roues). Les enjeux visés sont à l'échelle du flux ; cette hypothèse influence le choix de l'échelle de modélisation et de simulation. Notre étude vise à étudier différentes phases de déploiement. Les approches de contrôle proposées s'appuieront donc sur différents niveaux d'automatisation.

1.3 Approche méthodologique

L'ensemble des contributions de notre travail valorise la simulation comme un véritable outil d'aide à la décision. Ce type d'outil permet de concevoir de nouvelles solutions (par exemple, pour des problématiques de contrôle) mais permet aussi d'étudier le fonctionnement de systèmes complexes, avant leur déploiement réel. Afin de mettre en place cet outil d'aide à la décision, nous avons suivi une démarche méthodologique partant de la modélisation physique jusqu'aux problématiques de contrôle.

1.3.1 Aperçu

Le présent manuscrit est organisé en trois chapitres principaux reprenant les contributions portant sur les thématiques de la modélisation, de la simulation et du contrôle (figure 1.8).

Un second axe de lecture possible est lié aux phases de déploiement des Systèmes Coopératifs. Une première forme de contrôle local correspond à un cadre de déploiement plus futuriste où tout ou partie du flux est automatisé. Dans un futur plus proche, une stratégie de contrôle décentralisée permettra de mieux s'adapter à l'apparition progressive de nouveaux comportements accompagnant les phases de déploiement, en rendant les systèmes plus autonomes.

1.3.2 Organisation du manuscrit

Le **chapitre 1** permet de prendre un peu de recul sur l'état actuel du déploiement des Systèmes Coopératifs et les technologies impliquées. Les premiers projets pilotes sont encore limités à l'étape

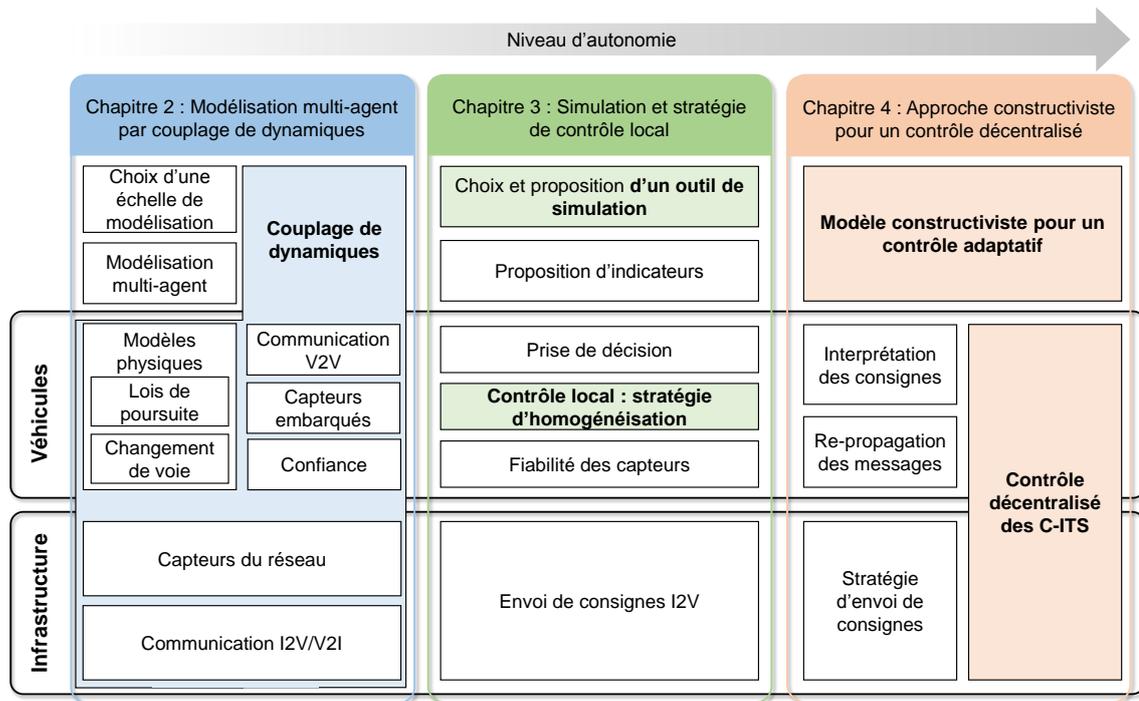


Fig. 1.8: Illustration de l'organisation du manuscrit, permettant de visualiser les contributions en rapport avec les différentes phases de déploiement des Systèmes Coopératifs.

de prototype. Pourtant, il est crucial de montrer le potentiel de ces systèmes pour motiver des déploiements à plus grande échelle.

Nous proposons dans le **chapitre 2** une étude approfondie des approches dédiées à la modélisation dans la littérature. La complexité des interactions et l'analogie des C-ITS avec les systèmes complexes et les systèmes distribués nous orientent vers l'étude du paradigme multi-agent. Les analogies ne s'arrêtent pas à la topologie des systèmes mais touchent aussi le cœur des interactions dans les Systèmes Coopératifs : la communication. Le besoin d'un système de gouvernance nous amène à étudier les modèles basés sur les concepts de confiance et de réputation. La coexistence des différents modèles, régissant les dynamiques du système impose une modélisation particulière. Nous proposons ainsi une modélisation des C-ITS par couplage de ses dynamiques.

La modélisation proposée dans le cadre de ce travail sert de base à l'élaboration d'une solution de simulation dédiée à l'étude et à la proposition de solutions pour les Systèmes Coopératifs. Nous proposons dans le **chapitre 3** un système local de régulation prenant la forme d'une stratégie d'homogénéisation d'un flux partiellement composé de véhicules connectés. Cette stratégie est en réalité la concrétisation du couplage des dynamiques du système. Les résultats obtenus par les implémentations des différents modèles permettent d'améliorer le flux tout en garantissant sa sécurité même dans un contexte de capteurs défaillants. Cette étude met en avant l'apparition de nouveaux comportements résultant de la coopération entre les entités du système. Les stratégies de contrôle et de régulation existantes montrent alors leurs limites dans ce contexte. Nous proposons donc d'étudier des nouvelles formes de contrôle permettant de profiter des bénéfices des C-ITS en intégrant les enjeux identifiés.

Le **chapitre 4** se concentre sur la proposition d'un modèle permettant à un système autonome de construire une représentation de son environnement. Cette représentation vise à lui apporter un meilleur contrôle de son environnement. La présentation de ce modèle est volontairement abstraite. Cette volonté s'explique par les analogies de notre problème avec les problématiques rencontrées en IA. La phase d'implémentation est décisive dans la phase d'exploitation du modèle. Afin d'illustrer les contributions associées, nous étudions le comportement du système dans deux environnements très distincts. La première étude vise à mettre en avant les capacités générales du modèle proposé d'un point de vue générique. Nous proposons donc des simulations dans un cadre dédié permettant de s'abstraire des biais d'un environnement réel. La seconde partie des expérimentations consiste à implémenter le modèle pour la problématique de contrôle des C-ITS. Cette dernière étape illustre l'intégralité de la réflexion ayant mené à l'élaboration de ce manuscrit. Les résultats produits ouvrent des perspectives à la fois du point de vue conceptuel et du point de vue des Systèmes Coopératifs.

Couplage de dynamiques pour la modélisation du trafic coopératif

” *The way to build a complex system that works is to build it from very simple systems that work.*

— **Kevin Kelly**
Écrivain

Le déploiement des systèmes de transports intelligents et plus particulièrement des systèmes coopératifs s'appuie sur la démocratisation des capteurs embarqués dédiés à la perception et à la communication. Comme nous l'avons vu dans le chapitre 1, ces nouveaux véhicules équipés et connectés pourront compter sur une infrastructure dédiée capable d'étendre leur perception de l'environnement. L'automatisation connaît un regain d'intérêt ces dernières années, et certains véhicules de série permettent déjà de soustraire tout ou partie de la tâche de conduite du conducteur. Du système avancé d'assistance au conducteur au véhicule autonome (Google, Tesla), le socle technologique semble propice au développement de nouveaux usages d'une infrastructure qui souffre de plus en plus d'un usage égo-centré des véhicules. Dans ce contexte, une voie d'amélioration possible est d'émuler une forme de coopération entre véhicules et avec l'infrastructure. Cette nouvelle forme d'interaction doit prendre en compte toutes les sources d'information (capteurs, communication) et s'intégrer dans la chaîne d'actions, du gestionnaire jusqu'au conducteur, à travers le comportement de son véhicule.

La phase de déploiement actuelle des systèmes coopératifs constitue une phase de prototypage. Grâce à des projets démonstrateurs, le réseau routier commence désormais à être parcouru par quelques véhicules capables d'échanger des informations à travers une interface unifiée (section 1.1.3). Les possibilités offertes par ces véhicules ne sont encore pas pleinement exploitées. En effet, plusieurs degrés de liberté sont à explorer : automatisation, optimisation, contrôle, régulation, anticipation, etc. Il est crucial de prendre en compte tous les acteurs finaux impliqués dans le déploiement et l'utilisation de ces futurs systèmes. Enfin, la sécurité des conducteurs (et en général, des usagers du réseau) étant prioritaire, une phase de conception, de test et de validation est préalable au déploiement. Ainsi, l'objectif est double : il s'agit de proposer de nouveaux comportements répondant aux objectifs précédents, mais aussi de convaincre les acteurs (gestionnaires, constructeurs, conducteurs) de l'intérêt à porter aux systèmes coopératifs. La première étape de ce travail passe par une étude en simulation. Dans un environnement virtuel, il est possible de reproduire le comportement d'un système observé mais aussi de le confronter à des situations inédites. Dans l'optique d'un déploiement progressif de véhicules coopératifs sur un réseau routier, ces deux étapes sont nécessaires. Dans un premier temps, il est possible de modéliser le réseau ainsi que les flux de trafic. Cela impose de choisir un niveau de modélisation adapté à l'étude des comportements à différentes échelles (section 2.1). En effet, les systèmes coopératifs se fondent sur de nouvelles formes d'interactions entre ses acteurs pour permettre une meilleure utilisation du

réseau à l'échelle des flux. Le domaine de la modélisation du trafic propose des solutions permettant, à travers le choix ou la proposition de modèle, de reproduire ce type de comportements. Le défi pour notre travail est de concilier les besoins d'une modélisation fine avec les enjeux de haut niveau que nous avons identifiés.

Notre proposition s'appuie sur les travaux récents du domaine de la modélisation du trafic, couplés avec les technologies issues du monde de l'Intelligence Artificielle. La sécurité des communications est une préoccupation principale des C-ITS. Nous l'intégrons dès la modélisation à travers le concept de confiance (section 2.2). Les modèles choisis s'intègrent dans un couplage multi-agent des dynamiques des systèmes coopératifs (section 2.3), qui permet de proposer une stratégie locale de régulation allant vers l'homogénéisation d'un flux de véhicules connectés (section 2.4).

2.1 Approches distribuées et modélisation du trafic

La modélisation du trafic, comme toute modélisation, a pour objectif de reproduire la dynamique des flux de trafic, ce qui ouvre la voie vers l'élaboration de stratégies de gestion du trafic. La première question qu'un modélisateur doit se poser est relative à l'échelle de la modélisation. Le système de trafic étudié doit-il être étudié à une grande échelle ou cherche-t-on à analyser les mouvements d'un peloton de véhicules ? Ces niveaux de modélisation ne sont pas clairement segmentés. Pourtant, tous ont un fondement théorique très proche : il s'agit de la théorie du trafic. Notre étude présente les approches des plus globales aux plus locales. Après une revue des différents modèles existants et de leur niveau de modélisation associé (section 2.1.1), nous proposons d'étudier les approches distribuées et notamment les systèmes multi-agents (SMA). Ce paradigme de l'Intelligence Artificielle Distribuée (IAD) semble en effet apporter des solutions pour les besoins de modélisation de systèmes aussi complexes que les C-ITS. L'analogie entre les disciplines de la modélisation de trafic et des SMA s'exprime aussi dans les approches existantes mêlant les deux problématiques. Pourtant, nous montrons que peu de travaux existants proposent de solutions allant dans ce sens, ce qui motive notre réflexion vers une nouvelle forme de modélisation.

2.1.1 Modélisation du trafic

Le domaine de la modélisation du trafic est étudié depuis le début du développement et de la conquête de l'automobile [191]; dans une infrastructure routière qui devient de plus en plus limitée. Beaucoup de modèles ont été proposés, étudiés et comparés dans la littérature [88]. Ces modèles s'intéressent à la reproduction des comportements des flux de trafic ou/et des véhicules. L'objectif de cette section est de recenser les travaux existants, afin de choisir une échelle de modélisation adaptée aux systèmes coopératifs permettant d'étudier et de comprendre l'impact de la coopération sur le flux. Comme notre travail se focalise sur une période de cohabitation des véhicules classiques et des véhicules coopératifs (trafic mixte), il est important que leurs comportements soient proches.

Pour trouver le point de départ de l'attrait du domaine de la modélisation du trafic, il faut remonter aux années 30, avec les travaux précurseurs de Greenshields qui proposa des relations fondamentales entre le débit, la densité et la vitesse d'un flux de véhicules [79]. Tous les modèles sont issus de cette relation (section 2.1.1.1) ou plus généralement du "Diagramme Fondamental" (FD), et se sont ensuite distingués en trois échelles de modélisation, comme représentées dans la figure 2.1. Les

points noirs correspondent aux publications et les lignes noires aux successions de publications (mêmes modèles ou variantes).

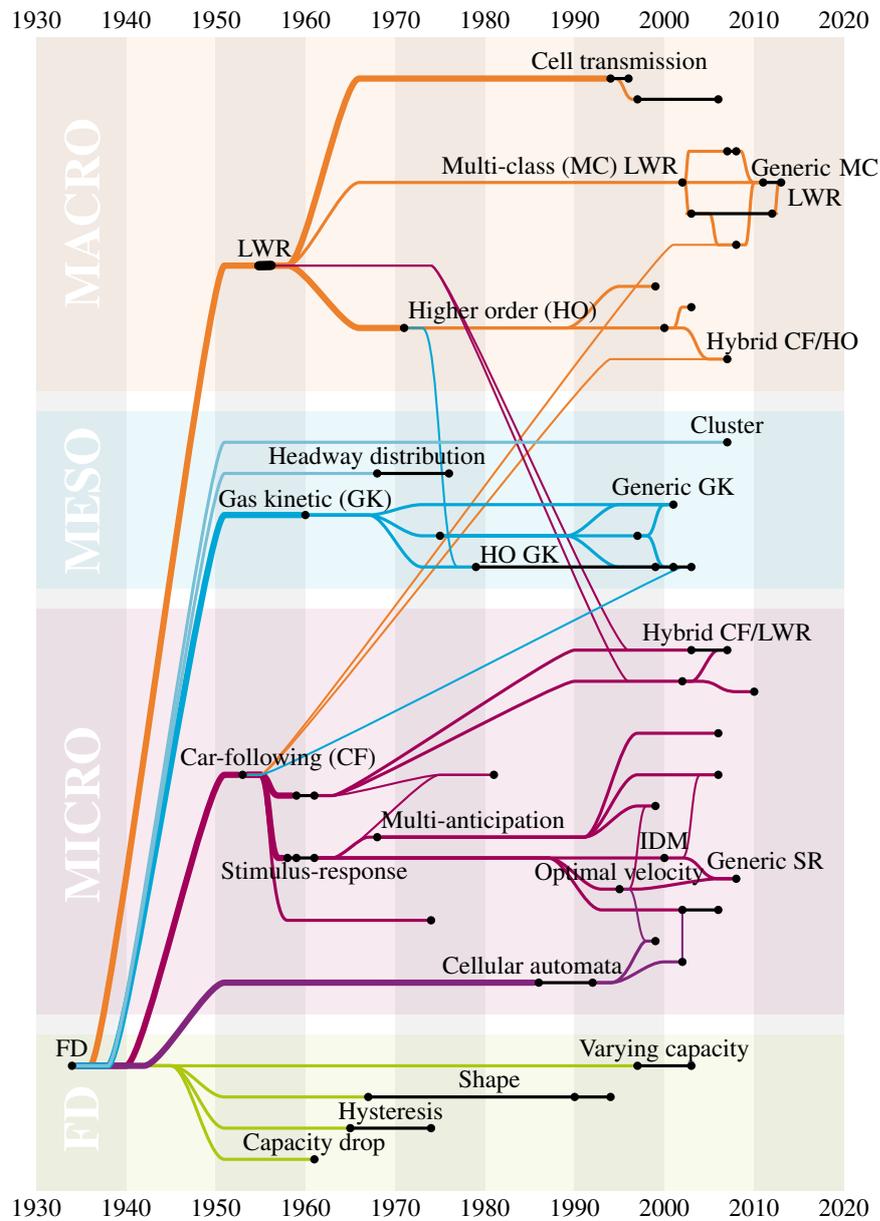


Fig. 2.1: "Arbre généalogique" des modèles de trafic, tiré des récents travaux de thèse de van Wageningen-Kessels [201].

Généralement, trois échelles de modélisation permettant de classer les modèles de trafic se distinguent. Les modèles microscopiques (section 2.1.1.4) visent à reproduire le comportement individuel de chacun des véhicules alors que les modèles macroscopiques (2.1.1.2) représentent le flux agrégé de véhicules. Le niveau mésoscopique (2.1.1.3) est intermédiaire et permet de caractériser un flux avec des données macro et micro ou des distributions statistiques (e.g. véhicules partageant des caractéristiques semblables : origine-destination, vitesse, ...). Bien que largement admise, cette classification n'est pas unique et d'autres critères permettent de comparer les modèles : variables

discrètes ou continues, niveau de détail, représentation déterministe ou stochastique, ou encore l'échelle d'application (taille du réseau) [92, 195].

2.1.1.1 Diagramme fondamental

La plupart des modèles de trafic reposent sur la mécanique des fluides. Les modèles reproduisent le comportement de l'écoulement du flux de trafic. Rapidement, des travaux se sont basés sur l'hypothèse de l'existence d'une forme de relation liant la distance entre les véhicules (*i.e* inter-véhiculaire) et leur vitesse. Cette relation d'abord caractérisée par Greenshields [79] a ensuite pris le nom de relation fondamentale. Prenons l'exemple d'une portion (section) de route sans entrée ni sortie (sans rampe d'insertion/sortie). Alors la relation fondamentale (simplifiée) qui lie le débit Q (véhicule/h), la concentration K (véhicule/km), et la vitesse moyenne du flux V (km/h) s'écrit :

$$Q = K.V \quad (2.1)$$

Si l'on observe les trajectoires d'un flux de véhicules au cours du temps (diagramme espace-temps représenté par la figure 2.2), on retrouve la définition de ces variables à partir des travaux de Edie [57]. Le débit Q est considéré comme le volume de véhicules qui traversent une zone Z de longueur dx durant une période de temps dt . Ce débit dépend donc du nombre de véhicules ayant parcouru la zone et de la distance d_n parcourue par chaque véhicule durant son temps de trajet t_n . La concentration K utilise les mêmes variables et s'exprime comme un nombre de véhicules par unité de distance répartis sur une zone Z spatiale donnée. Enfin, la vitesse V dans la zone Z s'exprime alors comme la distance totale parcourue divisée par le temps total passé.

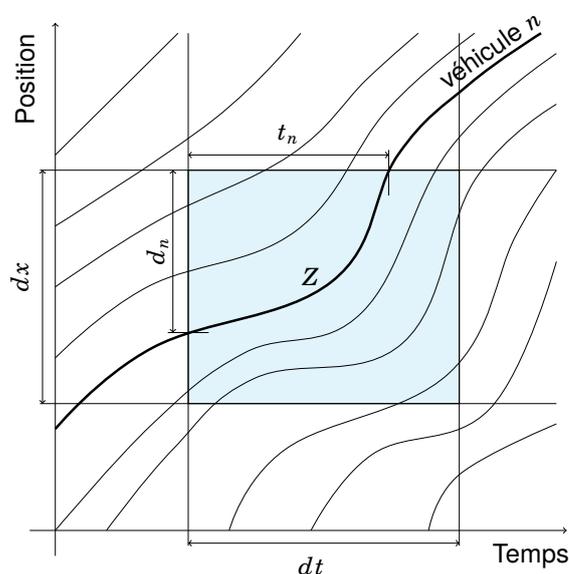


Fig. 2.2: Exemple d'observations spatio-temporelles de trajectoires individuelles ou "diagramme espace-temps".

La relation liant débit, vitesse et concentration peut être représentée sous la forme d'un diagramme fondamental qui permet d'estimer les débits maximaux et de distinguer les régimes fluide (à gauche

de la valeur de concentration critique) et congestionné (partie droite), comme décrit dans la figure 2.3.

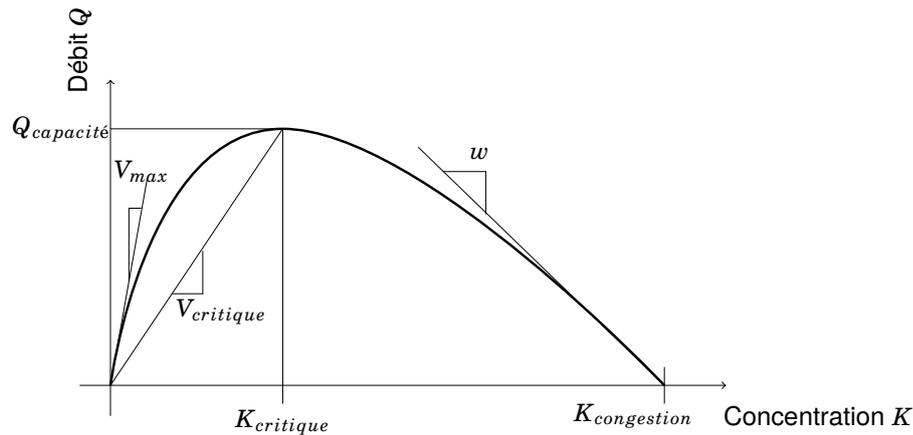


Fig. 2.3: Schéma d'un diagramme fondamental (débit/concentration)

Le diagramme fondamental représente les paramètres inhérents à la relation fondamentale débit/concentration/vitesse. On y retrouve certaines des propriétés énoncées par Del Castillo [44] :

- La vitesse est comprise entre zéro et V_{max} ;
- La concentration évolue entre zéro et $K_{congestion}$;
- Les vitesses pour les valeurs de concentration extrêmes sont $V(0) = V_{max}$ et $V(K_{congestion}) = 0$;
- Les débits pour les valeurs de concentration extrêmes sont $Q(0) = Q(K_{congestion}) = 0$;
- La vitesse maximum V_{max} et la vitesse de remontée de la congestion w sont les pentes du diagramme fondamental aux valeurs de concentration extrêmes.
- La capacité $Q_{capacité}$ d'un arc du réseau (section) est limitée et est atteinte pour des valeurs de vitesse $V_{critique}$ et de concentration $K_{critique}$ qui peuvent être connues empiriquement ou estimées.

Ces propriétés dépendent de paramètres qui sont fixés par le réseau étudié. Dans la suite de notre travail, nous serons amenés à nous reporter à ces propriétés qui permettent de caractériser et d'agir contre la congestion.

À des fins calculatoires, le diagramme fondamental est souvent simplifié par sa forme triangulaire [89] qui permet plus facilement d'analyser numériquement les phénomènes au niveau macroscopique. Ce diagramme servant de base aux modèles de trafic, plusieurs évolutions successives [39, 73, 38] ont permis d'appliquer cette relation à l'échelle macroscopique, permettant d'étendre la zone d'analyse à un réseau complet.

2.1.1.2 Modèles Macroscopiques

La modélisation macroscopique vise à caractériser et reproduire les phénomènes observés à grande échelle. Une première théorie du trafic a été construite à partir de l'hydrodynamique (mécanique des fluides) par Lighthill et Whitham [121, 122], puis Richard [173]. Les modèles macroscopiques

sont principalement des dérivés de la théorie de Lighthill-Whitham-Richards (LWR). Après avoir été considéré comme référence, ce modèle a récemment été critiqué [37], notamment de par les lacunes en termes de reproduction de phénomènes essentiels associés au trafic : la formation et la propagation d'ondes de congestion. D'autres travaux ont ensuite porté un nouvel intérêt aux modèles du second ordre comme le modèle de Aw-Rascle [10].

Le point de vue adopté par ce type de modèle rend difficile la modélisation des interactions fines induites par la communication. Seuls quelques travaux très récents basés sur le MFD (diagramme fondamental macroscopique), introduit par Geroliminis et Daganzo [73, 38], commencent à émerger [214]. Ils proposent d'évaluer l'effet de la coopération et des échanges d'information à l'échelle d'un réseau, mais se limitent à la reproduction de comportements observés (le modèle se base sur des données réelles).

2.1.1.3 Modèles Mésoscopiques

Le niveau intermédiaire dénommé mésoscopique ou cinétique, fait le lien entre les niveaux microscopique et macroscopique. L'approche consiste à considérer non plus le flux dans la globalité ou par ses individus, mais plutôt à modéliser le comportement soit de paquets de véhicules (pelotons), soit de distributions de véhicules ayant des caractéristiques proches (en termes de vitesse et/ou d'accélération).

On retrouve des modélisations à ce niveau basées sur les équations de Boltzmann comme le modèle de Prigogine et Andrews [166, 167], qui présente l'avantage d'être dérivable en modèle macroscopique ou microscopique. Des améliorations de ces modèles ont été proposées par la suite [160]. D'autre part, Illner *et al.* [95] ont montré qu'il était possible de dériver le modèle macroscopique de Aw-Rascle [10], ce qui confirme les transitions possibles entre les modèles.

Ce niveau de modélisation est intéressant puisqu'il permet d'estimer des paramètres individuels à partir de données agrégées (flux). Grâce aux transitions possibles entre les modèles, il est possible de modéliser les comportements de véhicules automatisés au niveau mésoscopique [119], mais la complexité des interactions (notamment la communication) oriente davantage vers les modèles microscopiques.

2.1.1.4 Modèles Microscopiques

La modélisation microscopique du trafic s'intéresse à la description des comportements individuels des véhicules (vision "particulaire" [88]). Dans cette modélisation, chaque véhicule se voit associer une loi de poursuite qui permet de modéliser sa cinématique en fonction du véhicule qui le précède. Si le véhicule courant est noté n , les véhicules prédécesseurs (*leader*) sont numérotés à partir de $n + 1$ et les suiveurs (*follower*) à partir de $n - 1$, comme représenté dans la figure 2.4.



Fig. 2.4: Numérotation des véhicules dans les modèles microscopiques.

En reprenant les critères de comparaison de Hoogendoorn et Bovy [92], les modèles peuvent être soit discrets soit continus. On parle alors d'automates cellulaires pour le cas discret. Les modèles continus prennent la forme de modèles nommés "lois de poursuite" (*car-following*), où le véhicule suiveur adapte sa vitesse et son accélération en fonction de la perception de son véhicule prédécesseur (ou de ceux de son entourage dans le cas de la multi-anticipation [118]). Ces deux types de modèles se limitent à reproduire le comportement longitudinal (accélération); ils sont complétés par des modèles de changement de voie qui permettent de modéliser le comportement latéral des véhicules.

Automates cellulaires

Parmi les modèles microscopiques, on distingue les modèles discrets (temps/espace) des modèles continus. Les automates cellulaires sont des modèles qui utilisent une représentation discrète de l'espace (le réseau). Ces modèles, plus anciens que les lois de poursuite, sont généralement plus rapides à calculer [110]. Le principe des automates cellulaires est de découper la voie de circulation en cellules (figure 2.5) pouvant contenir un véhicule ou aucun (la longueur couramment utilisée est 7,5 mètres, ce qui correspond à la longueur moyenne d'un véhicule plus une distance pare-choc à pare-choc à l'arrêt).



Fig. 2.5: Exemple de discrétisation d'un réseau pour les automates cellulaires.

Le fonctionnement simplifié d'un tel modèle consiste à estimer si un véhicule va ou non se déplacer d'une (ou plusieurs) cellule(s), vers l'aval. Le premier modèle est celui de Cremer et Ludwig [36]. Après quelques évolutions, le modèle de Nagel et Schreckenberg [144] est désormais considéré comme le prototype des automates cellulaires [110]. Bien que certains de ces modèles soient capable de reproduire le flux de trafic, la limitation intrinsèque de la représentation du réseau les rend difficilement utilisables pour modéliser des interactions fines (*e.g.* calcul de distance inter-véhiculaire). Ce niveau de détails est en revanche atteignable par les modèles de type "lois de poursuite".

Lois de poursuite

La littérature recense différentes classes de modèles microscopiques continus (voir la revue dans [194]). Ces modèles aussi appelés lois de poursuites fonctionnent sur un espace et avec un temps tous deux continus. L'accélération longitudinale est calculée à partir des paramètres du véhicule prédécesseur et des paramètres intrinsèques du véhicule courant, comme illustré dans la figure 2.6. Le véhicule courant accélère (ou décélère) en fonction du véhicule qui le précède sur une ligne virtuelle représentant l'axe médian de la voie de circulation courante (modélisée en bleu sur la figure 2.6).

Parmi les modèles les plus utilisés, le modèle de Gipps [76] constitue une des premières approches de la description du comportement individuel des véhicules en déterminant l'accélération d'un véhicule suiveur au temps $t + 1$ par rapport à l'accélération et à la vitesse du véhicule précédant au temps t . Deux régimes sont distingués : un régime fluide où le véhicule limite sa vitesse (législation)

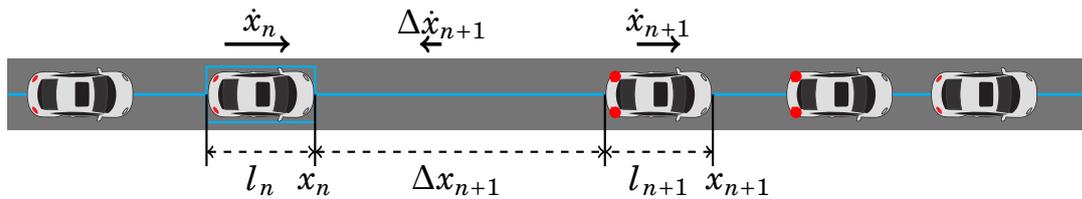


Fig. 2.6: Interprétation graphique des paramètres utilisés dans les lois de poursuites.

et un régime contraint où il cherche à maintenir une "distance de sécurité". D'autres modèles comme celui de Newell [146] se basent sur cette notion de distance cible, avec des paramètres simplifiés comme le temps de réaction.

Les modèles les plus récents, mais aussi les plus utilisés fonctionnent en temps continu. Il s'agit de modèles "réactifs" : l'hypothèse retenue est que le conducteur réagit dynamiquement aux trois *stimuli* représentés par les paramètres suivants : la vitesse de son véhicule, l'inter-distance avec le véhicule prédécesseur, et leur vitesse relative. Ainsi, ces lois de poursuite mettent en équation le calcul de l'accélération \ddot{x}_n du véhicule courant n (voir figure 2.6) suivant la formule :

$$\ddot{x}_n = f(\dot{x}_n, \Delta x_{n+1}, \Delta \dot{x}_{n+1}) \quad (2.2)$$

où f est l'équation du modèle considéré, \dot{x}_n la vitesse du véhicule courant, Δx_{n+1} l'inter-distance entre le véhicule courant et son prédécesseur (qui nécessite de connaître leurs positions x_n et x_{n+1} ainsi que leurs longueurs l_n et l_{n+1} respectives), et $\Delta \dot{x}_{n+1}$ leur vitesse relative. On retrouve clairement l'apport du diagramme fondamental (section 2.1.1.1) dans les premiers modèles comme *Optimal Velocity Model* (OVM [13]) et ses évolutions : *Optimal Velocity with Relative Velocity* et *Improved Optimal Velocity Model* (OVRV [98] et IOVM [120]); qui exploitent les régimes critiques de la théorie du trafic (section 2.1.1.1).

Les paramètres des modèles ont ensuite gagné en intelligibilité et en stabilité : ils décrivent le comportement des conducteurs par l'intermédiaire de leurs paramètres propres (agressivité, anticipation, etc.) et sont souvent sans collision. C'est par exemple le cas de *Intelligent Driver Model* (IDM) proposé par Treiber et Kesting [192, 108]. Face à l'augmentation significative du nombre de paramètres, le calage [103, 137] des paramètres des modèles vise à rapprocher les comportements simulés de comportements réellement observés (en partant de données de trajectoires réelles).

Les lois de poursuite sont utilisées dans le cadre d'un réseau modélisé sous la forme d'un graphe. Chacune des voies de circulation est alors une ligne virtuelle sur laquelle évoluent les véhicules modélisés. Dans le cas de réseaux plus complexes, il est nécessaire de modéliser les transitions entre les différentes lignes de ce réseau, c'est le rôle des modèles de changement de voie.

Modèles de changement de voie

L'amélioration des modèles microscopiques est passée par l'incorporation de modèles de changement de voie : on parle alors de modèles multi-voies. Cette problématique a d'abord été abordée au niveau macroscopique, comme en témoignent de nombreux travaux notamment recensés dans [116].

L'intégration de modèles de changement de voie aux lois de poursuite a aussi été très étudiée. Treiber, Kesting et Helbing ont proposé le modèle MOBIL (*Minimizing Overall Braking decelerations Induced by Lane changes*) intégré à leur modèle [192] et utilisable pour les autres modèles de type lois de poursuite [104]. Le principe est généralement similaire et dépend de la représentation du réseau. Les changements de voies sont instantanés et basés sur un processus de décision qui prend en compte une perception étendue (véhicule prédécesseur et véhicules présents sur la voie cible). Les trajectoires sont souvent interpolées pour affiner le mouvement latéral (comme illustré dans la figure 2.7, où les trajectoires interpolées sont représentées en pointillés).

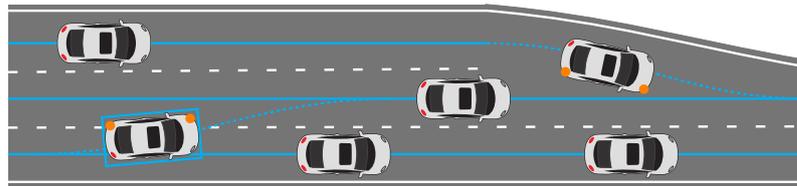


Fig. 2.7: Représentation multi-voie d'une section et modélisation du changement de voie.

Les modèles de changement de voie sont fondés sur la prise en compte d'un ou plusieurs véhicules (dans le cas de la multi-anticipation) avec la définition d'un intervalle minimum disponible pour entreprendre un dépassement [190]. Comme représenté dans la figure 2.7, certaines approches comme [104] permettent de distinguer les changements de voie obligatoires (*i.e.* imposés par le réseau : fermeture de voie, insertion, etc.) des changements de voie volontaires (dépassements). La décision du changement de voie se fait souvent suivant un arbre de décision (ou une machine à états finis), mais des travaux plus récents apportent des solutions plus complexes et plus réalistes [188].

2.1.1.5 Bilan des modèles classiques de trafic

Cette section a permis la synthèse des classes de modélisation de trafic existantes. Parmi les trois niveaux de modélisation, le niveau microscopique qui permet de modéliser les comportements individuels, apparaît comme le plus adapté à notre travail et le plus apte à proposer une grille d'analyse cohérente avec les objectifs de la thèse. Deux types de modèles sont couplés afin de modéliser le comportement final du véhicule : une loi de poursuite calcule l'accélération (comportement longitudinal) du véhicule et un modèle de changement de voie régit son comportement latéral. Nous avons montré que ces deux familles de modèles étaient compatibles, mais la plupart des travaux existants se limitent à une version simplifiée des interactions du véhicule. En effet, l'expression générale d'une loi de poursuite n'est basée que sur les paramètres d'un seul véhicule prédécesseur (équation 2.2). Cette hypothèse, qui permet de réduire les temps de calcul, est relativement naïve et ne correspond pas forcément au comportement observé des vrais conducteurs. D'autre part, nous cherchons à modéliser des comportements de véhicules coopératifs, c'est-à-dire potentiellement connectés et/ou autonomes. Il est donc important que le modèle utilisé soit compatible avec des sources d'information partagées ainsi qu'à des formes de contrôle adaptées. Les lois de poursuite présentées visent à modéliser le comportement d'un conducteur humain ; quelles modifications sont nécessaires pour reproduire le comportement d'un véhicule équipé ?

2.1.2 Modèles pour le trafic coopératif

Les modèles classiques de trafic ne prennent en compte ni la coopération entre les véhicules permise par la communication ni l'automatisation des véhicules. Dans ce travail, nous cherchons à modéliser finement ces nouveaux types d'interaction. Ce n'est que récemment, grâce à la généralisation des systèmes d'aide à la conduite et le développement de la communication inter-véhiculaire, que des approches de ce type ont vu le jour dans la littérature. Notre étude se focalise sur les modèles microscopiques, qui sont les plus représentés, même si la modélisation de la coopération est possible à d'autres échelles [147, 148].

2.1.2.1 Modélisation des systèmes ADAS

Un des premiers pas vers l'automatisation est de déléguer une partie de la tâche de conduite à un système embarqué. On parle alors de systèmes avancés d'assistance au conducteur (ADAS – section 1.1.1.2). Aujourd'hui, le système le plus représenté dans les véhicules est le régulateur de vitesse à contrôle de distance (*Adaptive Cruise Control* – ACC [211]). ACC repose sur l'utilisation d'un capteur embarqué à l'avant du véhicule capable de mesurer l'interdistance et la vitesse relative au véhicule prédécesseur, comme illustré dans la figure 2.8.

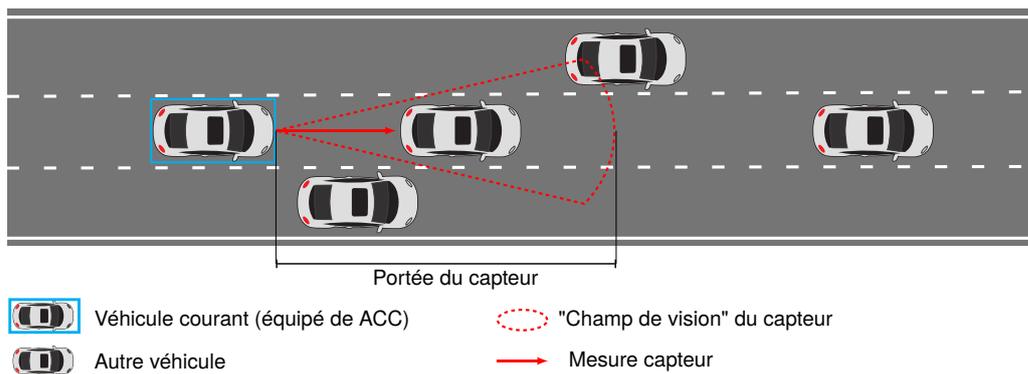


Fig. 2.8: Représentation de la perception d'un véhicule équipé du système ACC.

Comme il s'agit d'un système développé par différents constructeurs, le modèle permettant d'ajuster dynamiquement l'accélération du véhicule en fonction de ces données n'est souvent pas accessible. Les chercheurs en modélisation du trafic se sont tout de même attelés à reproduire son fonctionnement. Kesting et al [108] ont proposé un modèle basé sur IDM (éq. 2.4). IDM est conçu pour reproduire le comportement d'un conducteur ; le modèle n'est donc pas adapté aux systèmes automatiques qui ont des temps de réaction plus faibles et ne souffrent pas d'erreurs d'appréciation (capteur) ou de phénomène de relaxation (observé après les dépassements notamment). Les auteurs ont donc proposé un modèle visant à maintenir une accélération plus constante, autorisée par la réactivité du système. En plus des paramètres utilisés dans IDM, le modèle propose de moduler le comportement de ACC grâce à une valeur d'inter-distance temporelle cible et à un facteur de confort qui permet une conduite moins agressive. Les détails du modèle sont présentés dans l'ouvrage très complet de Treiber et Kesting [194]. Ce modèle a été implémenté en tant que processus de contrôle d'un système ACC.

Les résultats montrent des effets positifs et négatifs sur le flux [41]. Les études en simulation ou sur le terrain ont montré une augmentation possible de la capacité sur autoroute (grâce à une inter-distance réduite puisque le temps de réaction du système est très faible) mais une propension à générer ou propager de la congestion dès que le réseau est plus complexe (lorsque les changements de voie sont plus nombreux). D'autres expérimentations en conditions réelles [128] préconisent ironiquement de déconnecter le système ACC en conditions de trafic dense, car un peloton de véhicules équipés du système (fort taux de pénétration) serait plus instable qu'un peloton de véhicules classiques. Afin d'améliorer la stabilité du système, on a cherché à intégrer plus d'informations dans le contrôle du véhicule.

2.1.2.2 Modèles coopératifs

La coopération entre les véhicules se traduit par l'utilisation de données partagées par le flux dans le processus de contrôle du véhicule courant. L'objectif de cette forme de coopération est souvent égo-centré (il s'agit en premier lieu d'augmenter la sécurité du véhicule courant seul) mais les informations échangées sont profitables à l'ensemble du flux. Un premier pas vers la coopération est le modèle C-ACC, extension coopérative de ACC.

Cooperative Adaptive Cruise Control

Le principe reste très similaire à ACC, mais le véhicule courant ne limite pas sa perception à son seul capteur embarqué. Les informations supplémentaires proviennent du véhicule prédécesseur, connecté, qui échange ses paramètres intrinsèques (vitesse et accélération) grâce à la communication V2V. Les véhicules peuvent alors former des pelotons, dont l'existence est assurée par des fonctions de contrôle locales. De la même manière que pour ACC, beaucoup de modèles différents

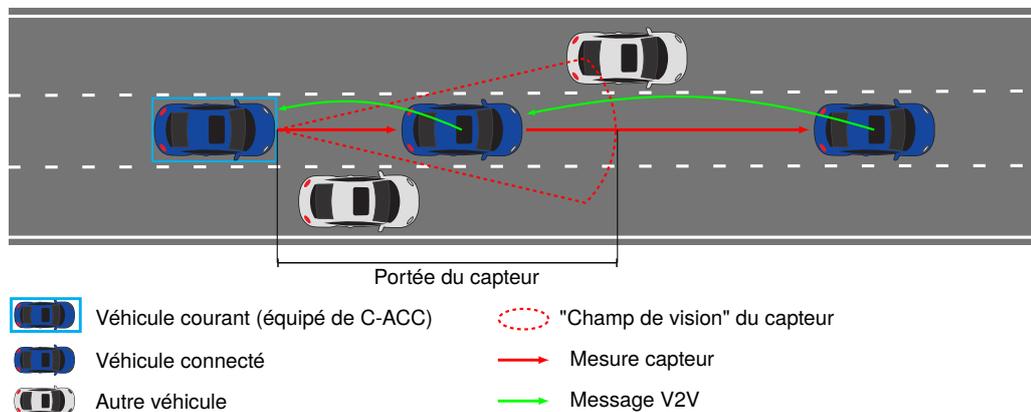


Fig. 2.9: Représentation de la perception d'un véhicule équipé du système C-ACC.

visent à proposer un contrôle d'un véhicule équipé du système C-ACC [42, 165]. L'avantage du système est de propager le comportement du premier véhicule dans le peloton dans un temps restreint puisque les informations transitent par la communication. Ce type de modèle permet d'améliorer significativement la stabilité des pelotons de véhicules (on parle de *string stability* [152]). Une nouvelle fois, l'inter-distance temporelle cible est choisie par le conducteur et influence directement le comportement du véhicule (progressivité de l'accélération).

Ces effets ont été reproduits en simulation [199] et vérifiés en expérimentation réelle [134]. Les perturbations, le plus souvent induites par les changements de voie [4], peuvent être réduites grâce au système. En contre-partie, l'écart entre le comportement du système et celui observé chez un conducteur entraîne des comportements qui vont à l'encontre de son bon fonctionnement : dépassements, occupation non justifiée d'une voie, etc. Le fait que le comportement du peloton soit régulé de proche en proche limite aussi les effets bénéfiques du système C-ACC à des faibles taux de pénétration (<50% [178]). Une généralisation du principe de ces modèles a donc vu le jour ; il s'agit de la multi-anticipation.

Multi-anticipation

Sans être évident pour un observateur externe, les modèles microscopiques de trafic sont en réalité relativement naïfs. En effet, seules les informations d'un véhicule prédécesseur (ou aucun) entrent en jeu dans le calcul de l'accélération du véhicule courant. Pourtant, votre expérience de conducteur vous a sans doute mis plusieurs fois dans cette situation plutôt inconfortable. C'est par exemple le cas lorsque vous suivez un véhicule de fort gabarit (poids-lourd, véhicule de livraison). Le comportement de conduite résulte en réalité d'une perception bien plus étendue : panneaux de signalisation, chaussée, véhicules environnants, éléments distrayeurs. Parmi les véhicules environnants, se trouve bien sûr le véhicule prédécesseur, mais les rétroviseurs permettent aussi de percevoir les suiveurs et la vision directe une partie des véhicules prédécesseurs. Les véhicules situés sur les côtés sont importants dans la prise de décision liée au changement de voie (section 2.1.1.4). L'exemple précédent nous confirme cette impression : en tant que conducteur, nous cherchons à anticiper le comportement des autres véhicules et non pas seulement le véhicule situé juste devant. C'est pourtant l'hypothèse suivie dans les systèmes ACC ou même C-ACC. Les observations sur la stabilité mitigée du flux sont donc logiques. La classe de modèles qui s'intéresse à étendre la perception du véhicule se base sur ce principe de la multi-anticipation [118]. Il est alors possible de considérer dans un modèle microscopique l'apport des véhicules situés devant [154] et même derrière [210] le véhicule courant.

Ce principe permet d'étendre significativement la perception d'un véhicule, comme illustré par la figure 2.10. En repartant de l'équation générale des lois de poursuite 2.2, il est possible d'étendre le

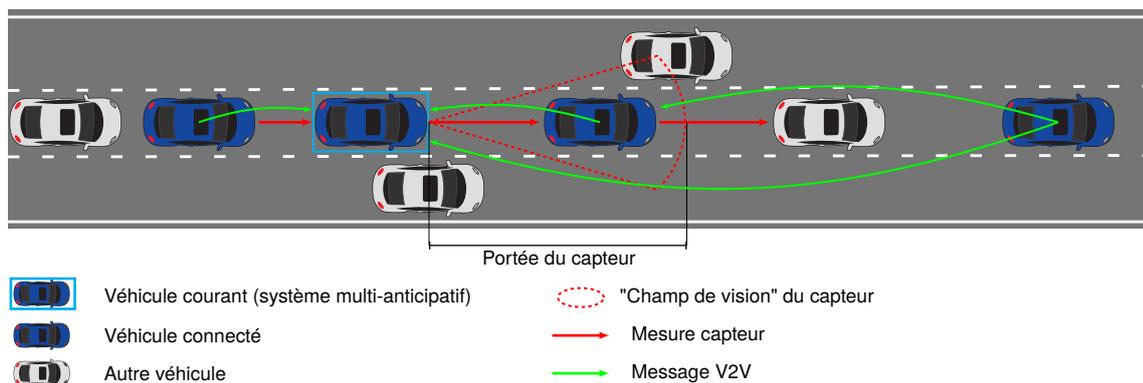


Fig. 2.10: Représentation de la perception d'un véhicule équipé d'un système de contrôle basé sur le principe de la multi-anticipation.

calcul de l'accélération du véhicule courant pour prendre en compte plusieurs véhicules (en amont

et en aval). La communication permet alors d'étendre les modèles microscopiques [71] afin de prendre en compte ces véhicules grâce à la multi-anticipation :

$$\ddot{x}_n = f \left(\dot{x}_n, \sum_j a_{n+j} \Delta x_{n+j}, \sum_j b_{n+j} \Delta \dot{x}_{n+j} \right) \quad (2.3)$$

Par exemple, une bonne base de modèle microscopique est ACC, qui devient MACC (*Multi-anticipative ACC*) dans sa version multi-anticipative [205, 206]. De tels modèles permettent de prendre en compte les enjeux environnementaux (effets sur la consommation grâce à des stratégies d'éco-conduite) et sur la sécurité (réduction de la congestion). Ces deux enjeux sont majeurs dans le déploiement des systèmes coopératifs.

La complexité des interactions nécessaires à la perception et à la prise de décision du véhicule nécessite cependant une architecture particulière dans la modélisation de ces systèmes. Comment intégrer la communication dans un modèle microscopique ? Comment modéliser les capteurs des véhicules ? Est-ce que le modèle va passer à l'échelle ? Les solutions apportées par les travaux cités plus haut imposent souvent une architecture centralisée qui gère les flux d'information. Cette solution est problématique pour le déploiement de ces modèles dans des véhicules autonomes. La littérature portant sur la modélisation de trafic et plus particulièrement les modèles multi-anticipatifs nous oriente directement vers l'utilisation des systèmes multi-agents (SMA). Opposés aux modèles calculatoires, ils permettent de modéliser un flux de véhicules en définissant le comportement de chaque individu [51] comme un agent autonome et ouvrent la voie vers les bénéfices apportés par ce type de paradigme.

2.1.3 Les Systèmes Multi-Agents – SMA

Les systèmes multi-agents (SMA) font référence aux approches distribuées fondées sur la définition de plusieurs entités autonomes et intelligentes appelées "agents". Initié par les recherches en robotiques [197], ce concept provient du monde de l'Intelligence Artificielle Distribuée (IAD) et regroupe une grande quantité de concepts s'articulant autour du paradigme agent [63]. La notion d'agent se réfère à une entité autonome évoluant en interaction avec son environnement, souvent dynamique et imprévisible. La modélisation de ces agents et leurs interactions trouve son inspiration dans l'observation de systèmes biologiques complexes (sociétés animales organisés comme les colonies de fourmis [46], les nuées d'oiseaux [172] ou les bancs de poissons). L'utilisation des SMA ne se limite pourtant pas à la simulation de vie artificielle [171], ils sont aussi un outil très performant dans la résolution de problèmes complexes [97]. Les systèmes multi-agents sont composés d'un ensemble d'agents autonomes en interaction les uns avec les autres, par l'intermédiaire de l'environnement qu'ils partagent. L'intérêt porté aux SMA réside dans leur capacité à générer des comportements émergents (à travers les phénomènes comme l'auto-organisation) à partir des interactions individuelles.

2.1.3.1 Les agents

Une des premières définitions est celle de Russel *et al.* [174] : un agent est défini comme toute entité capable de percevoir son environnement par l'intermédiaire de capteurs, et d'agir sur cet

environnement de manière autonome grâce à ses effecteurs. Cette définition a ensuite été successivement complétée. D'autres travaux [213] ajoutent une dimension sociale et pro-active. L'autonomie de l'agent signifie qu'il évolue sans intervention extérieure et qu'il contrôle ses états internes et ses actions. La notion de langage régit les interactions inter-agents. Ils sont aussi réactifs aux changements de l'environnement et pro-actifs dans la satisfaction de leurs objectifs.

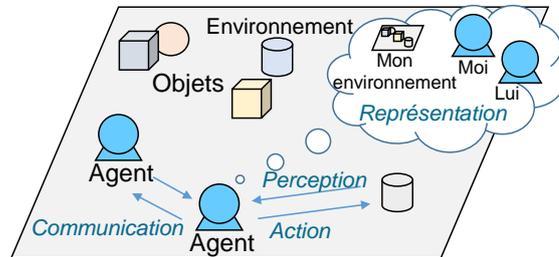


Fig. 2.11: Schéma simplifié des interactions d'un agent dans son environnement et conceptualisation de la notion de représentation

Enfin, les travaux de Ferber [63, 64], qui font référence dans ce domaine, convergent vers une définition plus étendue, qui peut être illustrée par la figure 2.11 ; un agent est une entité physique ou virtuelle :

- qui est capable d'agir dans un environnement,
- qui peut communiquer directement avec d'autres agents,
- qui est animée par un ensemble de tendances (objectifs, satisfaction, etc.),
- qui possède ses propres ressources,
- qui est dotée d'une capacité de perception de son environnement,
- qui n'a qu'une connaissance partielle et locale de son environnement,
- qui possède des compétences et peut offrir des services,
- qui peut éventuellement se reproduire,
- dont le comportement tend à satisfaire ses objectifs, en prenant en compte les ressources et compétences dont elle dispose, en fonction de sa perception, de ses représentations et des communications qu'elle reçoit.

Cette définition permet d'énoncer le concept de système multi-agent, à savoir : un système dans lequel plusieurs agents autonomes évoluent. Les notions de but individuel, de communication et de compétence font entrevoir des possibilités de coopération entre agents. L'étude des systèmes multi-agents permet d'en comprendre les fondements.

2.1.3.2 De l'agent aux systèmes multi-agents

D'après la définition de Demazeau [45], un Système Multi-Agent (SMA) est caractérisé par :

- Un environnement E : un espace disposant d'une métrique et d'un support de communication.

- Un ensemble d'objets O : ces objets sont situés dans l'environnement E . Il est donc possible à tout moment, d'associer à O une coordonnée dans E .
- Un ensemble d'agents A : les entités actives et autonomes du système.
- Un ensemble d'interactions I : elles permettent aux agents A de modifier l'environnement E
- Un ensemble d'opérateurs O_p : ils sont capables d'opérer sur l'environnement et de provoquer des réactions.

Cette définition (qui fait écho à l'illustration de la figure 2.11), associée à celle d'un agent [64], met en évidence l'existence de buts individuels et/ou collectifs au sein du système multi-agent. Les agents peuvent donc exprimer des propriétés de collaboration, ou coopération qui nous intéressent dans le cadre de notre travail. Ces propriétés sont souvent dépendantes de l'environnement des agents, support des interactions.

L'environnement

L'environnement est au centre des interactions des agents. À des fins de modélisation, il est possible de décomposer l'environnement en sous-environnements [150] : on distingue alors l'environnement physique, qui situe les agents, de l'environnement social qui porte les caractéristiques des relations entre les agents. Suivant le choix porté lors de la modélisation, l'environnement constitue alors une première étape d'abstraction [209] qui réunit les conditions nécessaires à l'existence des agents.

Il peut ainsi être limité à une stricte interface, support des interactions. C'est le cas du modèle d'environnement sous la forme d'un "tableau noir" [35] qui permet à une population d'agents de partager des informations en modifiant cet "environnement". L'idée était de fragmenter le problème en sous-problèmes plus simples de manière récursive. Chacun des agents pouvant ainsi résoudre une partie du problème correspondant à ses capacités.

À l'inverse, un environnement qui supporte des interactions plus complexes exhibera davantage de propriétés. Parmi ces propriétés, nous pouvons citer certaines parmi celles proposées par Russel *et al.* [174] :

- Accessible ou inaccessible : les agents ont accès à tout ou partie de l'état de l'environnement ;
- Déterministe ou non déterministe : les transitions entre les états de l'environnement sont uniquement dépendantes de l'état actuel ou dépendent aussi des interactions ;
- Statique ou dynamique : au sens de l'agent, c'est-à-dire que l'environnement n'évolue pas pendant sa prise de décision ;
- Discret ou continu : on peut distinguer ou non un nombre limité et connu *a priori* d'états de perception.

Les problèmes les plus complexes sont très souvent modélisés avec des environnements continus, non-prévisibles, dynamiques, et où les agents n'ont accès qu'à une partie de l'état de l'environnement. Pour ce type de problèmes, les agents doivent alors construire et maintenir une représentation locale, ce qui garantit une robustesse de leurs actions au cours du temps, puisque l'environnement peut évoluer durant leur phase de décision. Le cas de la modélisation et du contrôle du trafic coopératif nous place dans ce type de problématique. L'environnement, qui est celui de navigation des véhicules, est par nature complexe et dynamique. Pour pallier cette contrainte, il convient d'étudier en détail quelles interactions et dynamiques doivent être mises en place pour contrôler le comportement global du système.

Les interactions

La notion d'interaction est centrale dans les systèmes multi-agents, où l'autonomie des entités passe par leurs échanges avec l'environnement et où les échanges entre agents peuvent mener à des comportements collectifs [45]. Les interactions sont ainsi des actions réciproques qui peuvent modifier le comportement ou la nature des entités du système. Généralement, on distingue les interactions directes (agent-agent) des interactions indirectes (agent-environnement).

Les interactions indirectes sont les formes de communication qui s'appuient le plus souvent sur des modifications de l'environnement partagé par les agents. Ces modifications peuvent porter sur un environnement physique ou un environnement virtuel. Un environnement physique peut porter les traces des actions des agents. Les phéromones, qui tirent leur inspiration de la biologie, sont la forme la plus connue de ces traces. Dans un environnement virtuel, défini par le concepteur, les agents peuvent échanger des informations dans un espace dédié nommé "espace de marquage" (par exemple, le tableau noir [35]). La création d'un espace de marquage permet de proposer une abstraction des interactions des agents, ce qui peut être très utile pour catalyser les échanges et donc faciliter l'apparition de phénomènes émergents. On parle alors de stigmergie [19, 22] quand ces interactions mènent à l'apparition de forme d'organisation. Certains enjeux des systèmes coopératifs peuvent être vus comme des buts à atteindre pour le système. L'utilisation d'un espace de marquage pourrait être une solution pour les satisfaire.

Les interactions directes regroupent les formes de communication qui nécessitent la mise en place d'un protocole, qui peut prendre la forme d'un langage. C'est ce type d'interaction que l'on retrouve dans les systèmes coopératifs où des véhicules connectés échangent des informations sous la forme de messages V2V. Les véhicules peuvent alors être modélisés comme des agents autonomes

L'usage des interactions directes et indirectes n'est pas exclusif. Les deux formes semblent en effet présenter des avantages utiles à la modélisation et au contrôle des véhicules connectés grâce aux concepts multi-agents. Nous cherchons avant tout à produire des formes d'organisation permettant un meilleur écoulement à l'échelle du flux.

2.1.3.3 Émergence, organisation, auto-organisation

Les agents sont des entités autonomes évoluant dans un système (le SMA) lui-aussi autonome. Ces deux niveaux de lecture laissent entrevoir la possibilité de co-existence de comportements à l'échelle du système émanant des interactions des agents. Ce type de phénomène se rapporte à la notion d'émergence, qui peut s'accompagner de l'apparition d'une forme d'organisation, prémice de l'auto-organisation.

Organisation et auto-organisation

L'autonomie d'un agent se traduit par sa capacité à poursuivre ses propres buts [28]. Avec les SMA, la dynamique de l'environnement est alors régie par la succession de modifications, fruits de l'interaction avec les agents. Il paraît difficile dans ce contexte de contrôler le comportement du système.

C'est ici qu'entre en jeu la notion d'organisation. Pour faire simple, une organisation spécifique qui fait quoi et comment [85]. En réalité, il y a beaucoup de façons de décrire une organisation dans les SMA. D'ailleurs, le terme "organisation" désigne à la fois la façon d'organiser le système et le résultat de cette action. Il y a donc l'implication d'un élément extérieur au système capable de contrôler le processus d'organisation.

C'est principalement ce qui oppose organisation et auto-organisation. L'auto-organisation est un processus permettant au système, sans contrôle extérieur explicite, de modifier son organisation au cours de son exécution, afin de s'adapter aux changements survenant dans son environnement [50]. On distingue généralement deux approches permettant de concevoir des systèmes auto-organisés : l'approche bio-inspirée [127] et l'approche sociale [86]. Dans les deux cas, un système auto-organisé peut être caractérisé par les propriétés suivantes [49] :

- **Absence d'un contrôle externe explicite.** Cette propriété garantit l'autonomie du système. L'organisation dépend seulement de processus internes (qui peuvent être influencés par l'environnement).
- **Contrôle décentralisé.** Cette autonomie n'empêche pas une forme de contrôle décentralisé.
- **Exploitation dynamique.** L'auto-organisation évolue au cours du temps. Comme cette évolution est indépendante de toute forme externe de contrôle, cela implique une continuité dans le processus d'auto-organisation.

Cette liste de propriétés peut être complétée par d'autres caractéristiques [49] telles que : adaptation, redondance, hiérarchies ou encore émergence.

La notion d'émergence

Les systèmes multi-agents sont fondés sur la définition de comportements individuels simples, où les interactions (entre agents, avec l'environnement) prennent une place cruciale : elles fondent la structure d'une organisation. L'intérêt des SMA est ainsi de reproduire un système au comportement global complexe dont le fonctionnement semble régi par des règles globales qui n'ont pas été définies au préalable. On parle dans ce cas d'*émergence* [48] ou d'*auto-organisation* [70].

Dans le cas du trafic, on peut observer ce genre de phénomène lors de la formation d'une congestion. On peut considérer le flux comme une collection d'"agents" véhicules indépendants, mais le phénomène de propagation d'une onde de congestion est observable directement à l'échelle macroscopique [182].

Les notions d'émergence et d'auto-organisation paraissent liées, mais sont en fait deux notions à différencier [212]. Les SMA sont très souvent à la fois émergents et auto-organisés puisque l'auto-organisation peut être vue comme un phénomène émergent. Lorsque l'organisation s'établit dans le but de satisfaire un but commun, on parle alors de *coopération*.

2.1.4 Les SMA appliqués aux transports

La transposition des notions orientées agent au domaine des transports est en accord avec les caractéristiques propres aux deux domaines. On y retrouve en effet des caractéristiques d'autonomie, de comportements distribués et d'environnement partiellement observable. Les approches existantes

s'intéressent aux propriétés des systèmes multi-agents : émergence, auto-organisation, coopération. De plus, l'évolution d'un grand nombre de véhicules sur un réseau routier partagé correspond parfaitement aux problématiques de conflit de ressources étudiées via les SMA [197]. Il est donc logique que l'on retrouve des utilisations des concepts tirés du paradigme agent dans le domaine du trafic et des transports [31], à travers des applications très variées [24, 40].

2.1.4.1 Modélisation SMA

Alors que la modélisation à des fins de simulation semble être l'application la plus évidente des SMA dans les transports, on ne retrouve que peu de travaux dans ce domaine. Pour la simulation, l'approche comportementale est préférée [52, 53] et permet de reproduire les comportements en transposant les modèles de suivi sous la forme de comportements d'agents. L'intérêt de l'utilisation des SMA reste ici limité : les interactions, éléments clés de tels systèmes, ne sont pas assez développées pour profiter des bénéfices d'une modélisation agent. Il est donc logique que les applications des SMA au niveau du trafic se placent à un plus haut niveau. Beaucoup de modèles à base d'agent se consacrent donc à traiter des problématiques bien identifiées dans les transports [16] : gestion, régulation, contrôle.

2.1.4.2 Gestion du trafic

La gestion du trafic vise essentiellement à faire converger l'offre du transport à la demande avérée ou estimée. Pour ce faire, elle propose une panoplie de stratégies pour réduire les externalités négatives comme la congestion, les accidents, l'augmentation des temps de parcours ou des émissions causés par la croissance de la demande. Il n'existe que deux solutions à ces problèmes : augmenter la capacité actuelle du réseau ou utiliser de nouvelles technologies pour optimiser l'usage de l'infrastructure routière existante [65].

Les systèmes de transports intelligents offrent un ensemble d'outils, s'appuyant sur les avancées en termes de puissance de calculs, de communication et de perception, afin de produire un système supervisé, intégré, universel et abordable. Les systèmes multi-agents semblent constituer la meilleure voie d'amélioration à la fois en termes d'interaction et de calcul distribué.

Gestion des intersections

Pour agir en faveur de la réduction des temps de trajet, beaucoup de chercheurs se sont intéressés à l'amélioration de l'efficacité des systèmes d'intersections situés aux nœuds des réseaux routiers. De par leur nature conflictuelle, les intersections sont en effet une cause bien connue de congestion en milieu urbain. Les feux de signalisation (feux tricolores) sont adaptés à l'introduction d'un contrôle dynamique [33, 113], où les cycles des feux sont modulés en temps réel à la demande provenant de chaque arc. Les agents sont dans ce cas définis au niveau de l'intersection et/ou pour chaque feu, et leur coopération permet d'optimiser les cycles face à la demande.

Régulation du trafic

La régulation du trafic s'opère par l'action sur les itinéraires empruntés par les véhicules (restrictions d'accès, suggestion d'itinéraire alternatif, etc.). Ce contrôle implique un aspect prédictif de l'état

de trafic afin d'anticiper les situations menant à une saturation du réseau (on parle de contrôle prédictif). Dans cette optique, l'introduction du concept agent a été expérimentée, avec par exemple dans Oliviera *et al.* [43] une formulation basée sur la commande prédictive (MPC : *Model Predictive Control*). D'autres travaux [2, 3] introduisent l'aspect coopératif des SMA pour prendre en compte l'acceptation des consignes du gestionnaire de trafic par les usagers du réseau.

Gestion de la congestion

L'une des caractéristiques les plus étudiées en gestion du trafic est le phénomène de congestion, et plus particulièrement les transitions entre régime fluide et régime congestionné. Beaucoup de travaux ont été menés afin d'agir sur ce phénomène [47], principalement afin d'éliminer ou tout au moins retarder son apparition. Les systèmes multi-agents trouvent différentes utilisations dans ce domaine. On retrouve par exemple des applications destinées à assister le gestionnaire pour une exploitation supervisée [123, 200, 32] via un mécanisme de détection/décision SMA.

2.2 Réputation, confiance et coopération

La coopération repose sur le concept de délégation, qui est une notion critique dans les systèmes multi-agents [29]. Le cas d'étude des systèmes coopératifs est un parfait exemple de système qui repose sur la délégation car l'action d'un véhicule est directement dépendante des actions des autres véhicules. Dans cette partie, nous anticipons les problématiques qui pourraient survenir lorsque les véhicules communicants feront reposer leur processus de décision sur des informations tierces. Nous proposons dans la suite de ce travail une modélisation des véhicules sous la forme d'agents autonomes. Les C-ITS peuvent alors être modélisés comme un système multi-agent et l'objectif est de trouver un modèle existant capable d'être intégré à la décision des agents, et qui permettrait de limiter les effets de mauvaise information ou de capteurs non-fiables dans les systèmes coopératifs.

2.2.1 Définitions

Les concepts de confiance et de réputation sont très utilisés dans les problématiques de prise de décision. Ces notions, qui trouvent une origine sociologique, ont aussi récemment prouvé leur intérêt pour le maintien en cohérence d'un réseau communicant.

- La **réputation** est l'opinion qu'une entité se fait d'une autre. C'est une évaluation à un moment donné de la fiabilité d'une autre entité en adoptant le point de vue individuel d'une entité. La réputation peut être partagée, individuelle, ou résulter de ces deux composantes.
- La **confiance** est le degré de croyance, espéré ou avéré, associé à une entité relative à la réalisation d'actions par une autre. En d'autres termes, la confiance caractérise le degré de croyance en l'exécution d'une action par une autre entité.

2.2.2 Coopération dans les SMA

La notion de coopération peut se traduire par plusieurs actions de la part d'un groupe d'individus. On peut parler de coopération lorsque la satisfaction d'un but commun implique l'action coordonnée de plusieurs éléments, ou même simplement lorsque les individus participent à la résolution d'un conflit. La nature conflictuelle des ressources partagées par les agents dans un environnement SMA correspond directement à une situation de coopération. L'aspect coopératif d'un système est observable à une échelle plus globale, sous la forme d'un comportement émergent.

Ferber [63] donne la définition suivante de la coopération dans les SMA. On dira que plusieurs agents coopèrent, ou encore qu'il sont dans une situation de coopération, si l'une des deux conditions est vérifiée :

1. L'ajout d'un nouvel agent permet d'accroître différenciellement les performances du groupe.
2. L'action des agents sert à éviter ou à résoudre des conflits potentiels ou réels.

Le pré-requis à toute forme de coopération est de disposer d'une forme de communication entre les individus afin qu'ils soient en mesure d'intégrer leurs voisins dans leur mécanisme de décision. En prenant l'exemple d'un réseau d'entités individuelles communicantes, il est possible de se retrouver dans plusieurs situations. Soit chacune des entités se contente d'utiliser toutes les ressources à sa disposition, et fait ainsi preuve d'un comportement égoïste ; soit les entités trouvent un intérêt à communiquer et partagent les ressources disponibles [151]. Bien sûr, il apparaît clairement que ne pas coopérer avec les autres entités du réseau peut être plus avantageux pour une entité. C'est pourquoi il est nécessaire d'introduire un caractère incitatif (sous forme de motivation) pour initier une forme de coopération. Cette motivation peut être apportée par les mécanismes de confiance et de réputation. Un agent peut en effet s'appuyer sur la réputation d'un autre agent pour déterminer s'il trouvera un intérêt à coopérer avec lui. À l'inverse, un agent est amené à coopérer pour maintenir sa réputation et donc la confiance que les autres lui accordent.

2.2.3 Systèmes fondés sur la confiance et la réputation

La problématique de la confiance et de la réputation est très présente dans les systèmes distribués (IAD). Ce type de modèle est donc très représenté dans la littérature. L'objectif de cette section est de comparer les modèles existants afin de trouver une solution applicable aux systèmes coopératifs. Les modèles dédiés aux SMA sont donc privilégiés, et notre comparaison se fonde sur un ensemble de critères de classement.

2.2.3.1 Critères de classement

Afin de décrire et de classer la grande variété des modèles basés sur des mécanismes de réputation/confiance, on peut définir différents critères de classification.

Nature du paradigme

D'après Sabater et Sierra [177], les modèles se classent suivant deux types de paradigme : cognitif ou numérique. Les modèles cognitifs (C) décrivent des degrés de croyance, et ont une représentation

explicite. Les modèles numériques (**N**) font appel à des valeurs plus abstraites pour représenter la notion de confiance et de réputation.

Sources d'information

Pinyol et Sabater-Mir [163] proposent de classer les modèles en fonction des sources d'information utilisées pour établir les valeurs de confiance ou de réputation :

- Informations Directes : on distingue les informations issues d'interactions directes (**DI** : *Direct Interactions*) ou d'observations directes (**DO** : *Direct Observations*).
- Informations Indirectes (**I**) : par opposition, on retrouve les informations indirectes, issues du témoignage d'autres agents.

Visibilité

La portée d'une valeur de confiance peut être partagée par l'ensemble d'un groupe (propriété globale **G**) ou se limiter à la seule connaissance d'un agent (propriété subjective **S**).

Dans le premier cas, le calcul de la valeur de confiance/réputation est réalisé à partir des opinions des agents qui ont précédemment interagi avec les individus évalués. La valeur est rendue publique pour tous les membres de la communauté, puis actualisée à chaque fois qu'une nouvelle mise à jour est effectuée par l'un des membres. Dans l'autre cas, chaque agent met à jour une valeur de confiance/réputation pour chaque autre membre à partir de leurs interactions propres : expériences directes, témoignages recueillis, relations connues entre membres, etc.

Contexte

La granularité d'un modèle basé sur la réputation et la confiance se réfère à la dépendance au contexte (**CD** : *Context Dependant* ou **CI** : *Context Independant*) de l'information de confiance/réputation. Le contexte est en général fixé par l'environnement, on peut donc voir les systèmes simple contexte comme un cas particulier des systèmes multi-contextes.

Tricherie et Mensonge

Sabater et Sierra [177] proposent de décomposer en 3 niveaux les modèles adoptés :

- Niveau **0** : le comportement de tricherie n'est pas considéré. Le modèle compte sur un grand nombre d'agents honnêtes pour contrer l'effet potentiel d'agents malveillants.
- Niveau **1** : le modèle prend en compte la possibilité que certains agents puissent cacher des informations, mais ils ne mentent pas.
- Niveau **2** : le modèle dispose de mécanismes spécifiques qui prennent en compte les agents qui mentent.

Nature des informations

On distingue les modèles qui manipulent des données discrètes voire booléennes (**B**) et ceux qui utilisent des valeurs continues (**CO**). Les premiers sont généralement basés sur des probabilités alors que les autres fonctionnent le plus souvent grâce à un mécanisme d'agrégation.

Type de modèle

Les modèles peuvent utiliser des mécanismes de confiance (**T** : *trust*) (calculée à partir d'une valeur de fiabilité), et/ou de réputation (**R**). Cette classification est revendiquée par les auteurs des différentes approches.

2.2.3.2 Comparaison des modèles

À partir des critères de classement établis dans la section précédente, on peut présenter les approches existantes sous la forme du tableau 2.1. Le grand nombre de modèles existants témoigne à la fois de l'intérêt des problématiques de confiance mais aussi des divergences en termes de représentations adoptées. Une petite proportion seulement prend en compte la possibilité d'occultation d'information ou de diffusion d'informations erronées. La sécurité de systèmes tels que les SMA, où l'échange d'informations via la communication structure les interactions, est justement un point clé de leur robustesse. La complexité des algorithmes utilisés est un point à ne pas négliger pour assurer une bonne coopération quelle que soit l'envergure de la simulation.

Pour les systèmes coopératifs, la quantité d'informations échangées doit être limitée. Cela est dû aux contraintes techniques liées aux technologies déployées. L'emploi de modèles cognitifs semble alors difficile. L'aspect local est une des caractéristiques principales des interactions des véhicules. Les représentations devront donc se limiter à des aspects subjectifs. La manipulation de données numériques continues semble plus adaptée à une évaluation dynamique de la confiance entre les véhicules communicants. Enfin, comme notre modélisation vise avant tout une phase précoce de déploiement des systèmes coopératifs, l'utilisation d'informations directes (interactions) et indirectes apparaît comme une bonne solution pour améliorer la robustesse du système. La sécurité (au sens du flux et du système) est primordiale : des mécanismes contre la tricherie seront sans doute nécessaires. À partir de ces observations, deux modèles se distinguent particulièrement. Il s'agit des réseaux de confiance [149] et de FIRE [94]. Nous proposons dans la suite de notre travail d'utiliser ces modèles afin de contrôler la fiabilité de l'information dans les communications inter-véhiculaires (section 2.3). Une comparaison des deux modèles permettra de montrer leurs avantages respectifs dans les expérimentations (chapitre 3) menées en simulation.

Approche	Modèle		Représ.		Informations			Visibilité		Contexte		Triche			Données	
	Confiance	Réputation	Numérique	Cognitif	Interactions	Observations	Indirectes	Subjectif	Global	Dépendant	Indépendant	Niveau 0	Niveau 1	Niveau 2	Booléennes	Continues
Modèle	T	R	N	C	DI	DO	I	S	G	CD	CI	0	1	2	B	CO
AFRAS [25]		R	N		DI		I	S			CI			2		CO
Carter et al.[27]		R	N				I		G		CI	0				CO
eBay		R	N				I		G		CI	0				CO
Dirichlet [99]		R	N				I		G		CI	0				CO
Esfandiari et al. [58]	T		N		DI	DO	I	S		CD		0				CO
FIRE [94]	T	R	N		DI		I	S		CD		0				CO
ForTrust [91]	T	R		C	-	-	-	S			CI	-	-	-		CO
LIAR [143]	T	R	N		DI	DO		S			CI			2		CO
Marsh [129]	T		N		DI			S		CD		-	-	-		CO
Mui [142]		R	N		DI		I	S		CD		0				CO
Padovan et al. [155]		R	N		DI		I	S			CI	0				CO
Rahman et Hailes [1]	T	R	N		DI		I	S		CD				2	B	
Rasmusson et al. [169]		R	N				I		G	CD				2		CO
Regan et Robin [170]	T		N		DI		I	S			CI			2		CO
Regret [176]	T	R	N		DI	DO	I	S		CD				2		CO
Repage [175]		R	N	C	DI		I	S		CD				2		CO
BDI + Repage [162, 164]	T	R	N	C	DI		I	S		CD				2		CO
Schillo et al. [179]	T		N		DI	DO	I	S			CI		1		B	
Sen et Sajja [181]		R	N		DI	DO	I	S			CI			2	B	
Sierra et Debenham [183]	T		N		DI		I	S			CI	0				CO
Social Trust [29]	T			C	-	-	-	S		CD		-	-	-		CO
HISTOS [217]		R	N		DI		I	S			CI	0				CO
SPORAS [218]		R	N				I		G		CI	0				CO
TrustSets [149]	T		N		DI		I	S			CI			2		CO
Yu et Singh [216]	T	R	N		DI		I	S			CI	0				CO

Tab. 2.1: Comparaison des modèles basés sur la confiance et la réputation parmi les plus utilisés de la littérature.

2.3 Modélisation multi-agent des C-ITS par couplage de dynamiques

La description des interactions dans les systèmes coopératifs a montré le besoin, dans la modélisation, d'adopter un point de vue microscopique. De plus, l'ajout de la communication apporte une nouvelle dynamique à modéliser dans le système. Le problème est que toutes les interactions (ou flux d'informations) sont liées et inter-dépendantes. Dans ce cas, il est difficile de modéliser chacune des briques qui constituent le système complètement en parallèle. Nous proposons de modéliser les flux d'information au sein du système des C-ITS sous la forme d'un couplage de dynamiques. Ce couplage est rendu possible par l'apport des systèmes multi-agents, qui permettent de modéliser les différents niveaux d'interaction, en proposant une définition individuelle des comportements des entités du système.

Le modèle que nous proposons est une extension des travaux de Monteil *et al.* [136]. Dans cette première approche, les auteurs ont proposé une première version du couplage dynamique de trois flux d'informations pour les C-ITS. Bien qu'inspirée par les travaux en simulation orientée agent, la modélisation était limitée aux communications V2V et la dynamique de contrôle était simplifiée. Cette base repose néanmoins sur l'idée intéressante suivante : il est possible de décomposer les éléments du système en couches inter-dépendantes et dont le couplage régit les échanges entre ces modèles grâce aux différentes dynamiques du système. Sur cette base, nous avons proposé [82] une modélisation des C-ITS sous la forme d'un couplage de trois couches inter-dépendantes, représenté par la figure 2.12.

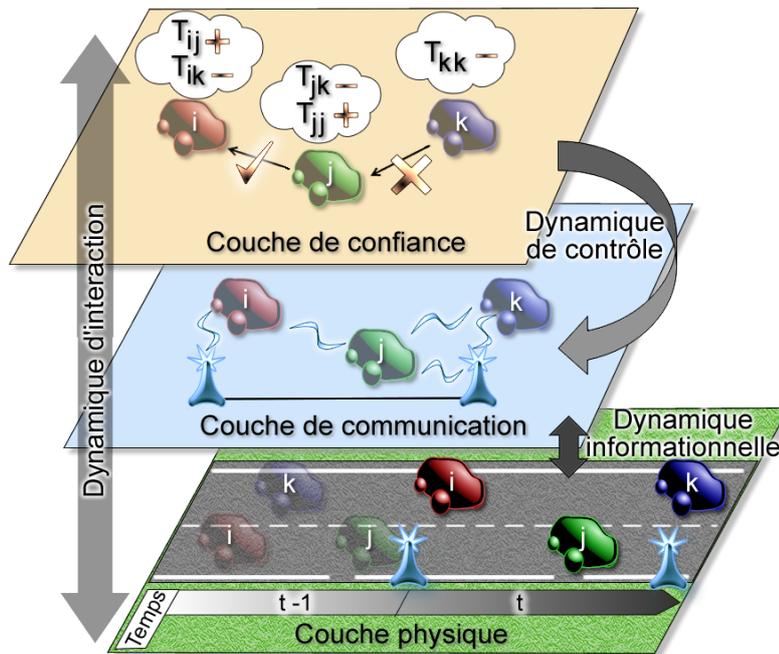


Fig. 2.12: Couplage de dynamiques pour les C-ITS

Dans ce modèle, trois couches permettent le couplage des différentes dynamiques du système afin d'intégrer une dimension de fiabilité de l'information (confiance), tout en propageant les informations (communication) qui agissent sur le comportement (physique) des véhicules. La couche physique régit le mouvement des véhicules grâce à des modèles microscopiques (lois de poursuite). La couche de communication gouverne les échanges d'informations grâce à des règles de proximité et de fiabilité, qu'il est possible d'exprimer en terme de probabilité. Enfin, la couche de confiance modélise la fiabilité de l'information et permet à un agent d'évaluer la confiance qu'il accorde aux autres. Nous parlons ici de couplage puisque les couches de communication et de confiance influencent les paramètres physiques, et donc le comportement des véhicules qui prennent en compte les échanges d'informations.

En partant de l'état de l'art des modèles de trafic, présenté dans la section 2.1.1, nous avons orienté notre modélisation vers le point de vue microscopique. Les lois de poursuite existantes permettent de modéliser le comportement longitudinal des véhicules classiques et peuvent être utilisés comme base de modèles coopératifs. L'intérêt des lois de poursuite est aussi la possibilité de les utiliser comme un élément de la prise de décision d'un agent autonome. La couche physique modélise aussi le réseau routier ainsi que l'infrastructure (les Unités de Bord de Route). Toutes les entités

(véhicules et UBR) peuvent être modélisées comme des agents autonomes, qui perçoivent leur environnement par l'intermédiaire de leurs capteurs. Les échanges d'information, sont gérés par la couche de communication. Dans notre modélisation, plusieurs flux sont inter-connectés et agissent sur les couches haute (confiance) et basse (physique) à travers un protocole de communication dédié. En repartant des standards définis pour les C-ITS (section 1.1.2), l'objectif est de proposer un cadre générique permettant une définition non limitée de tous types de messages (informations de l'infrastructure, ou échanges entre les véhicules). La couche de confiance agit enfin comme une dynamique de contrôle du système. Elle permet de représenter la fiabilité des informations échangées sous la forme d'une valeur numérique associée au concept de confiance. Le couplage des dynamiques du système permet ainsi de propager tous les flux d'informations et d'agir sur le contrôle final (*i.e.* le comportement physique) des véhicules. Comme une illustration du couplage présenté ici, nous proposons un modèle coopératif permettant à des véhicules équipés et connectés d'homogénéiser un flux de trafic autoroutier.

2.4 Application du couplage : un modèle multi-anticipatif bilatéral

Grâce au couplage proposé dans notre travail, il est possible d'étendre les comportements individuels des véhicules pour adapter dynamiquement leur processus de décision. Cette adaptation résulte de l'intégration des informations échangées à travers la communication (V2V et/ou V2I) et prend en compte la confiance accordée aux informations traitées. Dans cette partie, nous utilisons l'architecture en trois couches décrite précédemment pour proposer un modèle de contrôle local des véhicules coopératifs. Dans cette approche, nous considérons des véhicules semi-automatisés, capables de percevoir, par l'intermédiaire de leurs capteurs dédiés, des informations relatives aux véhicules de leur voisinage immédiat. Ces informations sont complétées par des données partagées via la communication V2V. Grâce au couplage avec la dynamique de contrôle, un modèle basé sur la confiance permet aux véhicules coopératifs d'intégrer la dimension de la fiabilité des capteurs dans leur prise de décision.

2.4.1 Objectifs

Les enjeux principaux des C-ITS (section 1.1.4) peuvent s'appréhender au travers de trois dimensions : sécurité, confort et environnement. Ces trois enjeux majeurs, qui peuvent apparaître antagonistes, peuvent cependant être réunis derrière le même objectif lorsqu'il s'agit de gérer un flux de véhicules : l'homogénéisation. Derrière ce terme, se cache un principe simple, parfois résumé par l'effet "*slower is faster*" [74]. Le comportement d'un flux de trafic répond très bien à cet effet, qui survient quand un système complexe exhibe de moins bonnes performances lorsque les individus tentent d'améliorer leurs propres objectifs (et tendre vers l'équilibre individuel). Ce type de phénomène est bien connu en trafic ; on parle alors de chute de capacité [185], qui peut survenir lorsque la topologie du réseau (le plus souvent des nœuds convergents) provoque un changement brusque de capacité. La relation entre les changements de voies et la vitesse des véhicules provoque une chute qui réduit la capacité à une valeur plus faible que celle de la section en aval.

Les C-ITS permettent aux véhicules d'étendre leur champ de perception grâce aux nouvelles informations échangées. Grâce à un processus de décision local, il est possible d'imaginer des solutions permettant d'aller dans le sens d'une amélioration de l'écoulement du flux. L'homogénéisation des vitesses [187], ou des comportements des véhicules en général, est une forme de régulation qui permet d'atteindre l'objectif d'homogénéisation du trafic. Grâce aux systèmes coopératifs, et en utilisant les propriétés des systèmes multi-agent, nous visons à homogénéiser les comportements au sein d'un flux de véhicules. Le défi principal réside dans le déploiement des systèmes coopératifs qui sera irrémédiablement progressif. Peut-on envisager que seule une partie du flux puisse provoquer une régulation à l'échelle d'une section routière entière ? D'autres questions émanent aussi de la faisabilité d'une telle approche : comment limiter les échanges d'informations au strict minimum ? Comment s'assurer que les informations échangées sont fiables ?

Nous proposons un modèle coopératif basé sur le principe de la multi-anticipation, qui considère la fiabilité des capteurs mis en jeu dans le processus de décision d'un véhicule semi-automatisé. Ce modèle, nommé loi bilatérale multi-anticipative (BMA) est une extension du modèle proposé par Monteil *et al.* [138]. Il peut être vu comme une généralisation de C-ACC (régulateur de vitesse à contrôle de distance coopératif), présenté en section 2.1.2.2. Le principe est d'atteindre une homogénéisation du flux grâce à la modification du comportement de certains véhicules, à travers une loi de poursuite coopérative, *i.e.* qui intègre des informations partagées par ses voisins.

2.4.2 Hypothèses

Le cadre général d'application du modèle est l'équipement progressif d'une partie d'un flux autoroutier par des capteurs de proximité et des dispositifs de communication. Le comportement longitudinal de ces véhicules est délégué à un système automatique qui se charge de réguler l'accélération du véhicule. L'originalité de la loi multi-anticipative bilatérale, qui peut être implémentée en tant que processus de décision d'un véhicule, est d'utiliser les informations provenant d'autres véhicules pour lisser son propre comportement. Dans ce cadre, nous considérons deux types de véhicules. La section étudiée est parcourue par des véhicules non-coopératifs, conduits par des conducteurs humains, et des véhicules coopératifs, qui exécutent le modèle BMA pour contrôler leur accélération. Les caractéristiques de ces deux types de véhicules peuvent être résumées par le tableau 2.2.

	Véhicule non coopératif	Véhicule coopératif
Capteurs	Conducteur humain	Télémètre longue portée (80 m) Dispositif de communication sans fil Récepteur GNSS
Informations échangées	Aucune	Mesures des capteurs Valeurs de confiance
Informations intrinsèques	Vitesse instantanée Limite de vitesse	Vitesse instantanée Limite de vitesse
Données capteurs	Distance avec le prédécesseur Vitesse relative avec le prédécesseur	Distances et Vitesses relatives avec les n prédécesseurs
Modèle physique	<i>Intelligent Driver Model</i> [192] Modèle de changement de voie [104]	Loi bilatérale multi-anticipative [138] Modèle de changement de voie [104]
Contrôle	Action du conducteur	Régulation semi-automatique de l'accélération/freinage

Tab. 2.2: Résumé des caractéristiques des véhicules coopératifs et non coopératifs.

Ce cadre d'étude seul ne suffit pas à générer un comportement d'homogénéisation, mais permet aux véhicules coopératifs d'accéder à suffisamment d'informations. Le modèle bilatéral multi-anticipatif peut se décomposer en trois étapes. La multi-anticipation s'exprime comme une loi de poursuite et

gère le comportement physique du véhicule. Une règle de proximité permet de pondérer l'influence des informations partagées par les autres véhicules communicants. Enfin, la confiance estimée par les agents intervient pour modéliser la fiabilité de l'information. Ce modèle est donc un parfait exemple qui permet d'illustrer le couplage de dynamiques présenté précédemment.

2.4.3 Interactions physiques

D'un point de vue physique, les véhicules non coopératifs sont modélisés par deux modèles, régissant leur comportement longitudinal (loi de poursuite) et latéral (modèle de changement de voie). Grâce à la revue des travaux existants dans le domaine de la modélisation, nous avons identifié des modèles permettant de reproduire fidèlement les comportements macroscopiques (notamment les instabilités) observés au niveau du flux.

2.4.3.1 Véhicules non coopératifs

Pour modéliser le comportement longitudinal des véhicules classiques, nous avons donc choisi d'utiliser IDM, *Intelligent Driver Model*, proposé par Treiber et Kesting [192, 108]. La fonction f_{IDM} s'écrit sous la forme générale des lois de poursuite (équation 2.2) :

$$f_{IDM}(\dot{x}_n, \Delta x_{n+1}, \Delta \dot{x}_{n+1}) = \ddot{x}_n = a \left[1 - \left(\frac{\dot{x}_n}{v_0} \right)^\gamma - \left(\frac{s^*(\dot{x}_n, \Delta \dot{x}_{n+1})}{\Delta x_{n+1}} \right)^2 \right] \quad (2.4)$$

où

- $s^*(\dot{x}_n, \Delta \dot{x}_i) = s_0 + \max(0, \dot{x}_i T + \frac{\dot{x}_n \Delta \dot{x}_i}{2\sqrt{ab}})$ exprime l'inter-distance temporelle désirée ;
- a et b sont respectivement l'accélération maximum et l'accélération désirée ;
- v_0 est la vitesse désirée ;
- s_0 est la distance inter-véhiculaire à l'arrêt ;
- T est l'inter-distance de sécurité (temporelle) ;
- γ est un coefficient d'agressivité de l'accélération.

Les changements de voies sont modélisés par une stratégie opportuniste nommée MOBIL (*Minimizing Overall Braking Induced By Lane-changes*), des mêmes auteurs [104]. Avant de mettre à jour leur accélération, les véhicules peuvent appliquer cette stratégie, qui intègre les changements de voie volontaires (dépassement) et les changements obligatoires (liés au réseau). La figure 2.13, tirée de l'article de Kesting *et al.* [104], décrit une situation gérée par le modèle.

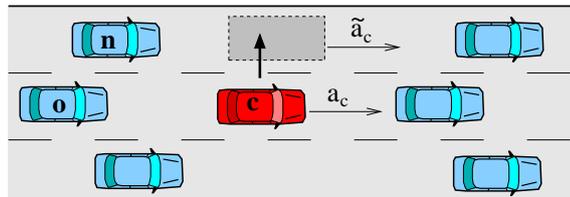


Fig. 2.13: Schéma (extrait de [104]) représentant le voisinage d'un véhicule courant c souhaitant effectuer un changement de voie en suivant la stratégie MOBIL. Les véhicules suiveurs – ancien o et nouveau n – sont pris en compte dans la décision. L'accélération du véhicule c passera, après application du changement, de a_c à \tilde{a}_c .

La stratégie MOBIL estime si un créneau disponible est suffisamment sûr (*i.e.* la décélération imposée au potentiel suiveur est limitée) en calculant l'inter-distance temporelle sur la/les voie(s) adjacente(s). Ce modèle vise à reproduire le comportement des conducteurs en intégrant des paramètres facilement interprétables, comme un facteur de politesse par exemple.

2.4.3.2 Véhicules coopératifs

Grâce aux informations échangées entre les véhicules coopératifs, leur perception peut être élargie significativement. Contrairement aux véhicules classiques, ils peuvent intégrer dans leur modèle longitudinal l'ensemble de leurs voisins connectés. A cet effet, les véhicules coopératifs exécutent la loi Bilatérale Multi-Ancipative [138] (BMA) qui s'exprime comme la combinaison des influences pondérées des véhicules voisins. Reprenant la forme générale des modèles basés sur la multi-anticipation (équation 2.3), elle s'écrit :

$$\ddot{x}_i = f_{IDM} \left(\dot{x}_i, \sum_j a_{ij} \Delta x_{i+j}, \sum_j a_{ij} \Delta \dot{x}_{i+j} \right), \quad (2.5)$$

où a_{ij} est le coefficient d'interaction permettant de pondérer l'information échangée entre le véhicule courant i et le véhicule j , et où f_{IDM} correspond à l'équation 2.4. Les coefficients d'interaction sont calculés à partir d'une règle de proximité (section 2.4.4) de la couche de communication et en intégrant l'estimation de la confiance associée à chaque véhicule. Le coefficient d'interaction s'exprime ainsi :

$$a_{ij} = p_{ij} \cdot T_{ij}, \quad (2.6)$$

où p_{ij} correspond à la règle de proximité (éq. 2.7), et T_{ij} représente la confiance que l'agent i a en l'agent j (éq. 2.8).

Le résultat du calcul et de l'application successive de la loi de proximité et de la confiance, qui contribuent aux coefficients d'interaction, est la position et la vitesse relative d'un véhicule "virtuel" que le véhicule courant suit. Le déplacement du véhicule virtuel par rapport au véhicule prédécesseur réel permet de faire reposer le modèle sur une loi de poursuite dont le comportement est validé. Ainsi, si le véhicule se trouve dans une situation où il est privé de communication (ou qu'aucun véhicule connecté n'est à portée), le comportement du modèle BMA va correspondre au comportement généré par IDM. Cela assure aussi la transition entre les différents environnements, où le taux d'équipement peut varier dans le voisinage local du véhicule.

La loi BMA utilisée modélise le comportement d'un système automatique, qui contrôle dynamiquement l'accélération des véhicules coopératifs. Le comportement latéral (changement de voie), laissé à la discrétion du conducteur, est modélisé en utilisant la même stratégie MOBIL [104]. Les véhicules coopératifs sont donc équipés, connectés et semi-automatisés, ce qui leur permet d'avoir un comportement allant vers l'homogénéisation du flux. Pour atteindre ce but, les informations échangées et leur fiabilité sont cruciales.

2.4.4 Communication

Comme nous l'avons étudié, la communication (directe ou indirecte) est préalable à toute forme de coopération (sections 2.1.3 et 2.2). Seuls les véhicules coopératifs, qui sont des véhicules

connectés, peuvent envoyer et recevoir des messages. Dans ce travail, nous avons défini deux types de messages : les messages de mesures et les messages de confiance. Les messages de mesures regroupent les informations issues des capteurs qui sont partagées par les véhicules (position et vitesse relative au prédécesseur local et position, vitesse dans l'espace global). Les messages de confiance contiennent les "témoignages", *i.e.* les valeurs de confiance partagées entre les agents.

Les informations contenues dans ces messages permettent aux véhicules connectés d'estimer l'importance qu'a un autre véhicule dans sa prise de décision. L'idée de base est de coupler les informations physiques (position, vitesse) aux informations de contrôle (confiance). Grâce aux messages échangés, le véhicule courant peut projeter ces nouvelles informations dans son espace local et ainsi construire une perception étendue de son environnement de navigation. Une règle de proximité p_{ij} permet de pondérer l'influence des véhicules voisins à partir de leur position et vitesses relatives :

$$P_{ij} = \frac{|\Delta \dot{x}_{i+j}|}{(\Delta x_{i+j})^\delta}, \quad (2.7)$$

et par normalisation : $p_{ij} = \frac{P_{ij}}{\sum_{k \in \mathcal{P}_i} P_{ik}}$;

où \mathcal{P}_i est l'ensemble des agents perçus par i et δ permet de contrôler l'interaction vers l'avant/arrière. Les valeurs respectivement retenues sont 2 et 0,5 pour assurer la stabilité de la chaîne de véhicules et respecter les caractéristiques du flux. Ces valeurs permettent de garantir la sécurité principalement liée aux interactions avec les véhicules situés à l'avant, tout en permettant d'obtenir une influence plus faible mais significative pour les véhicules situés à l'arrière (queue de bouchon).

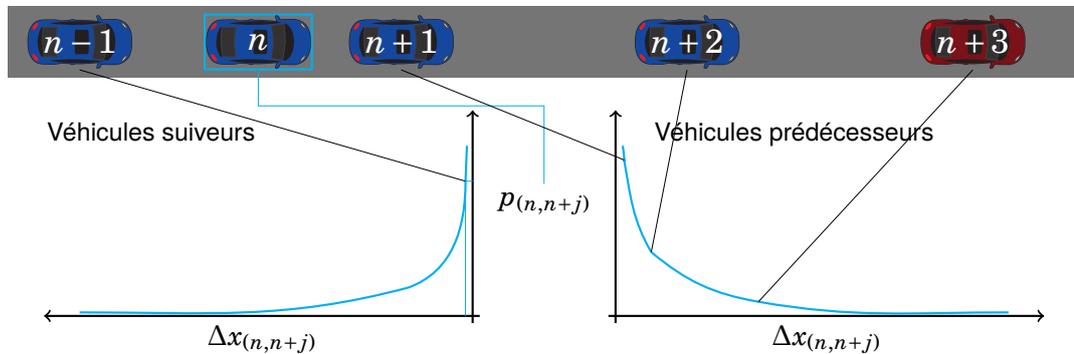


Fig. 2.14: Illustration de l'application de la règle de proximité. Le véhicule courant n pondère l'influence des véhicules suiveurs et prédécesseurs, connectés (en bleu) et non-connectés (rouge).

Le principe de la règle de proximité, représentée dans la figure 2.14, est de donner plus d'importance aux véhicules qui sont plus proches et/ou qui ont la vitesse relative la plus importante. Les véhicules coopératifs peuvent ainsi anticiper les comportements hétérogènes ; la stabilité et la sécurité du flux sont ainsi maintenus. Le capteur embarqué situé à l'avant du véhicule constitue une source d'information directe. Dans le cas où le véhicule connecté ne reçoit aucune information via la communication (défaillance du dispositif par exemple), son comportement, donné par le modèle BMA (éq. 2.5), correspondra au modèle non coopératif IDM (éq. 2.4). Cela permet aussi d'assurer

une continuité du comportement lorsque le véhicule traverse des zones où se trouvent beaucoup ou peu de véhicules connectés.

2.4.5 Fiabilité des capteurs

La fiabilité des capteurs est un point crucial dans le modèle BMA puisque les données échangées interviennent dans le calcul de la règle de proximité (éq. 2.7) et des coefficients d'interaction (éq. 2.6). Nous utilisons un modèle basé sur la confiance pour estimer la fiabilité des informations dans le système et ainsi modifier dynamiquement l'influence des agents qui en sont la source. Nous proposons d'étendre le modèle proposé par [149, 136], qui exploite une structure de données représentant un réseau de confiance nommée *TrustNet* [179]. Ce modèle permet de représenter numériquement la confiance à partir des expériences non agrégées des agents. Ce modèle évalue la crédibilité d'un agent à partir de la majorité des autres agents, ce qui correspond bien au cadre du trafic coopératif où les véhicules sont supposés fiables en moyenne. Bien qu'ils puissent avoir une influence négative sur le système, notre approche ne s'intéresse pas aux comportements égoïstes ou de tricheries, qui pourront faire partie des perspectives de nos travaux. La confiance calculée tient ici compte de la fiabilité des informations provenant des capteurs embarqués qui pourraient être défaillants et propager des erreurs qui devraient être visibles dans l'évolution des valeurs de confiance.

Pendant leur phase de décision, les agents comparent les valeurs de confiance dont ils disposent (celles qu'ils mémorisent et celles partagées dans les messages de confiance). Ainsi, chaque agent i peut mémoriser et mettre à jour une opinion locale de tous les agents j qu'il a précédemment perçus ; elle est représentée sous la forme d'une valeur de confiance T_{ij} telle que :

$$T_{ij} = \frac{ST_{ii}DT_{ij} + IT_{ij}}{ST_{ii} + 1}. \quad (2.8)$$

Cette valeur de confiance est calculée à chaque prise de décision de l'agent i . Il est possible de la formaliser à partir de trois composantes numériques, toutes dans l'intervalle $[0, 1]$:

1. **La confiance directe DT_{ij}** : qui représente l'erreur de mesure du capteur du véhicule. Elle représente l'évaluation de l'erreur en termes d'inter-distance et de vitesse relative qui a été mesurée par le véhicule j voisin :

$$DT_{ij} = \frac{|\Delta x_i + x_i - x_j|}{\Delta x_i + x_i}, \quad (2.9)$$

où Δx_i et x_i correspondent aux mesures provenant du véhicule courant i (respectivement depuis le télémètre et le dispositif de localisation), et x_j est la position dans le repère global partagée par j (estimée par son propre dispositif de localisation).

2. **La confiance indirecte IT_{ij}** : qui fait référence aux valeurs de confiance partagées par les autres agents dans les messages de confiance, *i.e.* les témoignages des voisins. Elle s'écrit :

$$IT_{ij} = \frac{\sum_{k \in \mathcal{W}_i} T_{ik} T_{kj}}{\sum_{k \in \mathcal{W}_i} T_{ik}}, \quad (2.10)$$

où \mathcal{W}_i désigne l'ensemble des agents k qui ont partagé la valeur de confiance qu'ils accordent à j avec l'agent i (pour les agents k situés dans le rayon de communication de i). Ainsi, les

agents k doivent avoir précédemment comparé leurs informations directes avec j , ce qui réduit la taille du réseau de confiance et limite l'ampleur des calculs.

3. **La confiance en soi** ST_{ii} : qui est modélisée comme la moyenne des valeurs de confiance que les autres agents ont relativement à l'agent i considéré. Cette valeur intègre donc les mises à jour successives lors des évaluations précédentes. La valeur de confiance en soi s'exprime :

$$ST_{ii} = \frac{\sum_{k \in \mathcal{C}_i} T_{kk} T_{ki}}{\sum_{k \in \mathcal{C}_i} T_{kk}}, \quad (2.11)$$

où \mathcal{C}_i désigne l'ensemble des agents k qui ont partagé leur valeur de confiance en i (les agents k sont donc situés dans le rayon de communication de i). Cette valeur de confiance en soi représente ainsi la confiance que l'agent considéré accorde en lui même. Elle peut être utilisée pour adapter le comportement d'un agent qui se déclarerait de lui même défaillant et pourrait donc limiter l'utilisation de ces capteurs ou stopper toute communication pour ne pas propager plus d'erreur.

Grâce au couplage mis en place, les véhicules coopératifs peuvent tirer profit des informations échangées. La fiabilité de ces informations est estimée par chaque véhicule tout au long de son parcours sur le réseau. Les échanges sous la forme de messages de confiance permettent d'étendre significativement leur perception. La fiabilité d'un véhicule, qui a été estimée sur un tronçon par un groupe de véhicules, est partagée avec les autres et leur permet très rapidement d'intégrer cette donnée dans leur processus de décision. La robustesse du système tient dans le partage systématique et la mise à jour des informations locales.

2.5 Bilan et perspectives

Un des objectifs de notre travail est de proposer un cadre de modélisation des systèmes coopératifs permettant d'intégrer toutes les interactions des nouveaux types de véhicules et de l'infrastructure intelligente. En effet, les C-ITS sont des systèmes complexes dont les dynamiques sont inter-dépendantes. Plusieurs acteurs se partagent une même infrastructure et l'ajout de capacités de communication risque de complexifier les échanges. La coopération entre les entités du système est la solution à mettre en place pour permettre à tous les usagers un usage de l'infrastructure plus efficace et avec une sécurité accrue. Même si le socle technologique semble désormais propice au développement d'approches aussi ambitieuses, les stratégies actuelles sont encore limitées. Cela tient pour beaucoup de la complexité des échanges et de l'aspect distribué de l'information. Dans ce chapitre, nous avons exploré les voies permettant de modéliser les différentes dynamiques des systèmes coopératifs : la dynamique physique, la communication mais aussi la dynamique de contrôle. Ce dernier point apparaît crucial dans l'objectif de crédibilité et d'acceptation par les usagers des nouveaux comportements qui pourraient émerger de la coopération entre véhicules. Pour modéliser toutes ces dynamiques, notre proposition s'appuie sur les récentes propositions de la théorie du trafic, couplées avec l'apport des approches décentralisées en Intelligence Artificielle. Grâce à l'apport des systèmes multi-agents, il est possible d'intégrer dès la modélisation des caractéristiques menant à l'autonomie des entités : perception par des capteurs, protocole de communication, représentation de la confiance. L'ensemble de ces modèles peut alors exhiber des comportements émergents à l'échelle du flux, fruit de la coopération entre les agents. Comme une illustration de ce type de phénomène, nous proposons un modèle multi-anticipatif bilatéral, capable de réguler un flux de véhicules connectés et partiellement automatisés. La dynamique de

contrôle prend la forme d'un modèle de confiance qui permet d'estimer la fiabilité des informations échangées.

Le prochain défi est de choisir comment implémenter ce couplage dans une plateforme de simulation. Cela passe d'abord par la sélection d'un outil existant qui sera compatible avec notre approche par couplage. Le fruit de ce travail est un simulateur qui intègre à la fois les bénéfices des systèmes multi-agent (perception orientée capteur, autonomie et phénomènes émergents) mais profite aussi des modèles de la communauté du trafic éprouvés et validés. Le cœur de la contribution se concentre sur l'aspect communication, central dans les systèmes coopératifs. Enfin, les objectifs liés à la régulation ou au contrôle de ces systèmes imposent d'explorer les possibilités offertes par les différents types de véhicules : connectés, autonomes, etc. La stratégie locale de régulation proposée peut être confrontée à différents environnements dynamiques grâce au simulateur, qui servira de socle à l'élaboration de stratégies de contrôle plus complexes.

Simulateur multi-agent du trafic coopératif

” *Learning by doing, peer-to-peer teaching, and computer simulation are all part of the same equation.*

— **Nicholas Negroponte**
Professeur et chercheur, MIT

Les systèmes coopératifs et les véhicules connectés n'en sont aujourd'hui qu'à leur phase de prototype. Les initiatives se multiplient et montrent des premiers résultats expérimentaux encourageants. Notre travail s'inscrit dans les prochaines phases de déploiement de ces systèmes. Nous souhaitons anticiper l'arrivée plus massive des véhicules connectés et automatisés sur les réseaux routiers. Cette phase passe par une étude reposant sur un outil de simulation. Notre analyse des modèles existants a mis en lumière des analogies prometteuses entre le domaine de la simulation de trafic routier et les approches basées sur les systèmes multi-agents. Les bénéfices possibles de ces systèmes distribués (auto-organisation, émergence) nous ont permis de proposer un cadre de modélisation par couplage des dynamiques des C-ITS. Ce couplage permet de modéliser les différents flux qui existent dans les systèmes coopératifs, en introduisant une dynamique de contrôle grâce au concept de confiance (chapitre 2). La suite de notre travail consiste à proposer un outil de simulation capable d'accueillir ce type de modélisation et d'évaluer la portée de nos contributions.

Plusieurs solutions sont envisageables dans la réalisation de cet outil de simulation. Proposer un nouvel outil sans réutiliser les contributions existantes semble contre-productif puisque l'ampleur de la tâche risquerait de nous détourner de nos objectifs. Les contributions des simulateurs existants peuvent sans doute former une bonne base. La question restante porte sur le type de simulateur, et notamment sur la communauté visée. Si le choix s'arrête sur un simulateur dédié aux technologies agent, le travail à accomplir sera principalement d'intégrer des connaissances issues de la théorie du trafic. Cette possibilité risque de limiter l'apport du domaine du trafic, qui regorge pourtant de modèles éprouvés et réalistes. Nous avons donc choisi l'option inverse qui consiste à démarrer notre réflexion à partir d'un simulateur de trafic parmi ceux de la littérature, à savoir MovSim (section 3.1). Nous proposons une extension de cet outil de simulation afin d'y intégrer les concepts multi-agents et permettre la modélisation par couplage proposée précédemment (chapitre 2). L'ensemble des modifications apportées a donné naissance à un outil d'aide à la décision (section 3.2). Comme une illustration du couplage, nous proposons un modèle microscopique visant à homogénéiser un flux de véhicules partiellement automatisés et aussi connectés. Grâce à la mise en place d'indicateurs (section 3.3), il est possible d'évaluer l'impact de telles stratégies sur le flux. Les résultats dans la partie 3.4 confirment le potentiel du couplage des modèles physiques et de l'apport de la confiance sur l'homogénéisation, et ce, même en présence de capteurs défaillants. Ces premiers résultats nous orientent ensuite vers l'élaboration de stratégies de contrôle à une échelle supérieure basée sur l'utilisation des unités d'infrastructure (section 3.5).

3.1 Approches distribuées et simulation : analyse de l'existant

Parmi les finalités principales de la modélisation du trafic et de ses évolutions, on retrouve la simulation et la prévision. La prévision utilise un modèle pour prévoir l'état futur d'un système, alors que la simulation s'intéresse à construire des scénarios afin de reproduire des comportements observés. Dans les deux cas, l'échelle de modélisation est d'une importance capitale puisqu'elle déterminera le niveau de détail atteignable. Dans le cadre de nos travaux, nous avons identifié un clair intérêt à s'orienter vers des modèles microscopiques. Nous souhaitons étudier de nouveaux comportements mettant à profit les technologies ITS dans un environnement connecté et coopératif. Dans ce cadre, la simulation offre une première étape vers le développement et la modélisation de nouveaux comportements ou de nouvelles stratégies de contrôle. La suite du travail présenté a fait l'objet d'une implémentation au sein d'un simulateur de trafic. La littérature existante confirme que beaucoup d'efforts ont été portés au développement de plateformes dédiées à la simulation de trafic. Repartir de zéro impliquerait de perdre une expertise acquise par la communauté scientifique. Bien que notre approche se distingue clairement des approches de modélisation classiques, nous avons sélectionné un simulateur qui permettra d'implémenter et de valider les approches développées dans ce travail. Cette section compare les plateformes de simulation existantes et justifie le choix d'un simulateur.

La simulation microscopique de trafic a suscité beaucoup d'intérêt en recherche et développement, aussi bien pour des plateformes commerciales que pour des logiciels libres (*open-source*). Parmi les simulateurs les plus utilisés dans la communauté de la modélisation de trafic, se trouvent SUMO [114], AIMSUN [14], VISSIM [62], DYNASMART [126], TRAFFSIM [11], et MovSim [193]. La modélisation des systèmes coopératifs partage des points communs avec la modélisation des réseaux Ad-Hoc de véhicules (VANET) pour lesquels beaucoup de simulateurs existent (le lecteur intéressé trouvera une revue complète dans l'article de Martinez *et al.* [130]). Certaines plateformes comme MATSim [12] exploitent les concepts multi-agents (paradigme d'Intelligence Artificielle distribuée présenté dans la section 2.1.3). La plateforme ARCHISIM [59] modélise aussi les conducteurs grâce à des comportements individuels d'agents autonomes, et se distingue de MATSim qui adopte un point de vue plus orienté flux de trafic. L'usage des concepts issus des SMA dans la modélisation de trafic n'est pas récente [18], mais ne s'est considérablement développé que très récemment [16]. Cette tendance recoupe les conclusions de notre étude des modèles de trafic (section 2.1.1) avec les besoins spécifiques des systèmes coopératifs (section 1.1.4).

Dans le chapitre 1, nous avons identifié les sections à caractère autoroutier comme les plus propices à bénéficier d'un impact positif lors de l'introduction des véhicules coopératifs. Pour les besoins de notre travail, la plateforme choisie ne devra donc pas forcément proposer une définition très détaillée d'un réseau maillé. Nos affinités avec les logiciels libres et communautaires nous poussent à retenir deux candidats principaux : SUMO et MovSim, qui proposent en plus beaucoup de fonctionnalités de par leur stade de maturité par rapport aux plateformes précédentes. SUMO est très utilisé et bien représenté dans les travaux de recherche, et MovSim est une plateforme plus récente et pour le moment moins représentée. Le simulateur MovSim a été développé à partir des concepts les plus récents en théorie du trafic [194, 105]. Concernant les communication inter-véhiculaires V2V, une extension de SUMO est en cours de développement et vise à étudier les effets des applications

embarquées sur le comportement des conducteurs [114]. À cela s'ajoute le projet VEINS qui vise à proposer un ensemble de modèles dédiés aux communications inter-vehiculaires (IVC) dans SUMO [184]. Du côté de MovSim, les développeurs ont imaginé l'intégration de ces formes de communication dès le début du projet [180, 189, 106]. Enfin, chacune de ces deux plateformes est codée dans un des langages de programmation les plus usités : C++ pour SUMO et Java pour MovSim.

Dans ce travail, nous avons sélectionné le *Multi-model Open-source Vehicular-traffic SIMulator* (MovSim) proposé par Treiber et Kesting [193]. Nous avons proposé une extension de ce simulateur afin d'y intégrer une modélisation des systèmes coopératifs. L'architecture générique de MovSim et les nombreux modèles microscopiques implémentés ont motivé notre choix. Ce simulateur propose une représentation unidimensionnelle du réseau (les véhicules se déplacent sur des lignes), et utilise des lois de poursuite (section 2.1.1.4) grâce à une vision décentralisée. La modélisation orientée agent a inspiré l'architecture de MovSim [105]. L'application comprend une interface graphique permettant de visualiser le résultat des simulations. Chaque simulation est décrite dans un fichier de scénario écrit en XML, qui regroupe tous les paramètres utilisés (définition du réseau, paramètres des véhicules, flux d'entrée/sortie, etc.). Malgré la jeunesse de la plateforme, plusieurs travaux récents montrent un attrait vers MovSim. Par exemple, Van Willigen *et al.* [202] proposent une approche de prévision de l'état de trafic qui est améliorée grâce à des données simulées depuis MovSim. Dans un autre registre, Yan *et al.* [215] proposent une extension de MovSim dédiée à l'étude de solutions innovantes de conduite en convoi (*platooning*) basées sur les technologies ITS. Des travaux comme ceux de Schönhof *et al.* [180] étudient des solutions d'anticipation de la congestion grâce à une estimation locale de l'état de trafic et un algorithme coopératif de remontée d'information V2V.

L'étude de plateformes de simulation microscopique de trafic, en plus de confirmer l'intérêt des chercheurs pour les nouvelles technologies C-ITS, a révélé une tendance qui tend vers la pluridisciplinarité. En effet, les avancées récentes en génie logiciel et en informatique en général ont profité aux plateformes. Certaines se sont notamment tournées vers l'intelligence artificielle distribuée et un paradigme associé nommé "systèmes multi-agents". Cette tendance s'explique par la complexité croissante des interactions induites par la communication et la coopération. La modélisation sous la forme d'agents autonomes permet de simplifier la phase de conception (définition de comportements individuels au niveau microscopique) tout en espérant des propriétés émergentes (visibles au niveau macroscopique). Nous pensons que ce paradigme est la clé vers une modélisation plus fine des systèmes coopératifs, et que les travaux en intelligence artificielle peuvent significativement inspirer la recherche de solutions innovantes répondant aux enjeux des C-ITS. Ce vaste domaine est présenté dans la section 2.1.3. Malgré tous les efforts récents allant vers l'introduction de la communication dans les modèles, aucune plateforme n'est à l'heure actuelle capable d'offrir toutes les fonctionnalités préliminaires à notre travail. L'infrastructure intelligente (ou communicante), par exemple n'est pas ou peu représentée dans les plateformes. Le choix de MovSim s'inscrit donc pleinement dans les objectifs de notre travail. Nous proposons ainsi d'étendre la plateforme pour modéliser le trafic coopératif grâce aux concepts multi-agents. Cette extension a donné naissance à la plateforme MASCAT que nous présentons plus en détail dans la section 3.2.

3.2 Modèles et implémentation

Le niveau de déploiement actuel des systèmes coopératifs ne permet pas aujourd'hui une étude à grande échelle à partir de données réelles. L'intérêt principal de la simulation est de placer un système dans une configuration nouvelle, difficile à mettre en place, à reproduire ou à maîtriser et de tester des scénarios difficilement accessibles dans la réalité (e.g. variation du taux d'équipement). Beaucoup d'approches se limitent à la reproduction la plus exacte possible de phénomènes observés dans un cadre de simulation. Pourtant, une architecture propice à l'élaboration de solutions innovantes en simulation, même souffrant d'un décalage avec la réalité (le problème du *reality gap*, très connu en robotique [112]), peut constituer un outil d'aide à la décision pour les acteurs du déploiement de systèmes complexes comme les C-ITS. Le cadre d'étude des systèmes coopératifs soulève en effet des questions préalables au déploiement : quelles informations échanger ? Où équiper l'infrastructure ? Comment contrôler les véhicules ? Quand mettre en place les stratégies de régulation ? Le positionnement de notre travail est d'étudier ces questions en amont, en concentrant notre réflexion sur des scénarios qui ne seront pas réalisables avant quelques années. L'idée est d'anticiper les besoins et les apports des C-ITS, dans l'optique d'évaluer les bénéfices et les limitations de ces systèmes. Une partie de notre travail a donc consisté à produire un cadre de simulation générique nommé MASCAT (*Multi-Agent Simulator for Connected and Automated Traffic*) permettant de proposer et d'étudier différentes stratégies de contrôle des C-ITS.

3.2.1 Extension de MovSim : MASCAT

La revue des plateformes de simulation (section 3.1) a montré un intérêt croissant allant vers la modélisation des nouvelles technologies de communication inter-véhiculaires (IVC). La complexité des interactions entre les entités (mobiles ou fixes) des C-ITS a naturellement suivi la tendance de la simulation multi-agent, particulièrement adaptée à la modélisation de tels problèmes distribués. Pourtant, aucune plateforme existante ne rassemble une expertise dans les modèles de trafic combinée à une modélisation permettant un cadre de travail robuste et générique, incluant la communication (V2V et I2V/V2I). Dans le cadre de notre travail, nous sommes partis d'un simulateur microscopique de trafic, le *Multi-model Open-source Vehicular-traffic SIMulator*¹ – MovSim [193].

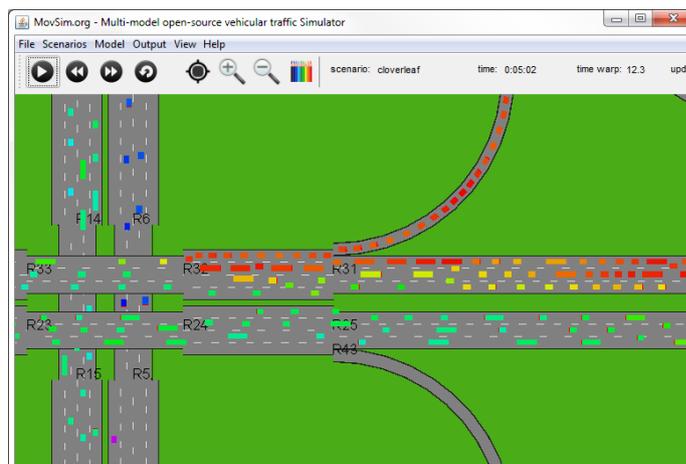


Fig. 3.1: Capture de l'interface du simulateur MovSim

1. Page web de MovSim : <http://www.movsim.org/>

MovSim propose plusieurs fonctionnalités qui ont motivé notre choix :

- Il s'agit d'un simulateur microscopique en temps continu,
- qui implémente plusieurs lois de poursuite (OVM [13], Newell [146], Gipps [76], IDM [192, 108], ACC [194], ...),
- et la stratégie de changement de voie opportuniste MOBIL [104].
- La représentation générique du réseau routier (formalisme OpenDrive [56]) inclut les limitations de vitesses, les feux de signalisation et permet de modéliser des intersections.
- Une définition détaillée des scénarios permet de générer et de modifier dynamiquement des profils de demande sur les sections d'entrée.
- L'interface graphique est paramétrable et adaptable.
- L'architecture choisie reprend les principes de la simulation multi-agent, et est implémentée en Java.

Les éléments développés dans le cadre de notre travail, prennent la forme d'une extension de MovSim [83, 81], baptisée MASCAT (*Multi-Agent Simulator for Connected and Automated Traffic*), qui permet :

- de modéliser les échanges d'informations grâce à un protocole de communication V2V et I2V/V2I ;
- de proposer un modèle agent pour le comportement des véhicules ;
- d'ajouter l'infrastructure en modélisant les Unités de Bord de Route comme des agents autonomes ;
- de fonder le processus d'exécution des agents sur une architecture perception-décision-action ;
- de décomposer la perception des agents grâce à des capteurs (véhicules et infrastructure) ;
- d'instancier le couplage à travers l'implémentation des modèles de confiance et la multi-anticipation ;
- de représenter les nouvelles fonctionnalités et d'observer les résultats de simulation à travers une interface graphique dédiée aux C-ITS ;
- de disposer d'un cadre de simulation stable pour développer des stratégies de contrôle des systèmes coopératifs.

Malgré l'ampleur de l'extension, la version originale de MovSim a constitué une base stable. Les modifications majeures se situent au niveau de l'implémentation de la perception et des mécanismes de communication.

Comme le montre la figure 3.2, nous avons souhaité garder une grande partie des composants originaux afin de ne pas remettre en cause le cœur du simulateur. Le modèle multi-anticipatif bilatéral (BMA) est implémenté grâce aux modifications de l'extension et reprend le principe du couplage des dynamiques proposé dans notre travail (section 2.3).

3.2.2 Architecture

Le principe général adopté dans MASCAT est de faire correspondre le cycle d'exécution des entités du système à un cycle perception-décision-action d'un agent autonome (figure 3.3). Les véhicules comme les entités d'infrastructures peuvent alors être modélisés individuellement et il

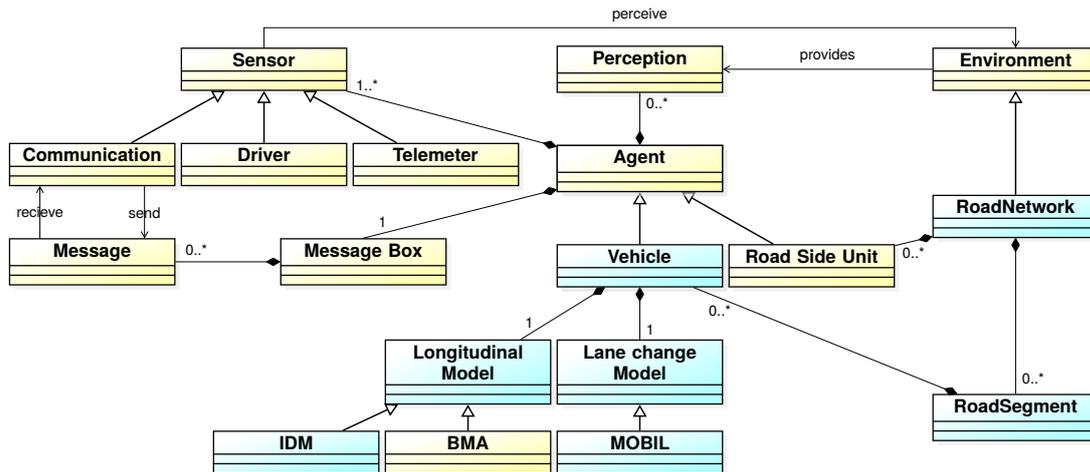


Fig. 3.2: Diagramme de classes simplifié de MASCAT, extension de Movsim. Les classes représentées en bleu sont les classes originales du simulateur.

est possible d'agir sur la définition de leurs comportement en modifiant les caractéristiques de leur environnement, qu'ils perçoivent par l'intermédiaire de leurs capteurs. Les actions, issues de leur processus autonome de prise de décision, sont la traduction dans le monde physique des influences qu'ils calculent.

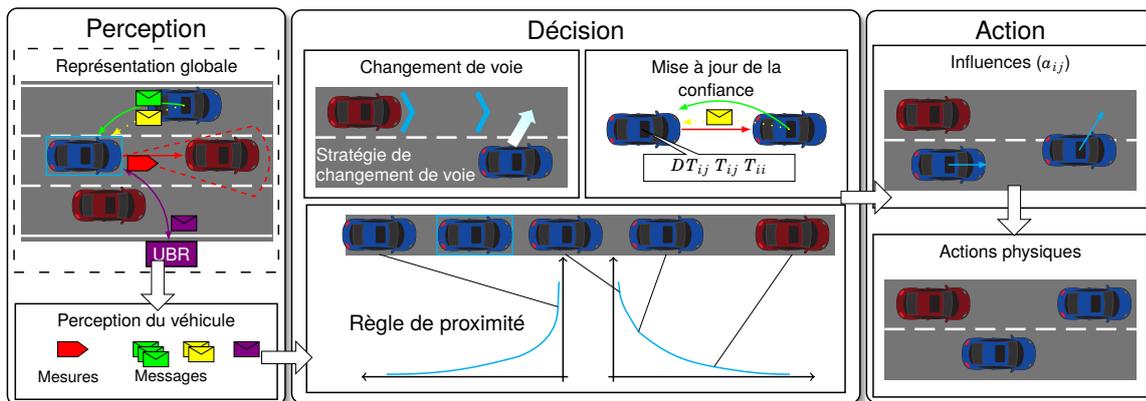


Fig. 3.3: Schéma d'opérations de la simulation dans MASCAT

Ainsi, l'environnement, qui était déjà au centre de l'architecture originale, reste le point central de la simulation. Afin de respecter le principe d'autonomie dans le processus de décision des agents (*i.e.* l'ordre d'exécution des agents ne doit pas avoir d'incidence sur le résultat), l'exécution est déléguée à un ordonnanceur qui gère le temps en pas discrets. En suivant les principes des systèmes multi-agents situés (section 2.1.3), le principe général est de considérer une situation fixe de l'environnement ; de calculer la perception de chaque agent ; de lancer leur processus de décision indépendamment ; et enfin d'appliquer les actions décidées par les agents.

3.2.3 Perception

La première phase de l'architecture multi-agent proposée repose sur l'environnement. La perception de chacune des entités (fixes ou mobile) est calculée indépendamment et au même moment. L'idée de base consiste à figer l'état de l'environnement et à exécuter chaque capteur de chaque agent afin de générer les données recueillies. Ces données sont ensuite traitées lors de la phase de décision des agents.

3.2.3.1 Capteurs

L'architecture retenue dans la modélisation des C-ITS se repose sur une définition générique des capteurs, dispositifs permettant aux agents (véhicules et infrastructure) de recueillir des données utiles à leur prise de décision. Comme représentés dans la figure 3.2, des capteurs spécifiques permettent d'équiper des véhicules connectés (communication) et/ou équipés (Télémètre) ou des véhicules classiques (le conducteur est alors considéré comme un capteur).

Véhicules

Afin d'étendre leur perception, les véhicules peuvent être équipés de capteurs. Il existe actuellement un grand nombre de capteurs embarqués dont la fonction est semblable : donner une représentation de l'environnement de navigation du véhicule. Les données provenant de ces capteurs sont souvent brutes, *i.e.* elles sont entachées d'un bruit de mesure propre au capteur et à ses caractéristiques. Un algorithme de fusion de capteurs permet d'agréger les données de plusieurs capteurs (le cas échéant) afin de générer un modèle de l'environnement.

Dans notre modélisation, nous avons choisi de modéliser les capteurs de manière générique et extensible. Seuls certains capteurs (les plus représentés dans les technologies ITS) sont implémentés. Au niveau des véhicules, il s'agit avant tout de télémètres laser. Ces dispositifs utilisent un (ou plusieurs) rayons optiques (émis dans un spectre non visible comme l'infrarouge) pour calculer la distance et la vitesses relative aux objets environnants. Comme il s'agit de capteurs embarqués, le référentiel des capteurs est local (*i.e.* relatif au véhicule).

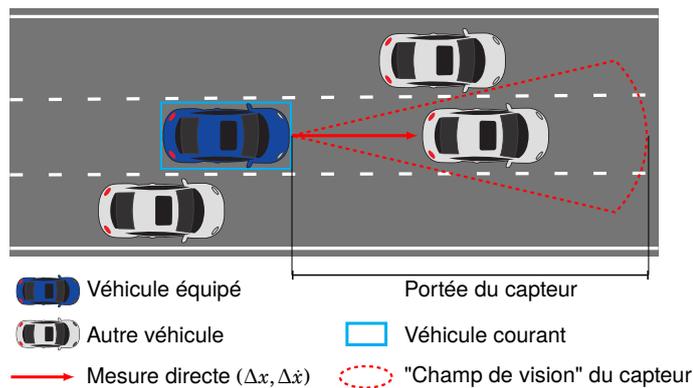


Fig. 3.4: Représentation de la perception directe d'un véhicule équipé d'un capteur (télémètre laser).

La figure 3.4 illustre le champ de vision d'un véhicule équipé d'un télémètre laser à balayage. En tenant compte des technologies actuelles, nous avons modélisé un capteur capable de mesurer

l'inter-distance Δx_{i+1} ainsi que la vitesse relative $\Delta \dot{x}_{i+1}$ avec le véhicule prédécesseur. La portée maximale du capteur est de 80 mètres ; au delà, les véhicules ne sont plus perçus. Nous faisons l'hypothèse qu'un algorithme embarqué dans le capteur est capable de mesurer directement ces valeurs à partir des faisceaux lasers.

Les capteurs réels effectuent des mesures avec une plage de précision donnée par le constructeur. Ils peuvent aussi être sensibles aux conditions extérieures (précipitations, luminosité) ou simplement être défaillants. Afin de modéliser la fiabilité des capteurs des véhicules, il est possible d'ajouter un bruit de mesure lors du calcul de la perception. Chaque capteur dispose d'un indice de fiabilité noté $\alpha_i \in [0,1]$. Lorsque le capteur mesure l'inter-distance avec le véhicule prédécesseur, un bruit blanc gaussien est ajouté à la mesure. L'écart-type σ_i de la distribution des mesures du capteur du véhicule i est alors :

$$\sigma_i = \Delta x_{i+1} \times (1 - \alpha_i) \quad (3.1)$$

Les capteurs parfaits (ce qui n'existe pas dans la réalité) peuvent être modélisés en utilisant la valeur $\alpha_i = 1$.

Unités de Bord de Route

Les Unités de Bord de Route sont des entités de l'infrastructure. Leur installation nécessite donc un raccordement au réseau (alimentation et information). L'infrastructure actuelle comporte déjà des capteurs permettant de remonter des informations relatives aux flux, majoritairement dans un objectif de supervision (*monitoring*), permettant à un opérateur (ou un système automatique) d'appliquer des stratégies de régulation. Ces capteurs sont des boucles électromagnétiques, capables, par induction, de détecter des véhicules circulant sur une voie donnée. En comptant les véhicules pendant un pas d'agrégation (de l'ordre d'une à quelques minutes), il est possible d'obtenir des données telles que le débit (véhicules par unité de temps), ou la concentration (véhicules par unité de distance). Si deux boucles sont couplées (double boucle), il est alors possible d'estimer une vitesse moyenne et même de discriminer le type de véhicule (à partir de leur longueur).

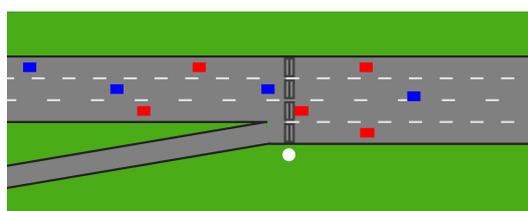


Fig. 3.5: Représentation des boucles électromagnétiques (rectangles gris) et d'une UBR (cercle blanc) dans l'interface de MASCAT.

La version originale de MovSim comportait déjà une implémentation des doubles boucles électromagnétiques, permettant une sortie dans les fichiers de simulation. Notre travail pour MASCAT a consisté à modéliser et à rajouter les UBR comme un nouveau type d'agent, et à assimiler ces boucles à des capteurs génériques. De cette façon, la perception des UBR peut intégrer de l'information provenant d'un ou de plusieurs capteurs, comme illustré dans la figure 3.5. Tout comme les véhicules connectés, les UBR peuvent échanger

de l'information via la communication. Les messages reçus sont utilisés par les agents pour étendre la perception directe, donnée par leurs capteurs respectifs.

3.2.3.2 Communication

L'architecture générale des C-ITS se repose principalement sur l'échange d'informations grâce à des technologies sans fil (voir la section 1.1.2 pour plus de détails sur les standards en matière de com-

munication inter-véhiculaire). La version originale de MovSim ne propose pas de modèle permettant l'échange d'informations. Nous avons donc, dans MASCAT, proposé une modélisation s'appuyant sur la formalisation multi-agent. La perception des agents repose ainsi sur des capteurs génériques. L'un de ces capteurs est un dispositif de communication, dédié à l'envoi (émission) et à la réception de messages. Afin de distinguer les émetteurs des messages (véhicules et infrastructure), chaque entité dispose d'un identifiant unique (conformément au standard en vigueur). Notre modélisation ne prend pas en compte l'échange partiel, les messages modifiés ou la perte d'information, ce qui est assuré par le standard. Cependant, les messages parviennent à leur(s) destinataire(s) avec une latence d'environ 100 millisecondes (ce qui correspond aux valeurs observées sur le terrain [153]). Enfin, les rayons de communications sont fixés à 150 mètres (ce qui correspond aux portées actuelles, ne dépassant pas 300 mètres dans le meilleur des cas).

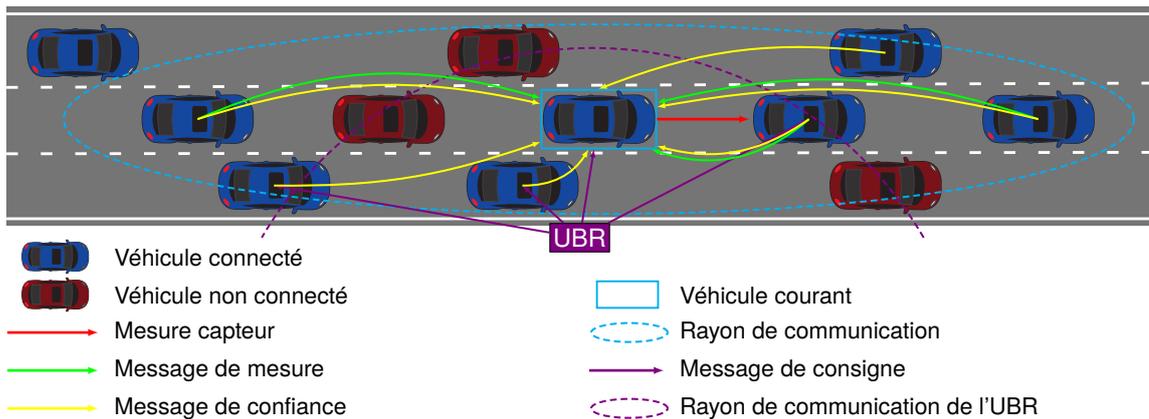


Fig. 3.6: Schéma des communications (V2V et I2V/V2I) possibles dans MASCAT.

Les informations sont échangées sous la forme de messages, qui sont des objets contenant au minimum l'en-tête suivant : identifiant du message, identifiant de l'émetteur, date d'envoi, date d'expiration. En plus de ces informations minimales, différents types de messages (illustrés dans la figure 3.6) permettent de partager d'autres informations. Notre modélisation intègre quatre types de messages :

- Les **messages de mesure** (V2V et V2I) : contiennent des informations, issues de capteurs, partagées par un véhicule à ses voisins (diffusion) ;
- Les **messages de confiance** (V2V) : portent le triplet $(i, j, T_{i,j})$, *i.e.* la valeur de confiance qu'un agent i a en un autre agent j ;
- Les **messages de consigne** (I2V) : permettent aux UBR de transmettre des recommandations afin de contrôler un flux de véhicules connectés.

Le protocole retenu pour les échanges d'information dans les systèmes coopératifs est basé sur de la diffusion (*broadcast*). Ainsi, les messages envoyés n'ont pas de destinataire connu *a priori* ; *i.e.* l'information est partagée avec tous les véhicules présents dans le rayon de communication de l'émetteur. Afin d'étendre la portée effective des messages, leurs en-têtes peuvent contenir des informations relatives à la pertinence de l'information partagée : une zone de pertinence, ainsi qu'une date d'expiration, définis par l'émetteur original du message. Dans le cadre d'un déploiement progressif des technologies ITS et des véhicules connectés, il est nécessaire d'augmenter le nombre de véhicules capables de recevoir les informations échangées. Nous avons implémenté le principe

de rétro-propagation (*re-broadcasting*, ou encore *geo-broadcasting* si la propagation est localisée) permettant aux véhicules de transmettre des messages reçus à d'autres véhicules.

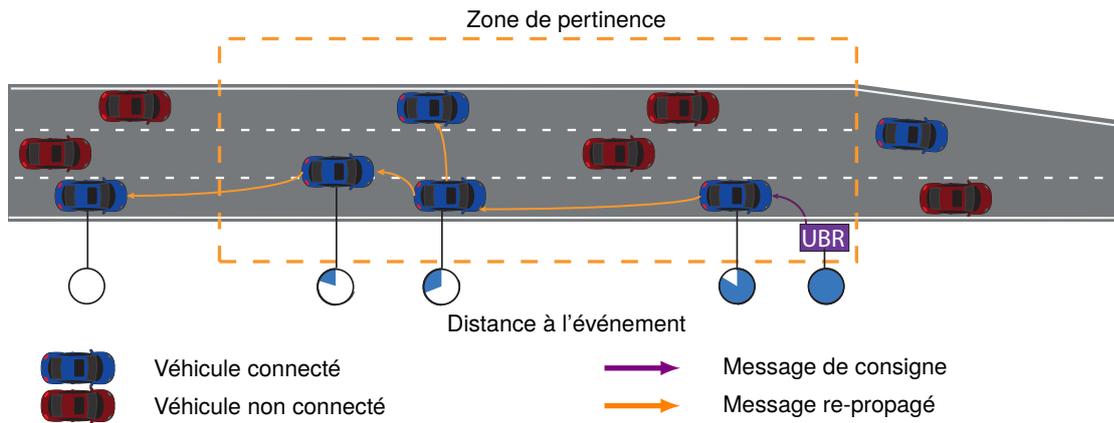


Fig. 3.7: Principe de rétro-propagation (*re-broadcasting*) des messages en V2V.

L'implémentation se base sur la définition de la zone de pertinence du message afin de réguler localement la zone de diffusion V2V du message (figure 3.7). Le principe consiste à estimer localement la distance du véhicule courant à l'événement (ou plus précisément à vérifier si le véhicule se trouve dans zone de pertinence décrite dans le message). Le véhicule décide de retransmettre le message (sans modification). L'information est ainsi propagée dans toute la zone (si suffisamment de véhicules s'y trouvent). L'utilisation du flux en sens inverse est tout à fait envisageable.

Toutes les informations encapsulées dans des messages permettent aux agents (véhicules et UBR) d'étendre leur perception en vue de leur processus de décision.

3.2.4 Décision

La prise de décision est une notion centrale dans les systèmes multi-agents, puisqu'elle définit directement le degré d'autonomie des agents. Dans les systèmes coopératifs, ce degré d'autonomie dépend des entités et des applications. Les véhicules peuvent être modélisés comme des agents réactifs, *i.e.* dont la décision est vue comme une réponse aux stimuli de l'environnement. Les Unités de Bord de Route peuvent décider d'une stratégie de régulation plus complexe. Dans les deux cas, le traitement des informations perçues et reçues est un point central. Ce traitement passe par l'utilisation d'une boîte de réception (visible sur le diagramme de la figure 3.2). Les messages reçus durant le dernier pas de temps de perception sont stockés avant d'être lus, lors de la phase de décision de l'agent. Cette phase aboutit à la sélection d'une action qui sera exécutée dans l'environnement.

3.2.4.1 Infrastructure

Du côté de l'infrastructure, le processus de décision peut prendre plusieurs formes. Si l'on considère les UBR comme de simples relais d'information, elles peuvent re-propager les messages (par

exemple, les message de confiance). Dans notre travail, nous proposons d'utiliser les UBR comme des agents en charge de réguler et de contrôler une section.

Relais d'information

Une UBR permet, grâce à son installation (lien avec le réseau électrique), une portée de communication supérieure aux véhicules. Dans l'optique d'un déploiement progressif des véhicules coopératifs, les UBR permettront d'améliorer la couverture du réseau en relayant les messages transmis entre véhicules (V2V). Dans notre implémentation, le principe est simplifié. Tous les messages envoyés par les véhicules et reçus par une UBR sont placés dans sa boîte de réception. Lors de la phase de décision de l'UBR, tous ces messages sont lus et disparaissent de la boîte de réception. Dans le cas d'un rôle de relais, l'information contenue dans les messages n'est pas utilisée. Les messages lus sont donc directement copiés et déplacés vers une boîte d'envoi des messages. Lors de la prochaine phase de perception, les véhicules connectés à portée recevront tous les messages contenus dans la boîte d'envoi de l'UBR. Cette boîte d'envoi sera vidée après calcul de la perception de tous les véhicules. La résolution des messages éventuellement reçus en doublon s'effectue lors de la phase de perception des véhicules, en utilisant l'identifiant du message contenu dans son entête.

Stratégie de contrôle et de régulation

Grâce à la modélisation mise en place dans notre travail, il est possible d'imaginer d'autres processus de décision plus complexes au sein des UBR. Nous présentons, dans le chapitre 4, un modèle d'agent autonome permettant de proposer des stratégies de contrôle par apprentissage. Ce modèle est implémenté comme le processus de prise de décision d'une UBR chargée de la régulation d'un flux autoroutier partiellement connecté. Le résultat de la prise de décision de l'unité de bord de route est la sélection et la propagation d'une consigne. Cette consigne est intégrée dans un message de consigne et envoyée à tous les véhicules connectés à portée. Les paramètres et le type de consigne sont définis par l'UBR en fonction de sa perception de l'environnement (capteurs et messages V2I). Les consigne reçues sont interprétées par les véhicules qui modifient leur comportement final (phase d'action).

3.2.4.2 Véhicules

Les informations disponibles dans la boîte de réception des véhicules sont triées grâce aux données disponibles dans leur entêtes. Les messages périmés sont ainsi ignorés, les doublons sont supprimés, et les messages à re-propager sont mis de côté.

Extension de la perception

Le principe du modèle multi-anticipatif bilatéral (BMA) repose sur la notion d'anticipation. Cette anticipation est rendue possible par une extension de la perception du véhicule. Le processus de décision des véhicules connectés permet de convertir les informations échangées (dans le repère global) en données locales, qui peuvent être utilisées par le modèle. Ainsi, chaque véhicule connecté estime la vitesse relative et la distance à son prédécesseur direct (grâce à son capteur), puis partage ces données en plus de ses mesures intrinsèques (*i.e.* sa propre position et sa vitesse).

Chaque véhicule connecté reçoit ainsi les informations provenant de son voisinage direct. Seuls les messages provenant de véhicules circulant sur la même voie de circulation sont considérés dans le modèle BMA. Lors de la phase de décision, les véhicules effectuent un passage du référentiel global vers leur référentiel local (figure 3.8).

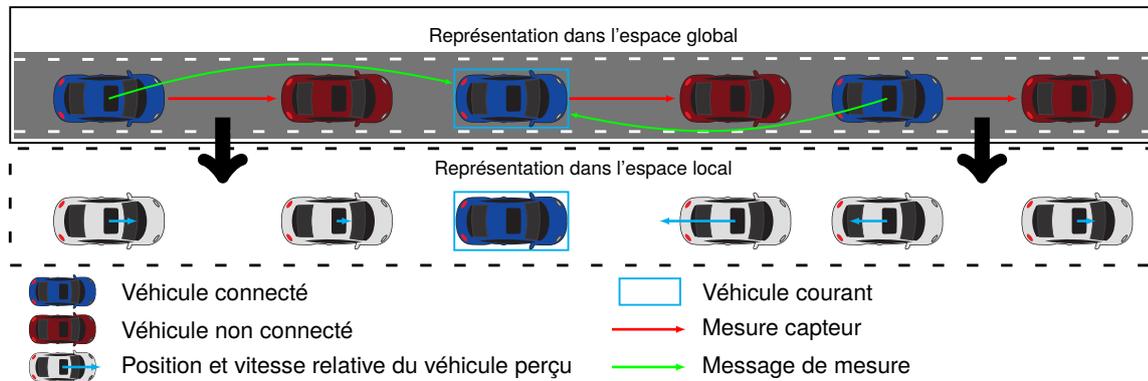


Fig. 3.8: Passage du référentiel global au référentiel local du véhicule.

Les positions et vitesses perçues directement (capteur) ou indirectement (messages) sont transformées sous la forme d'influences (positions et vecteur vitesse relative). Ces données sont utilisées dans le calcul des coefficients de proximité du modèle multi-anticipatif bilatéral (section 2.4). Les messages de confiance interviennent enfin dans la mise à jour du réseau de confiance. Durant le processus de décision, un véhicule peut ainsi décider de stopper toute forme de communication si son capteur est jugé trop peu fiable par les autres.

Re-propagation des messages

L'autre partie de la prise de décision des véhicules concerne la re-propagation (figure 3.7) des messages (confiance et consigne). Ce traitement est réalisé par des règles simples permettant de déterminer si un message est pertinent (*i.e.* non périmé et si le véhicule se trouve dans la zone de pertinence) ou non. Les messages qui sont déterminés comme pertinents disparaissent de la boîte de réception et sont placés dans une autre : la boîte d'envoi. Ces messages seront ensuite échangés par les véhicules présents dans le rayon de communication du véhicule courant lors de la prochaine phase de perception.

Changement de voie

Dans la version originale de MovSim, les véhicules appliquent d'abord le modèle de changement de voie avant de mettre à jour leur accélération, en exécutant leur loi de poursuite. Cela permet d'assurer une continuité dans le comportement longitudinal des véhicules. Dans le cadre de l'extension, le processus d'exécution des véhicules est modifié. La partie "décision" du modèle de changement de voie fait désormais partie de la prise de décision du véhicule et l'exécution finale (qui prend en compte le caractère réalisable de la décision) est réalisée lors de la phase d'action. Cette modification permet de surcharger le modèle de changement de voie. C'est notamment le cas lors de la réception par un véhicule connecté d'une consigne I2V de changement de voie. Le modèle original (stratégie MOBIL [104]) s'appuie, dans son implémentation, sur la distinction entre les

changements de voie opportunistes et obligatoires. Les changements obligatoires sont déterminés par la topologie du réseau (par exemple, un véhicule qui se trouve sur une voie d'insertion doit rejoindre la voie située à sa gauche). Les changements opportunistes participent à la volonté de dépassement d'un véhicule. Le modèle estime si un changement de voie est bénéfique en fonction de la perception actuelle du véhicule et de ses voisins. Dans MASCAT, un niveau intermédiaire permet d'intégrer les changements de voie qui sont motivés par des consignes I2V reçues par les véhicules connectés. L'ordre d'exécution des types de changement de voie est ainsi soumis à l'ordre de priorité : obligatoire, consigne, opportuniste. La réalisation du changement de voie (*i.e.* mouvement latéral) est traitée dans la phase d'action de l'agent.

3.2.5 Action

La phase finale de l'exécution du cycle de vie des agents dans le simulateur consiste à appliquer les actions décidées par les entités face à leur nouvelle perception de leur environnement. Cette phase aboutit à l'exécution d'une action, qui peut prendre la forme d'un mouvement (agents mobiles comme les véhicules) ou mener à l'émission de messages (UBR et véhicules). Cette phase est cruciale pour garantir un réalisme et une stabilité dans l'exécution des agents. En effet, la phase de décision pourrait aboutir à des conflits ou à des situations impossibles (accélération non bornée, changement de voie impossible, etc.).

3.2.5.1 Infrastructure

L'action d'une UBR ayant pour rôle le contrôle d'une section consiste en l'envoi d'un message de consigne aux véhicules connectés à portée (message qui est propagé dans sa zone de pertinence). Ces messages décrivent un type de consigne et un paramètre associé. La liste des consignes implémentées dans MASCAT est la suivante :

- **Consigne de changement de voie** (paramètre : *voie à éviter*) : indique aux véhicules destinataires du message une voie de la section courante à éviter (accident, fermeture de voie, facilitation d'une insertion, etc.). Cette consigne permet aux véhicules d'anticiper un événement ponctuel ou récurrent perçu par l'UBR et ayant une incidence sur le réseau.
- **Consigne d'inter-distance** (paramètre : *inter-distance cible*) : demande aux véhicules connectés de réduire ou d'augmenter leur inter-distance avec le véhicule prédécesseur (le cas échéant). L'objectif est de faire varier la densité du trafic, ce qui permet de faciliter ou de réduire la fréquence des changements de voie au sein du flux.
- **Consigne de vitesse** (paramètre : *vitesse limite*) : à la manière des limites de vitesse variables (VSL) qui pourraient être indiquées par des Panneaux à Message Variable (PMV), cette consigne change dynamiquement la vitesse maximum autorisée sur la section. Une telle consigne, appliquée partiellement par une fraction de véhicules connectés, vise à éloigner le flux d'un régime critique menant à l'apparition de congestion (principe de la régulation dynamique des vitesses).

3.2.5.2 Véhicules

Les véhicules exécutent un comportement longitudinal modélisé sous la forme d'une loi de poursuite. Le modèle retenu pour ce travail est *Intelligent Driver Model* [192] (équation 2.4). Ces modèles disposent de paramètres qui permettent de représenter la variabilité des comportements de conduite.

Interprétation des consignes

L'interprétation des consignes de vitesse et d'inter-distance par les véhicules consiste en une modification temporaire (durant la durée de pertinence du message de consigne) des paramètres du modèle longitudinal. Lorsqu'un véhicule reçoit un message de consigne de vitesse, le paramètre représentant la vitesse désirée (paramètre v_0 de IDM) est immédiatement mis à jour par la valeur de consigne. Le véhicule reste contraint par les conditions de circulation si l'objectif était d'accélérer, mais diminue progressivement sa vitesse si la consigne est plus faible que sa vitesse courante. Cette modification fonctionne sur le même principe que les limitations de vitesses modélisées en bord de route (signalisation verticale).

Si une consigne d'inter-distance est reçue, le véhicule met à jour la valeur du paramètre correspondant de son modèle longitudinal (noté T dans l'équation 2.4). Le modèle est capable d'ajuster dynamiquement l'inter-distance effective sans causer d'instabilité, puisqu'un tel changement est interprété comme une modification de la position du véhicule prédécesseur.

Les messages de consigne de changement de voie indiquent aux véhicules une voie à éviter. Les véhicules recevant ce message vont surcharger leur comportement de changement de voie pour intégrer cette consigne. Pour cela, nous avons ajouté un nouveau type de changement de voie dont le niveau de priorité se situe entre les changements opportunistes et obligatoires. Ainsi, le véhicule change de voie pour quitter la voie ciblée tant que les conditions le permettent et tant qu'aucun changement obligatoire n'est imposé. Afin de rendre l'action la plus efficace possible, les véhicules ne se contentent pas de quitter la voie à éviter, mais ils tentent de s'en éloigner autant que possible (par exemple, sur une route à 3 voies, si la voie de gauche est à éviter, les véhicules vont tenter de rejoindre la voie la plus à droite).

Application des modèles

La version originale de MovSim propose déjà un calcul fin des positions et des vitesses des véhicules en fonction de leurs paramètres actuels et de leurs modèles. Le mouvement latéral (changement de voie) consiste en un changement de la voie de circulation. Celui-ci prend effet immédiatement mais est graphiquement interpolé lors de la projection sur le réseau en 2 dimensions (vue graphique du simulateur). Le calcul de la position de chaque véhicule est le résultat de l'interpolation de la vitesse, calculée à partir de l'accélération issue du modèle longitudinal. Le simulateur fonctionne en pas de temps discrets durant lesquels l'intégration du processus en trois étapes perception-décision-action permet de garantir l'autonomie des agents et d'assurer la cohérence des mouvements résultant des actions choisies par les agents lors de leur processus de décision.

3.2.6 Autres améliorations du simulateur et perspectives

En plus des modèles et de l'architecture multi-agent ajoutés au simulateur, d'autres développements ont été réalisés et implémentés dans MASCAT. À des fins d'étude et d'analyse, les fichiers d'entrées ont été étendus pour permettre l'utilisation des nouvelles fonctionnalités, et les fichiers de sortie permettent désormais de récupérer plus de données de simulation. L'interface initiale de MovSim a aussi été mise à jour pour visualiser le trafic coopératif. D'autres développements sont aussi envisagés.

3.2.6.1 Entrées

La définition des scénarios dans la version originale de MovSim se base sur l'édition de deux fichiers de configuration au format XML. Le premier fichier est dédié à la définition du réseau ; il peut donc être commun à plusieurs scénarios. Il reprend le formalisme OpenDrive [56] qui permet de décorrélérer la représentation logique (sous la forme d'un graphe orienté composé de sections décomposées en voies de circulation) de la représentation graphique (projection en 2 dimensions du graphe logique) du réseau. Le standard permet de modéliser toute forme de réseau, plus ou moins complexe, mais il est nécessaire d'y représenter toutes les connexions possibles (ce qui complique la tâche de modélisation dans le cas de réseaux fortement maillés). Le second fichier décrit le scénario en tant que tel, et permet de choisir un grand nombre de paramètres, parmi lesquels :

- les types de véhicules, leurs modèles (longitudinal, changement de voie, consommation) ;
- les paramètres généraux du scénario (durée, pas de temps, etc.) ;
- les caractéristiques des points d'entrée du réseau (débits entrants, composition, configuration initiale) ;
- les sorties du scénario (traces de véhicules, consommations, capteurs du réseau, etc.).

Suite aux ajouts de notre extension, d'autres fonctionnalités peuvent être activées directement depuis les fichiers de configuration des scénarios. La figure 3.9 met en valeur (en couleurs) les ajouts disponibles dans les fichiers de configuration.

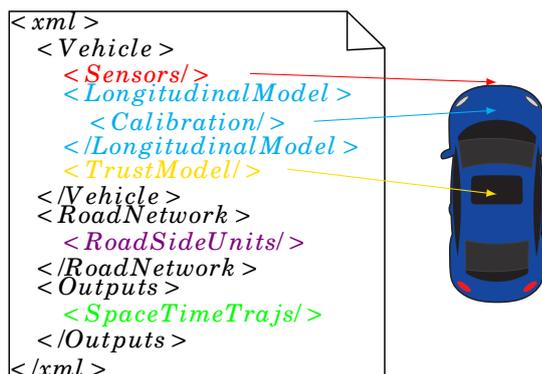


Fig. 3.9: Illustration des paramètres des scénarios dans MovSim (en noir) et son extension MASCAT (en couleurs).

Parmi les nouveaux paramètres définis dans MASCAT, il est possible de choisir :

- les capteurs des véhicules (et leurs paramètres : portée, fréquence, fiabilité) ;

- le fichier de paramètres calés pour les modèles longitudinaux ;
- le modèle de confiance utilisé (parmi ceux implémentés) ;
- le type et l'emplacement des unités d'infrastructure (UBR), leurs capteurs et paramètres associés ;
- les nouveaux fichiers de sortie (trajectoires, confiance, contrôle).

Ces ajouts permettent d'instancier les modèles développés dans notre travail directement depuis les fichiers de configuration. Cela assure aussi une base de comparaison saine entre les modèles coopératifs et non-coopératifs, qui peuvent être testés sur le même réseau, et dans les mêmes conditions de simulation (débits d'entrée, concentration initiale, etc.).

3.2.6.2 Sorties

MovSim propose dans son implémentation une gestion des fichiers de sortie générique qui permet de générer des fichiers au format CSV (fichier texte dont les colonnes sont séparées par un caractère marqueur : "virgule", "point-virgule", etc.). Ces fichiers ordonnent les données de sortie du simulateur en colonne, et chaque ligne représente un pas de temps. Le pas de temps des fichiers de sortie peut être décorrélé du pas de temps de simulation (donc plus grand mais pas plus petit), ce qui permet de pré-traiter les données (agrégations, etc.). Ces données sont cruciales à des fins d'analyse et d'étude des résultats obtenus par les modèles proposés dans notre travail. En ce sens, les nouveaux comportements et les informations échangées par les véhicules coopératifs nécessitent de nouveaux indicateurs. L'interface de MovSim a aussi été étendue pour visualiser l'évolution des véhicules coopératifs et l'infrastructure.

Interface

En reprenant le code de couleurs défini dans la partie modélisation (chapitre 2) de nos travaux, nous avons proposé une extension de l'interface de MovSim. Comme l'illustre la figure 3.10, il est désormais possible de distinguer les véhicules équipés des véhicules classiques, de visualiser l'infrastructure, ou encore de représenter les rayons de communication.



Fig. 3.10: Capture d'écran de la vue trafic adaptée pour le trafic coopératif dans MASCAT. Le code couleur est le même que celui de la figure 3.6. Le rectangle jaune correspond à la position du véhicule "virtuel", représentation de l'application des coefficients de proximité dans le modèle bilatéral multi-anticipatif. Les marqueurs verts sont les véhicules perçus par le véhicule courant (en cyan) dans cette configuration.

À des fins de vérification et de développement de la stratégie locale d'homogénéisation (modèle BMA), il est aussi possible de sélectionner un véhicule, afin d'afficher les données de sa perception

et de sa décision actuelle. L'interface affiche alors la position des véhicules perçus, ainsi que la position du véhicule virtuel, résultat des coefficients d'interactions du modèle. Cette position permet une représentation graphique de l'effet d'anticipation qui permet au véhicule de mieux adapter son accélération grâce à son champ de perception étendu.

Diagrammes espace-temps

Notre travail se concentre sur la définition de stratégies de contrôle (locale ou décentralisée), où les interactions individuelles des agents en coopération provoquent des phénomènes émergents au niveau du flux. Bien qu'ils puissent s'accompagner de bénéfices à l'échelle de l'individu (réduction du temps de trajet, augmentation du confort), les objectifs de ces systèmes coopératifs sont souvent plus globaux. Très utilisés pour visualiser les données d'un grand nombre de véhicules évoluant sur une même section, les diagrammes espace-temps représentent toutes les trajectoires au cours du temps. Nous avons donc choisi de générer dynamiquement ces diagrammes, en proposant en supplément l'affichage des vitesses des véhicules. Un exemple d'un diagramme espace-temps généré dans MASCAT est disponible dans la figure 3.11.

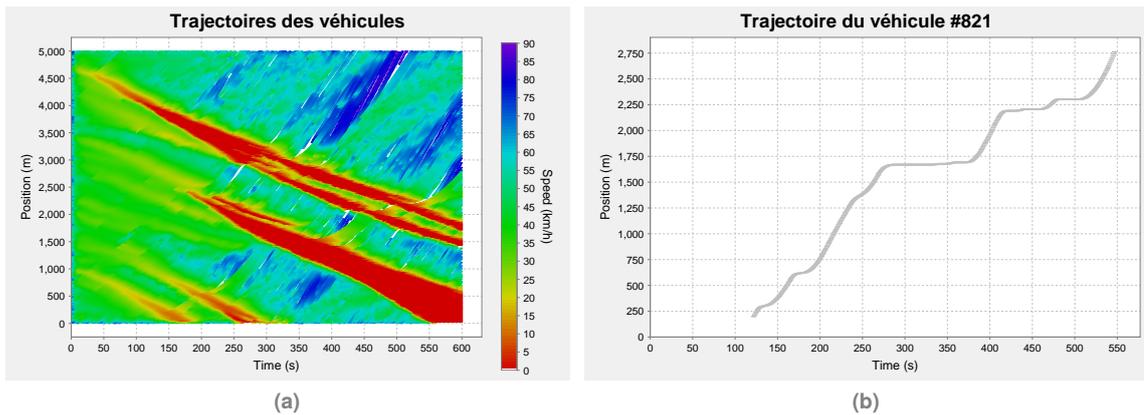


Fig. 3.11: Exemple d'un tracé des trajectoires sous la forme d'un diagramme espace-temps-vitesses dans MASCAT. Le diagramme (a) représente les traces de tous les véhicules sur une voie, et le diagramme (b) permet de mieux comprendre le type de trajectoire observé en affichant la trace d'un seul véhicule.

Ce tracé est rendu possible par la définition d'un nouveau fichier de sortie représentant les trajectoires. Le fichier regroupe différents paramètres pour chaque colonne (temps, identifiant du véhicule, vitesse, position, voie ...); et chaque ligne représente un point de la trace d'un véhicule. Grâce aux identifiants, il est possible de reconstruire la trajectoire de chaque véhicule. Le diagramme espace-temps est un tracé simultané de toutes les trajectoires des véhicules circulant sur une section et une voie donnée. Comme le montre le diagramme 3.11a, les différences de vitesses dans la chaîne de véhicules mettent en évidence l'apparition et la propagation de perturbation. La congestion (ampleur, durée et évolution) est lisible sur le diagramme en observant les zones où les vitesses sont très faibles (e.g. inférieures à 30 km/h pour un scénario autoroutier).

Indicateurs

La définition et/ou le choix d'indicateurs est préalable à l'évaluation de l'impact de l'introduction des véhicules connectés dans un flux de trafic. Ce thème a fait partie de notre réflexion, et sera présenté dans la partie 3.3. Les indicateurs permettent de générer des données, le plus souvent agrégées, capables de mettre en lumière une caractéristique particulière présentée par le système au cours de son évolution. Notre travail sur les indicateurs a nécessité de créer un nouveau type de fichier de sortie, qu'il est possible de paramétrer dans les fichiers de configuration. Ces fichiers contiennent des données recueillies dans le simulateur suivant un pas de temps d'agrégation défini par l'utilisateur. Ils peuvent ainsi être utilisés pour mener des analyses (sensibilité, impact) après avoir simulé un ou plusieurs scénarios.

3.2.6.3 Perspectives

MovSim est une plateforme dont le développement est très récent et certaines fonctionnalités sont encore en développement. Une amélioration majeure qui sera bientôt proposée portera sur une version bidirectionnelle du trafic. En effet, pour le moment, les sections modélisées sont des sens unique. Ce qui correspond bien à notre besoin, puisque le scénario autoroutier correspond à des cas où il est possible d'améliorer substantiellement la situation de trafic avec les systèmes coopératifs. L'ajout de la version bidirectionnelle ouvrira la voie vers une étude plus poussée des modèles de changement de voie (avec la reproduction de comportement de dépassement, par exemple). Dans le cadre des C-ITS, l'intérêt majeur est de pouvoir utiliser les véhicules connectés évoluant en sens inverse pour propager des informations critiques ou relayer des consignes. La manœuvre de dépassement pourrait aussi être effectuée avec plus de sécurité en augmentant le champ de perception des véhicules (*e.g.* qui tenteraient de dépasser un véhicule de grand gabarit).

Une autre perspective de notre travail de développement réside dans sa valorisation et son partage. Nous souhaitons dans un avenir proche proposer une version stabilisée de nos contributions en tant qu'extension multi-agent de MovSim (MASCAT). En restant sur la volonté de rendre les sources libres (Open-Source), la meilleure façon de contribuer serait de proposer les modifications sur le dépôt Git officiel de MovSim². Compte tenu de l'ampleur de l'extension, qui touche une partie du cœur du simulateur, nos développements risquent de ne pas s'intégrer directement à la version courante de MovSim, mais ils pourraient contribuer à une version multi-agent stable et dédiée à la modélisation des systèmes coopératifs.

Côté valorisation, MASCAT a été utilisé pour illustrer les problématiques des C-ITS dans plusieurs cours d'Outils pour la Mobilité Intelligente à l'ENTPE (École Nationale des Travaux Publics de l'État). Le simulateur permet en effet de reproduire des instabilités et d'en chercher la cause, afin de montrer le potentiel des systèmes coopératifs, même à un public non averti. Le frein majeur à ce type d'utilisation est l'accessibilité du logiciel, qui pour le moment est limitée à l'interface et aux fichiers de configuration. Une amélioration intéressante en ce sens pourrait être de proposer une interface d'édition des scénarios, permettant à un utilisateur non expert, grâce à des outils intégrés, de définir ses propres scénarios. En développant davantage cette idée, il est possible d'imaginer une interface générique permettant de définir de nouveaux types de messages, et de décorrélérer la prise de décision des agents (implémentée par l'utilisateur) dans un module externe au simulateur.

2. Dépôt Git officiel de MovSim : <https://github.com/movsim/movsim>

Le simulateur pourrait alors être considéré comme un outil d'aide à la décision capable d'orienter la réflexion des acteurs majeurs des C-ITS. Les questions et attentes des gestionnaires de réseau et des décideurs étant souvent liées aux impacts des systèmes et au dimensionnement des moyens (infrastructure, capteurs, etc.). En complément d'expérimentations réelles, le simulateur peut aussi être utilisé pour évaluer les effets de prototypes dans des conditions difficiles à maîtriser (météo, pannes, fiabilité).

3.3 Indicateurs de performance et d'évaluation

L'un des objectifs de notre travail est de montrer le potentiel des C-ITS en s'appuyant principalement sur des études en simulation. Grâce au simulateur MASCAT, nous avons proposé un cadre stable permettant d'étudier les effets de différentes stratégies de contrôle (local, et décentralisé) de ces nouveaux véhicules coopératifs. Afin d'orienter notre réflexion et pour obtenir des données tangibles relatives aux effets de ces systèmes, il est important de proposer une série d'indicateurs capables de quantifier les bénéfices attendus ainsi que leurs limitations. Nous faisons l'hypothèse que les effets de tels systèmes complexes sont souvent le fruit d'un compromis entre différents objectifs globaux, qui sont satisfaits par des entités ayant leurs propres objectifs. Nous proposons donc des indicateurs permettant d'évaluer les bénéfices de la coopération entre véhicules et faisant écho aux enjeux principaux des C-ITS : impact sur le flux trafic, la sécurité et les consommations. Les indicateurs sont calculés dans le simulateur en suivant un pas d'agrégation distinct du pas de temps de simulation. Toutes les valeurs observées résultent ainsi d'une moyenne des valeurs calculées à chaque pas de temps de simulation durant le pas d'agrégation. Ces indicateurs font l'objet d'un fichier de sortie dédié qui reprend les paramètres fixés dans le fichier de configuration.

3.3.1 Écoulement du flux

Les indicateurs qui caractérisent le flux de trafic permettent d'évaluer la qualité de l'écoulement. Dans MASCAT, nous avons implémenté les indicateurs suivants :

- **Vitesse moyenne V** : la vitesse moyenne de tous les véhicules sur le dernier pas de temps de simulation.
- **Débit sortant moyen Q_{out}** : le débit moyen mesuré par un capteur (de type boucle électromagnétique) situé à la fin de la section.
- **Temps Total Passé – Total Time Spent TTS** : la somme des temps passés individuels ts_i sur la section de tous les N véhicules présent. Ainsi :

$$TTS = \sum_{i=1}^N ts_i \quad (3.2)$$

- **Capteurs en congestion $det_{congestion}$** : le pourcentage de capteurs (type boucle électromagnétique) où a été observée de la congestion sur le dernier pas de temps. En simulation, les capteurs peuvent être virtuels donc répartis de manière homogène sur la section, mais cet indicateur dépendrait fortement de la répartition pour une application avec des données réelles.

- **Véhicules en congestion** $sec_{congestion}$: le pourcentage de véhicules dont la vitesse était inférieure à 30 km/h durant le dernier pas de temps (nécessite d'avoir accès aux traces – trajectoires – des véhicules).
- **Hétérogénéité du groupe – Group Disagreement Value** G_d : cet indicateur est lié à l'objectif d'obtenir un trafic fluide grâce à la coopération entre les véhicules. Adapté des travaux de Choi *et al.* [34], il permet d'évaluer l'homogénéité d'un groupe d'agents qui s'exprime :

$$G_d = \sum_{(i,j) \in N^2} b_{ij} (\dot{x}_j - \dot{x}_i)^2, \quad (3.3)$$

où b_{ij} est une valeur binaire d'adjacence (0 si les véhicules ne sont pas dans leur rayon de communication, 1 sinon), et N désigne le nombre de véhicules sur la section. L'indicateur G_d permet d'évaluer l'homogénéité d'un flux sur une section lorsque le nombre de véhicules est sensiblement semblable.

3.3.2 Sécurité

Les indicateurs permettant d'évaluer la sécurité sont plus largement représentés dans la littérature. MASCAT permet de calculer les indicateurs suivants :

- **Temps avant collision – Time To Collision** TTC : il s'agit du temps estimé avant qu'une collision entre un véhicule et son prédécesseur ne survienne en faisant l'hypothèse que leurs vitesses et trajectoires restent constantes [87]. Le TTC d'un véhicule i à l'instant t suivant le véhicule prédécesseur $i + 1$ s'écrit [135] :

$$TTC_i(t) = \frac{x_{i+1}(t) - x_i(t) - l_i}{\dot{x}_{i+1}(t) - \dot{x}_i(t)} \quad (3.4)$$

où x désigne la position, \dot{x} la vitesse et l la longueur du véhicule. En reprenant nos notations (section 2.1.1), le TTC peut s'écrire :

$$TTC_i(t) = \frac{\Delta x_{i+1}}{\Delta \dot{x}_{i+1}} \quad (3.5)$$

où $\Delta x_{(i,i+1)}$ et $\Delta \dot{x}_{(i,i+1)}$ désignent respectivement l'inter-distance entre les véhicules i et $i + 1$ et leur vitesse relative.

- **Temps exposé au TTC – Time Exposed to TTC** TET : la définition du TTC le rend très sensible aux faibles différences de vitesses. Cet indicateur est souvent remplacé par le TET qui consiste à estimer le temps pendant lequel un véhicule circule avec une valeur de TTC plus faible qu'un certain seuil [135] (nous considérons un seuil de 3 secondes, comme recommandé par les auteurs).
- **Temps inter-véhiculaire – Post-Encroachment Time** PET : il est défini comme la différence temporelle entre deux véhicules passant par une même zone spatiale. Le PET représente ainsi un risque direct de collision (qui survient lorsqu'il prend la valeur zéro). Le PET , s'écrit [135, 75] :

$$PET_i(t) = \frac{\Delta x_{i+1}}{\dot{x}_i} \quad (3.6)$$

- **Fréquence des changement de voie** : il s'agit du pourcentage de véhicules qui ont effectué un changement de voie sur la période d'observation.

3.3.3 Consommation et émissions

L'enjeu environnemental est un des points les plus critiques dans le déploiement futur des systèmes coopératifs. Un meilleur usage de l'infrastructure routière doit aussi passer par une réduction des émissions de polluants, et donc une réduction des consommations à l'échelle du flux. Beaucoup de modèles existent dans la littérature, mais le problème principal réside dans leur validation. Dans le cas des modèles microscopiques, on parle de modèles d'émission/consommation instantanés : *i.e.* ils estiment la consommation d'un véhicule à un moment donné en tenant compte de son accélération. Il faut garder à l'esprit que de tels modèles sont validés pour un véhicule donné, qui aura ses propres caractéristiques mécaniques (boîte de vitesse, régime moteur) et subira des conditions différentes (température, qualité du carburant, vent). Notre objectif n'est pas d'estimer de manière fine le gain attendu mais de donner une tendance concernant l'amélioration ou non attendue par les véhicules connectés. Dans ce cadre, nous avons choisi de faire reposer notre étude sur un modèle proposé dans l'implémentation originale de MovSim [196]. Il s'agit d'un modèle basé sur une représentation physique. Les valeurs de vitesse et d'accélération sont discrétisées et mises en correspondance avec les ratios de la boîte de vitesses. Cela permet de calculer le régime moteur (quel que soit le type de transmission). Ce régime moteur et la valeur d'accélération permettent de calculer la consommation instantanée du véhicule. Notre étude permet une estimation de la consommation au niveau du flux :

- **Consommation moyenne – Average Fuel Consumption AFC** ;
- **Consommation cumulée – Cumulated Fuel Consumption CFC**.

Nous faisons l'hypothèse que tous les véhicules sont identiques en termes de consommation lorsque cet indicateur est utilisé. Bien que l'hypothèse soit forte, elle reste compatible avec l'objectif de comparaison à l'échelle d'un flux, où seules des valeurs moyennes à grande échelle sont utilisées.

La batterie d'indicateurs proposée permet de comparer des situations analogues, où l'accent est porté sur les différentes phases du déploiement des systèmes coopératifs. En d'autres termes, il est possible de comparer deux scénarios en simulation présentant les mêmes conditions (débits entrants, topologie du réseau) mais avec des flux composés de proportions de véhicules connectés variables. Les indicateurs permettent de pointer les bénéfices et les limitations des approches proposées en évaluant les résultats obtenus au regard des enjeux principaux des C-ITS.

3.4 Applications : stratégie locale d'homogénéisation

Dans cette section, nous proposons une implémentation du modèle multi-anticipatif bilatéral (BMA), qui permet à un véhicule, grâce aux technologies de la communication et à un capteur embarqué, d'exhiber un comportement coopératif allant dans le sens de l'homogénéisation du flux. Cette stratégie de régulation, appliquée localement mais dont l'effet est multiplié par les véhicules l'appliquant au sein du flux, est une illustration du couplage des dynamiques du système, comme présenté dans la section 2.3. Grâce à la plateforme de simulation et ses développements récents (section 3.2), nous disposons d'une plateforme robuste et réaliste qui met à profit les concepts multi-agents pour modéliser finement les nouveaux types d'interaction dans les systèmes coopératifs. L'implémentation du modèle BMA a consisté en l'ajout d'un nouveau modèle se basant sur l'implémentation originale d'IDM, proposé dans la version originale de MovSim (mais dont les paramètres sont calés à partir de données réelles). Les expérimentations sont menées sur un même scénario (section 3.4.1). Dans un

premier temps, nous étudions l'effet d'homogénéisation (section 3.4.2) apporté par le modèle pour des taux de véhicules coopératifs variables. L'effet de la couche de confiance et du modèle associé sur l'évaluation de la fiabilité des capteurs des véhicules est ensuite examiné (section 3.4.3).

3.4.1 Scénario

Le scénario sélectionné et implémenté dans MASCAT s'inscrit dans l'optique d'un déploiement progressif des véhicules connectés. Le modèle BMA est implémenté comme un système avancé d'assistance au conducteur (ADAS), qui délègue une partie de la tâche de conduite du conducteur à un système semi-automatisé. Les véhicules connectés sont équipés d'un dispositif de communication ayant une portée maximale de 150 m (de rayon). Le trafic simulé consiste en un trafic mixte (*i.e.* composé à la fois de véhicules connectés et non connectés) qui évoluent sur une portion rectiligne de 10 km d'une autoroute péri-urbaine à trois voies de circulation pendant 30 minutes. La densité sur la section est initialisée à 40 véhicules/km qui circulent à 90 km/h, ce qui correspond à une situation de trafic dense où la congestion est susceptible de se former et de se propager dans le flux. Le débit d'entrée est maintenu à 1800 véhicules/heure/voie pour s'assurer que la congestion ne se réduira pas à cause d'une diminution de la demande. La figure 3.12 montre la vue graphique du scénario durant une simulation.



Fig. 3.12: Capture d'écran du simulateur avec le scénario de l'autoroute péri-urbaine à 3 voies pour un taux de pénétration des véhicules connectés de 20 % ($t = 3$ min 25 s). Le trafic évolue de gauche à droite et la section est modélisée comme continue (le saut de ligne est purement graphique). Les véhicules sont représentés par leurs vitesses, en gradient de couleur : vitesse nulle – noir ; faible vitesse – rouge ; vitesse élevée – bleu.

Tous les paramètres du modèle physique sont générés aléatoirement depuis un jeu de paramètres calibrés (nous invitons le lecteur à consulter [137] où la méthodologie de calage détaillée et le jeu de données initial sont présentés, accompagnés d'une analyse statistique de la distribution des paramètres). Les véhicules connectés sont équipés d'un capteur de type télémètre laser capable de

percevoir l'inter-distance et la distance relative à leur prédécesseur avec une précision de 5 %. Les objectifs des simulations sont d'évaluer l'influence du taux de véhicules connectés et celui de la fiabilité de l'information échangée. Nous avons mis en place deux expériences (sous-scénarios). La première vise à étudier l'influence de la stratégie coopérative sur l'homogénéité du trafic. La seconde a pour but la validation de l'utilisation de la confiance pour évaluer la fiabilité des capteurs.

3.4.2 Effets de l'homogénéisation

Pour cette première expérimentation, l'objectif est d'évaluer et de comparer l'impact de la stratégie coopérative qui émerge des interactions régies par le modèle bilatéral multi-anticipatif. Différentes phases du déploiement des véhicules connectés sont modélisées par l'introduction progressive de ces véhicules (variation du taux de pénétration) dans des simulations successives.

3.4.2.1 Observation des trajectoires

Les résultats de ces simulations (0-20-40-60-80-100 % de véhicules connectés) sont présentés dans la figure 3.13 sous la forme de diagrammes espace-temps, c'est-à-dire en traçant les trajectoires de tous les véhicules (sur la voie la plus à droite) et en mettant en valeur leurs vitesses. Une version dynamique de ces résultats est disponible dans la démonstration³ présentée dans [81, 80].

La simulation (a) de la figure 3.13 constitue le témoin de cette expérimentation. Les véhicules sont modélisés comme des véhicules non connectés (ou non coopératifs), dont le mouvement est donné par IDM. Dans les conditions du scénario (trafic dense), des bouchons fantômes se créent. Ils ont pour origine de petites perturbations au sein du flux, provoquées par les hétérogénéités des individus (vitesses désirées), et leurs comportements (changements de voie plus ou moins agressifs). Ces perturbations peuvent, dans des conditions de trafic dense, se propager sous la forme d'une onde de congestion, qui, suivant le trafic en amont (le sens de propagation de la congestion se faisant dans le sens inverse de l'écoulement), peuvent aller jusqu'à l'arrêt des véhicules. La simulation (a) montre bien cet effet : le trafic est initialisé stable (tous les véhicules roulent à la même vitesse), mais certaines perturbations sont amplifiées et s'étendent. Le modèle bilatéral multi-anticipatif a été imaginé pour améliorer ce type de situations en permettant à des véhicules équipés et connectés d'anticiper leurs réactions face à des ralentissements provoqués en aval. Les simulations suivantes montrent deux résultats intéressants. Pour 20 % (b) et 40 % (c) de véhicules exécutant le modèle, le comportement du flux va vers une réduction du nombre de perturbations propagées mais surtout vers une baisse significative de l'ampleur de ces perturbations et de la vitesse moyenne du flux. À partir d'un certain seuil (ici, peu après 40 %), les simulations suivantes (d-e-f) montrent que le modèle est capable de complètement lisser le flux (d'où la couleur de plus en plus uniformément verte). Même si des perturbations peuvent encore survenir (visibles en jaune sur le diagramme (d)), elles sont atténuées par la chaîne de véhicules connectés. Le modèle a donc un impact significatif, et ce, même pour des faibles taux de pénétration. On observe aussi un effet seuil, à partir duquel le trafic est tellement stable que les perturbations ne se propagent plus. La valeur du seuil dépend des conditions de trafic et se situe (observé empiriquement sur d'autres

3. Démonstration vidéo disponible en ligne :

— Version anglaise (AAAI 2015) : <https://youtu.be/uv60iu5hfMw>

— Version française (JFSMA 2015) : https://youtu.be/oZf_F0k6mmk

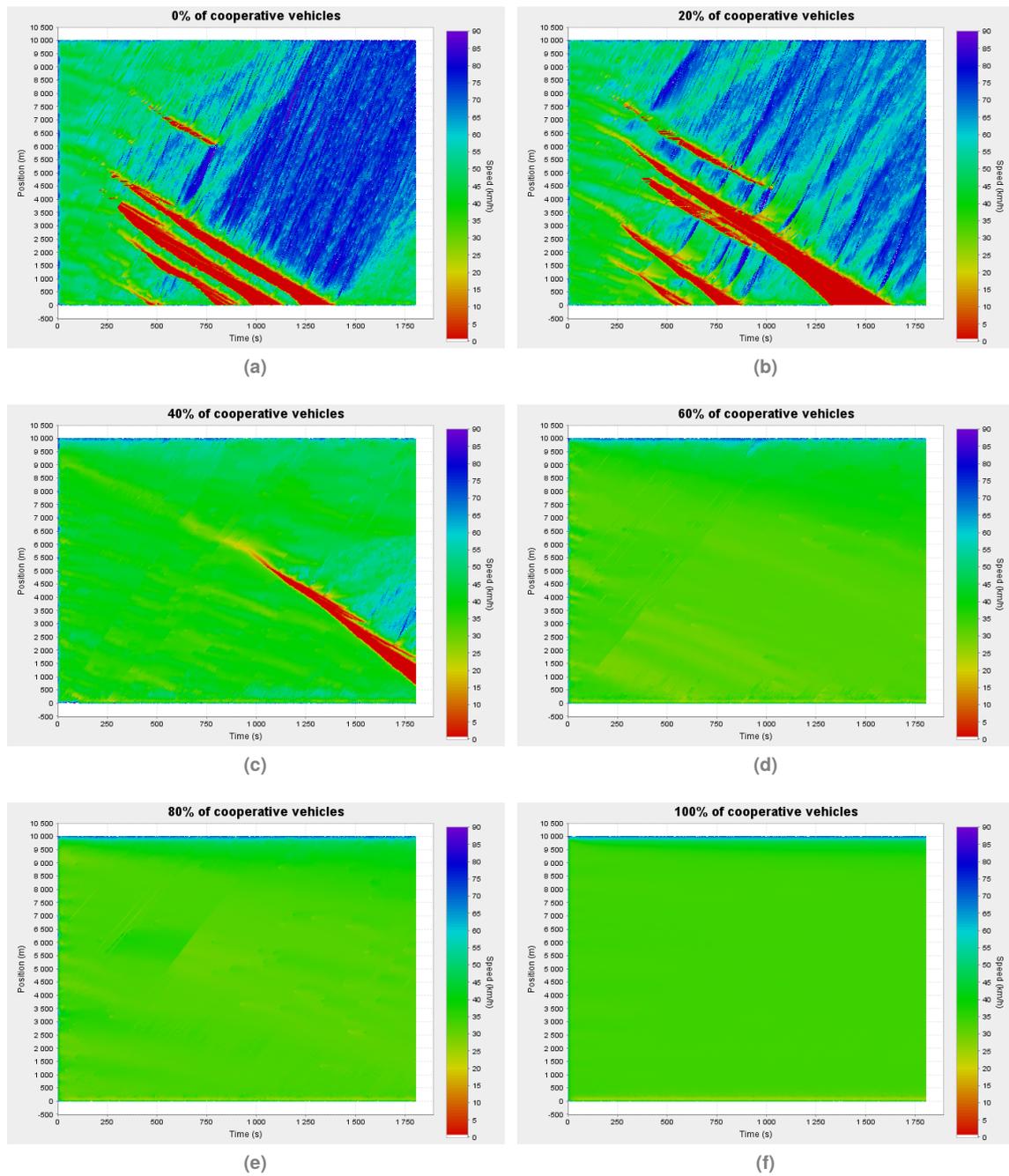


Fig. 3.13: Trajectoires et vitesses des véhicules sur la voie de droite au cours du temps pour différents taux de pénétration.

simulations) entre 30 % et 50 %. Enfin, la simulation (f), où tous les véhicules sont connectés, montre que le trafic peut devenir complètement stable malgré la forte demande. En contre partie, la vitesse moyenne du flux observée est plutôt faible, ce qui indique que les paramètres du modèle (ceux de la loi de proximité, équation 2.7) pourraient être dynamiquement ajustés en présence de forts taux de pénétration. L'effet de seuil lors de l'introduction de véhicules autonomes et connectés observé dans ces expérimentations est conforme aux observations d'autres travaux [186].

3.4.2.2 Comparaison avec le modèle ACC

Dans les expérimentations, le modèle bilatéral multi-anticipatif est utilisé par les véhicules connectés comme un système embarqué semi-automatique qui régule dynamiquement leur accélération. Ce type de contrôle existe depuis quelques années et équipe même certains modèles de série. Le système le plus répandu est le régulateur de vitesse à contrôle de distance (*Adaptive Cruise Control* – ACC). Grâce à un capteur embarqué situé à l’avant du véhicule, il est capable de percevoir la distance et la vitesse relative à son prédécesseur immédiat. Ce qui correspond exactement aux informations directement perçues par les véhicules connectés exécutant BMA. En revanche, ACC n’utilise pas d’informations partagées par la communication. L’objectif de cette simulation est de comparer, dans des situations très similaires, l’influence des deux systèmes sur le flux, et notamment l’apport de la multi-anticipation. La version originale de MovSim propose une implémentation d’un système ACC [194] (section 2.1.2.1).

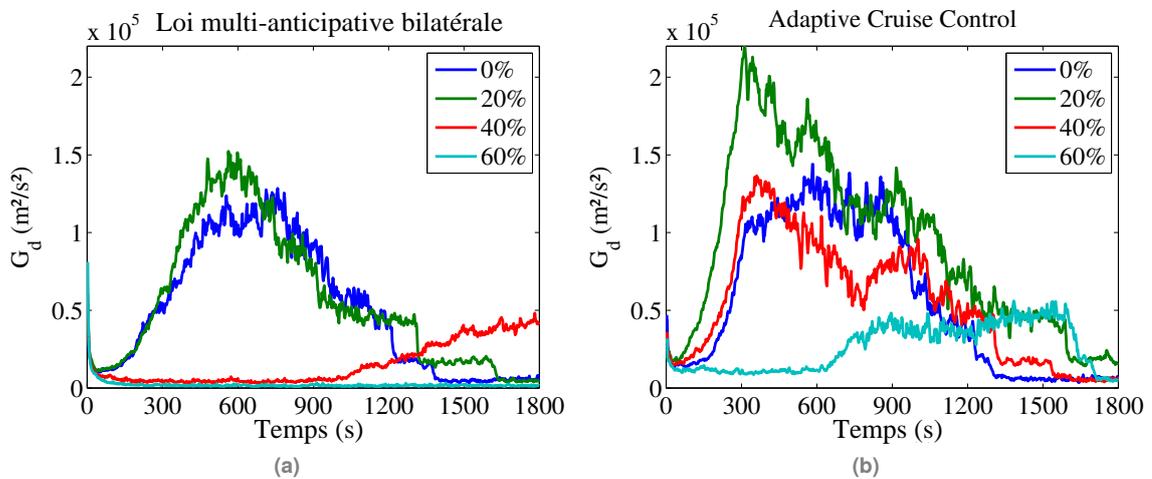


Fig. 3.14: Comparaison de la valeur d’hétérogénéité du groupe entre les modèles BMA (a) et ACC (b), pour différents taux de pénétration.

La figure 3.14 montre l’évolution de la valeur d’hétérogénéité du groupe (G_d , équation 3.3) pour différents taux de pénétration (mêmes résultats que l’expérimentation précédente) et avec les deux systèmes utilisés pour réguler l’accélération des véhicules équipés. La première observation concerne l’amplitude des hétérogénéités atteintes par les deux modèles : le modèle BMA a un impact plus important sur la réduction des perturbations pour de faibles taux de pénétration. Par exemple, le modèle retarde l’apparition de congestion avec 40 % de véhicules équipés (cf. le tracé des trajectoires de la simulation (c) de la figure 3.13). La transition entre régimes fluides et congestionnés est aussi plus lisse avec la stratégie coopérative. L’utilisation de plusieurs véhicules (en avant et en arrière) a donc un effet important sur l’homogénéisation du flux. Les résultats obtenus en simulation par cette implémentation de ACC [194] sont très comparables aux expérimentations réelles menées dans l’étude de Milanés *et al.* [133], où les limites de ACC sur la stabilité d’une chaîne de véhicules ont été montrées. L’extension coopérative de ce modèle, nommée C-ACC (*Cooperative Adaptive Cruise Control* [199]) permet d’ajouter les informations issues de la communication dans le modèle et donc d’améliorer la stabilité du flux [133]. Il conviendra de comparer BMA à C-ACC, qui se contente des véhicules situés devant le véhicule courant. L’apport de BMA réside aussi dans le couplage avec la dynamique de contrôle modélisée sous la forme d’un réseau de confiance.

Les expérimentations qui suivent se concentrent sur l'étude des effets du modèle en présence de capteurs défaillants.

3.4.3 Impact de la fiabilité des capteurs

Dans cette expérimentation, nous souhaitons mettre en évidence l'effet de la confiance sur la gestion des capteurs défaillants. En utilisant le même scénario, la simulation est modifiée afin que certains véhicules connectés soient équipés de capteurs défaillants qui fournissent des mesures entachées d'erreurs importantes (+/- 50%) pour l'inter-distance et la vitesse relative. L'objectif est de montrer l'impact de notre modèle de confiance (section 2.4.5), couplé au modèle BMA, implémenté dans le simulateur et exécuté au sein des véhicules connectés.

3.4.3.1 Impact sur le flux

Les premiers résultats visent à montrer l'effet de la couche de confiance lorsque certains véhicules sont équipés de capteurs défaillants. Le modèle BMA qui régit l'accélération des véhicules connectés repose sur le couplage des flux du système. L'influence des informations des capteurs est partagée par la communication, et les erreurs risquent d'être propagées. Notre modèle basé sur la confiance agit alors comme la dynamique de contrôle qui permet de minimiser l'impact de ces informations sur le système. Le modèle contribue ainsi à augmenter la sécurité, enjeu majeur qui permettra aux usagers d'accorder leur confiance aux systèmes coopératifs. Dans ce cadre, nous proposons d'utiliser le concept de confiance pour modéliser la fiabilité des capteurs des véhicules. Le scénario de cette expérimentation est toujours identique, mais le taux de pénétration (*i.e.* la proportion de véhicules connectés) est fixé à 60 %. Parmi les véhicules connectés, 10 % présentent un capteur défectueux (modélisé par un écart-type de 50 % contre seulement 5 % pour les capteurs désignés comme fiables). Nous comparons deux simulations. La première se déroule sans utiliser les valeurs de confiance : les valeurs sont propagées mais ne rentrent plus en compte dans le calcul des coefficients de proximité (α_{ij} , équation 2.6). La seconde est lancée avec le modèle de confiance complet, qui utilise et met à jour les différentes valeurs de confiance (directe, indirecte et confiance en soi, présentées en section 2.4.5).

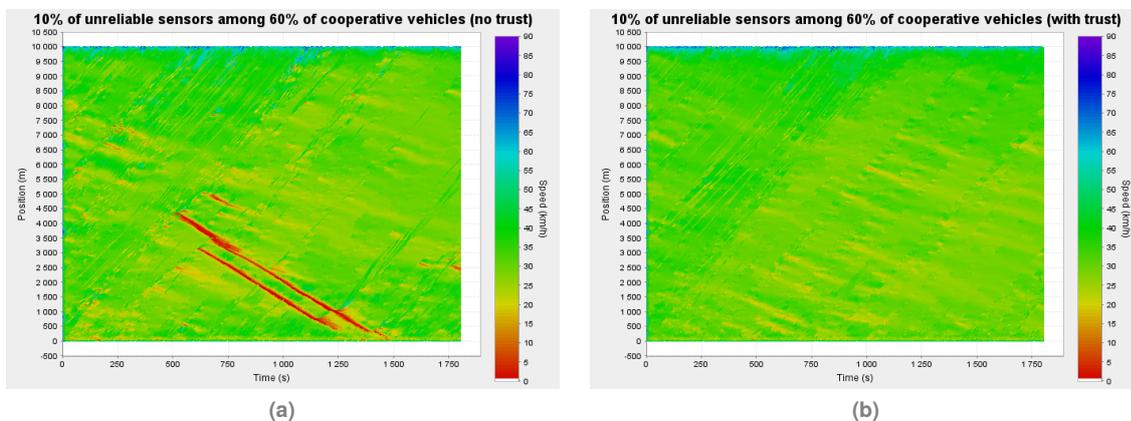


Fig. 3.15: Comparaison des trajectoires d'un flux de véhicules composé de 60 % de véhicules connectés dont 10 % sont équipés de capteurs défaillants, sans (a) et avec (b) le modèle de confiance.

Les diagrammes espace temps de la figure 3.15 montrent les trajectoires des véhicules pour un même scénario, avec et sans le modèle de confiance. Dans la simulation sans le modèle de confiance (a), les erreurs des capteurs se propagent rapidement et causent beaucoup d'instabilités dans le flux, même si seulement 10 % des véhicules connectés sont impactés. La densité du trafic le rend sensible aux perturbations induites par ces véhicules. Cela s'explique par le fait que les informations erronées sont partagées et entrent en considération dans la prise de décision (modèle BMA) des autres véhicules connectés, pourtant équipés de capteurs fiables. Dans cette configuration, il apparaît clairement que la fiabilité des capteurs est un paramètre important pour assurer la stabilité et l'homogénéité du flux. Notre modèle de confiance se présente comme une couche de contrôle permettant, grâce à son couplage avec la communication et la dynamique physique, de jouer sur l'effet de groupe et la coopération afin de détecter et de minimiser l'impact des informations non fiables à l'échelle du système. La figure 3.15 (b) montre les trajectoires obtenues lorsque le modèle de confiance est en place. Dans ces conditions de haute densité, les véhicules circulent en pelotons et sont ainsi suffisamment en contact (ou plutôt, à portée de communication) pour faire évoluer leurs valeurs de confiance. Cette mise à jour résulte en une modification et une adaptation des coefficients de proximité qui permettent au modèle BMA de minimiser l'influence des véhicules estimés comme non fiables dans le modèle multi-anticipatif. Les propriétés du modèle sont ainsi conservées (effet homogénéisation) même si les véhicules dont les capteurs sont défectueux continuent à perturber localement le flux. Cependant, le modèle coopératif, couplé au modèle de confiance, assure la stabilité du flux puisque les perturbations ne se propagent plus.

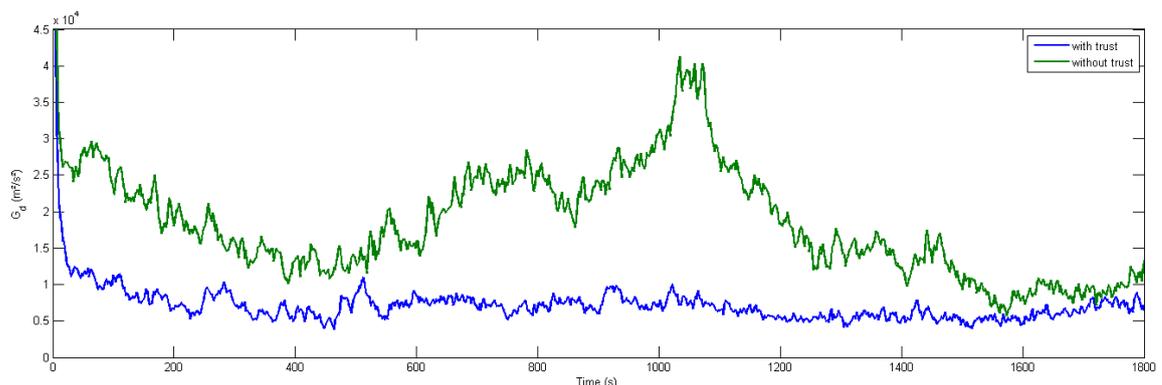


Fig. 3.16: Comparaison de l'hétérogénéité du groupe G_d d'un flux de véhicules composé de 60 % de véhicules connectés dont 10 % sont équipés de capteurs défaillants, avec et sans le modèle de confiance.

Si l'on observe maintenant l'indicateur permettant d'évaluer l'hétérogénéité du groupe (G_d , équation 3.3) dans les deux situations (figure 3.16), on observe que la stabilité du modèle de confiance n'est pas atteinte immédiatement. En effet, au début de la simulation, les véhicules sont placés sur la chaussée mais ils n'ont aucune connaissance de leur environnement (les valeurs de confiance ne sont pas initialisées). Les premières interactions permettent aux véhicules d'estimer les valeurs de confiance directe et de commencer à adapter leur propre comportement aux véhicules non fiables de leur rayon de communication. Ensuite, la propagation des valeurs de confiance indirectes dans les messages de confiance assure un recouvrement plus étendu de cette information, y compris à des véhicules qui n'ont pas encore interagi avec des véhicules non fiables.

3.4.3.2 Impact sur la sécurité

Parmi les indicateurs proposés dans la section 3.3, nous avons choisi le temps inter-véhiculaire (*PET*, équation 3.6). Cet indicateur est utilisé pour évaluer la sécurité des systèmes automatiques de régulation de vitesse à contrôle de distance [135]. Afin de valider le comportement de notre modèle de confiance, basé sur les réseaux de confiance (*TrustNets* [179]), nous avons choisi de comparer l'approche avec un autre modèle de la littérature. Le modèle retenu est FIRE [94]. Les deux modèles sont en effet comparables, puisqu'il s'agit de modèles de confiance calculatoires, adaptés à un calcul par des agents, et qui utilisent un mécanisme sans agrégation pour convertir les interactions entre agents en valeurs de confiance. Les agents qui utilisent FIRE stockent les résultats r_i (représentés sous la forme d'une notation) de leurs expériences passées dans leur mémoire (limitée). Chacun de ces résultats est l'évaluation d'une interaction entre deux agents à un temps donné, ce qui est similaire à la confiance directe de notre modèle lors de l'évaluation de l'erreur de mesure. Nous avons fixé la taille de la mémoire à 6000 résultats d'interaction (ce qui correspond à 10 minutes de temps de simulation avec un pas de temps de 0,1s). Chaque valeur de confiance est calculée à partir de la somme pondérée des expériences passées, et les poids décroissent en fonction du temps de la simulation. L'objectif étant d'accorder plus d'importance aux interactions récentes tout en gardant une mémoire des expériences passées (le module de réputation de FIRE n'est pas implémenté dans ce travail). Chaque valeur de confiance T_{ij} s'écrit donc :

$$T_{ij} = \sum_{r_i} \omega(r_i) \cdot v_i \quad (3.7)$$

où $\omega(r_i) = \exp - \frac{\Delta t(r_i)}{\lambda}$ est la fonction de pondération qui estime la pertinence ou la fiabilité associée à la notation du résultat et v_i est la valeur de la note associée au résultat r_i . En suivant la description du modèle FIRE [94], nous avons choisi $\lambda = -\frac{60}{\ln(0.5)}$, ce qui implique que $\omega(r_i) = 1$ après 0 seconde et $\omega(r_i) = 0$ après 10 minutes de simulation (ce qui correspond à la taille maximum de la mémoire de l'agent).

La figure 3.17 présente les distributions des valeurs de *PET* pour trois modèles différents : sans confiance, modèle FIRE et modèle TrustNet.

Le pourcentage de capteurs défaillants ne concerne que les véhicules connectés. Donc, si cette valeur est fixée à 25% avec une proportion de véhicules équipés de 25%, cela veut dire que tous les véhicules connectés sont équipés de capteurs défaillants. Le but pour le trafic coopératif serait d'obtenir des valeurs de *PET* homogènes et suffisamment importantes (des valeurs trop faibles indiqueraient une sécurité réduite). De haut en bas, la figure 3.17 révèle une augmentation des performances de notre modèle lorsque le pourcentage de véhicules connectés augmente. De manière plus générale, le modèle coopératif multi-anticipatif bilatéral (BMA) permet une homogénéisation des valeurs de *PET*, quel que soit le modèle de confiance employé, ainsi qu'une augmentation significative des valeurs de *PET* (ce qui traduit une sécurité accrue). Les résultats obtenus montrent que la couche de confiance renforce cet effet d'homogénéisation. Notre modèle (TrustNet) donne de meilleurs résultats que les approches avec lesquelles il est comparé. De gauche à droite, les performances de notre modèle sont d'autant plus visibles, comparées aux autres approches, que la fraction de véhicules équipés de capteurs défaillants augmente. Les résultats obtenus par FIRE dans notre architecture sont tout de même à souligner compte tenu de la relative simplicité de notre implémentation du modèle.

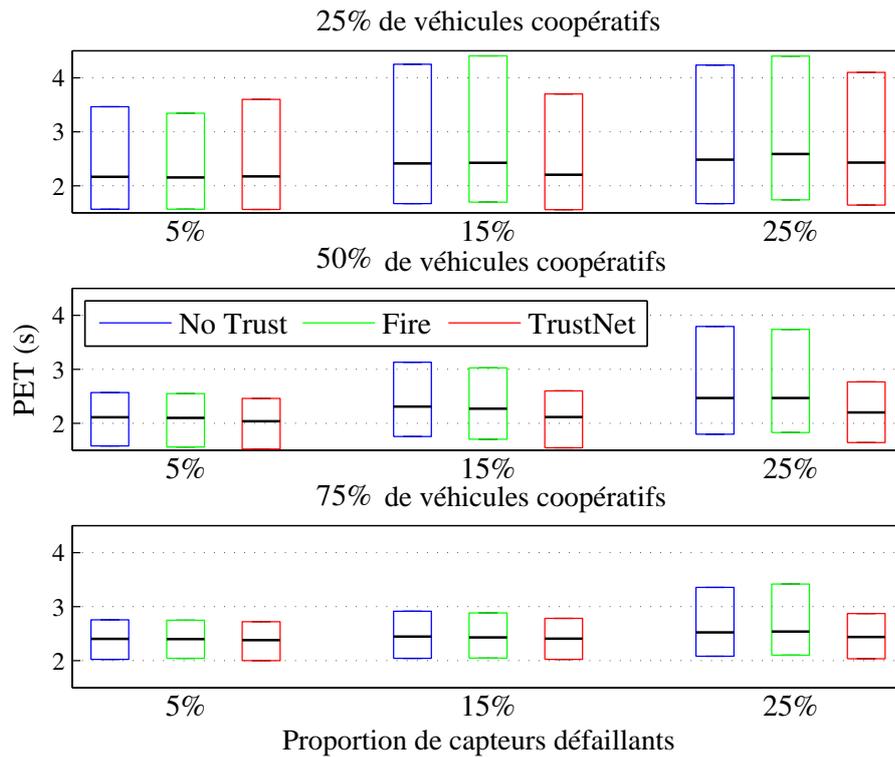


Fig. 3.17: Distributions des *PET* (médiane, 1^{er} et 9^{ème} déciles) pour trois variantes de modèles de confiance. Du haut vers le bas : augmentation du % de véhicules équipés. De gauche à droite : augmentation du % de capteurs défectueux.

3.4.3.3 Discussion

Le modèle bilatéral multi-anticipatif permet de soustraire une partie de la tâche du conducteur en déléguant l'accélération du véhicule à un système automatisé. Ce système repose à la fois sur des technologies de capteurs embarqués (perception directe) et aussi sur la communication (perception indirecte). Les résultats du modèle testé avec des taux d'équipement du flux variables, montrent plusieurs effets intéressants au niveau macroscopique. Un des effets émergents est l'homogénéisation du flux, qui comme tout phénomène transitoire dans les systèmes complexes, exhibe un seuil de percolation. Ce seuil, qui dépend des conditions de trafic, laisse entrevoir une amélioration substantielle de l'écoulement du flux grâce à la coopération, et ce dans un proche futur, correspondant à cette phase du déploiement. Bien sûr, les bénéfices apportés par la multi-anticipation (notamment face aux systèmes existants) sont très dépendants de la fiabilité de l'information, puisqu'elle est multipliée. Dans ce cadre, nous avons mis en place une stratégie de contrôle locale qui se base sur la majorité. Les véhicules mettent à jour un réseau de confiance, en échangeant leurs valeurs avec leurs voisins. Les résultats de ce modèle dans le cas où certains véhicules sont équipés de capteurs peu fiables confirment la nette diminution des interactions avec ces véhicules, ce qui contribue à l'effort d'homogénéisation du groupe.

Plusieurs critiques peuvent être apportées concernant les résultats présentés. D'abord, concernant le modèle BMA, les résultats pour des taux de pénétration importants confirment l'effet d'homogénéisation, mais cela s'accompagne par une baisse drastique des vitesses. Le lecteur averti pourrait alors se demander pourquoi ne pas simplement limiter la vitesse (e.g. via un panneau à message

variable), ce qui pourrait avoir le même potentiel. Dans toutes les simulations, nous avons gardé les mêmes paramètres pour tous les modèles, ce qui implique que l'effet d'homogénéisation est prépondérant. Ces paramètres sont plutôt destinés à des faibles taux de pénétration, c'est-à-dire à une phase précoce du déploiement. Il conviendrait d'envisager une évolution dynamique, prenant en compte le taux d'équipement local des voisins du véhicule courant, afin que son comportement s'adapte automatiquement aux conditions (*i.e.* les paramètres devraient évoluer en fonction du taux de pénétration mesuré ou estimé). Cela permettrait aussi de déjouer l'hypothèse d'une distribution uniforme des véhicules connectés dans le flux (localement, les taux de pénétration pourraient varier sensiblement).

Au sujet du modèle de confiance, bien que les résultats montrent un effet bénéfique sur la réduction des perturbations induites par des véhicules utilisant des informations non fiables, le modèle se limite à agir sur les coefficients de proximité. Un véhicule isolé recevra peu d'informations sur la confiance que les autres lui accordent et mettra plus de temps à intégrer cette information dans sa décision. Il pourrait être intéressant d'utiliser davantage la confiance en soi : un véhicule dont la confiance en soi d'effondre devrait choisir d'arrêter de communiquer des informations relatives à ces capteurs (mais il pourrait continuer à propager les valeurs de confiance, sans les modifier). Grâce à la définition du modèle BMA et de la couche confiance associée, un véhicule connecté qui n'interagirait plus avec les autres reprendrait le même comportement que les véhicules non équipés. En effet, les coefficients de pondération permettent la multi-anticipation, mais si un véhicule connecté décide de ne plus intégrer les messages provenant d'autres véhicules, il reprendra un comportement simple basé sur une loi de poursuite (comme un véhicule avec ACC activé).

Le dernier point soulevé par les expérimentations menées relève de l'observation de nouveaux comportements émergents à l'échelle du flux. Tout comme la congestion, l'homogénéisation est un effet généré par les interactions locales des véhicules. La question qui se pose alors relève du contrôle de ces effets. Les techniques actuelles de contrôle et de régulation visent à donner une consigne (panneau à message variable, etc.) à l'ensemble du flux. Or, les expérimentations ont montré que les meilleurs résultats étaient atteints en ne modifiant que les comportements d'une partie du flux. Dans l'optique d'un déploiement progressif, les stratégies actuelles de contrôle risquent d'être de moins en moins adaptées. Nous avons proposé une première forme de contrôle, basé sur la majorité des véhicules. Dans ce cadre, il n'y a pas d'autorité (les gestionnaires, par exemple) qui pourrait donner des objectifs de plus haut niveau. Il paraît difficile de faire évoluer le modèle de confiance seul, afin d'intégrer de tels objectifs dans la décision des véhicules. Face à l'évolution des comportements au sein du flux, il est donc crucial de proposer des stratégies de contrôle capables de s'adapter à l'évolution des technologies et à l'augmentation du nombre de véhicules coopératifs sur le réseau.

3.5 Vers l'élaboration de stratégies de contrôle I2V des véhicules connectés

La première étape vers l'élaboration de stratégies de contrôle des systèmes coopératifs passe une nouvelle fois par une étude en simulation. Les expérimentations focalisées sur une forme de contrôle local ont montré le potentiel d'une telle approche, mais ont aussi pointé ses limites. Les besoins en termes de contrôle émanent principalement des gestionnaires, qui bénéficient déjà d'une

grande quantité d'informations. L'architecture des C-ITS prévoit l'utilisation d'unités d'infrastructure nommées Unités de Bord de Route. Ces UBR sont des systèmes autonomes fixes qui peuvent avoir différents rôles : du simple relai d'information jusqu'à un contrôle localisé. Nous souhaitons étudier l'apport d'une stratégie de contrôle I2V basée sur la transmission de messages de consignes aux véhicules connectés. Le cas d'une section autoroutière non maillée est très bien régulée avec une stratégie locale (section 3.4). Mais lorsque le réseau devient plus complexe, le comportement du flux est plus difficile à maîtriser. Nous proposons d'étudier deux scénarios de ce type : une insertion (bretelle d'accès) et une fermeture de voie. Une version dynamique des résultats de ces scénarios est disponible dans la démonstration⁴ présentée dans [81, 80].

Les Unités de Bord de Route sont capables de recueillir des informations à partir de leurs capteurs et de la communication pour estimer l'état de trafic actuel sur leur section dédiée afin de transmettre des messages de consignes aux véhicules connectés. Les messages diffusés par les UBR indiquent les détails liés aux événements partagés (position, rayon, durée) ainsi que le type d'événement (fermeture de voie, voie d'insertion). Tant que les informations reçues sont pertinentes pour un véhicule (*i.e.* le véhicule se trouve dans le rayon de l'événement), elles sont re-propagées en amont (section 3.2.3.2). Cette dissémination des messages permet aux véhicules connectés de recevoir les informations plus tôt, et donc d'anticiper les difficultés de circulation à venir. Dans cette expérimentation, les véhicules reçoivent une consigne de changement de voie de la part de l'UBR implantée à proximité de la difficulté. L'idée est de faire générer un comportement de changement de voie par anticipation qui permet une meilleure insertion des véhicules situés sur les voies qui posent problème. Nous faisons l'hypothèse que tous les véhicules qui reçoivent une consigne tentent de changer de voie jusqu'à atteindre la plus extrême. Les scénarios sont paramétrés pour générer de la congestion en cas d'inaction des UBR (ou lorsque qu'aucun véhicule connecté n'évolue à proximité). L'objectif est d'évaluer l'impact des consignes sur le flux de trafic.

3.5.1 Insertion

La figure 3.18 (a) décrit une situation typique de convergent où les véhicules de la voie secondaire ne parviennent pas à rejoindre la section principale sans ralentir. Ce ralentissement, allant jusqu'à l'arrêt des véhicules sur la voie d'insertion, entraîne la propagation de congestion sur les voies adjacentes (voies de droite et voie du milieu). Les véhicules connectés appliquant les consignes de l'infrastructure devraient anticiper cette situation et faciliter l'insertion des véhicules entrants.

Dans la figure 3.18 (b), le carré blanc représente l'Unité de Bord de Route, située en amont de l'insertion. L'UBR envoie en continu des messages de consigne à destination des véhicules connectés, qui les re-propagent durant toute la simulation. Les messages contiennent la position de l'insertion (point critique) et incluent une consigne de changement de voie (de la droite vers la gauche). L'application des consignes (section 3.2.5.2) consiste à appliquer un changement de voie obligatoire (quand c'est possible), à la place de la stratégie opportuniste du modèle MOBIL [104]. Le comportement des véhicules connectés consiste donc à rejoindre les voies les plus à gauche avant l'insertion, et d'y rester tant qu'ils n'ont pas franchi la perturbation. L'idée est d'inciter une forme de changement de voie coopératif, que certains conducteurs déclenchent parfois spontanément. Sur

4. Démonstration vidéo disponible en ligne :

— Version anglaise (AAAI 2015) : <https://youtu.be/uv60iu5hfMw>

— Version française (JFSMA 2015) : https://youtu.be/oZf_F0k6mmk

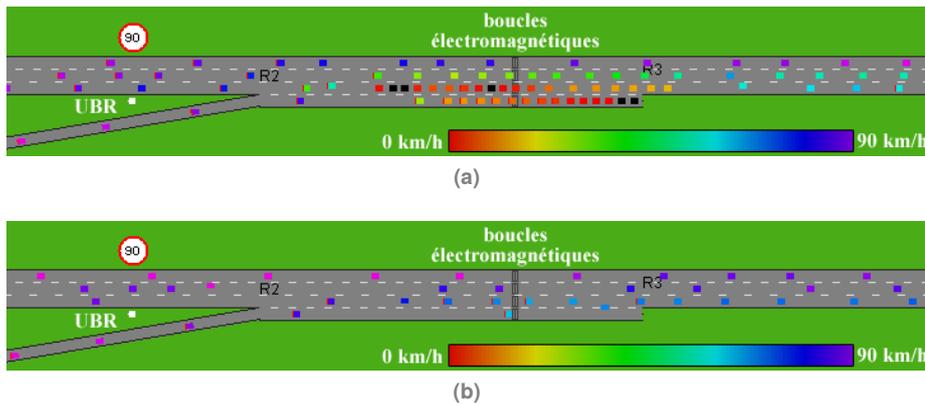


Fig. 3.18: Capture d'écran du simulateur pour le scénario d'insertion. (a) Instabilités du trafic avec 0% de véhicules connectés (b) Même situation avec 40% de véhicules connectés.

la situation de trafic décrite dans la figure 3.18 (b), il apparaît clairement que seulement 40 % de véhicules connectés suffisent à maintenir un trafic fluide (les captures sont réalisées aux mêmes temps pour les figures 3.18 (a) et 3.18 (b)). Comme le montre l'interface graphique, la section d'étude est équipée de capteurs de type boucle électromagnétique. Ces capteurs permettent d'agrégier les données de débit et concentration du flux au cours de la simulation. Ces données sont représentées dans la figure 3.19.

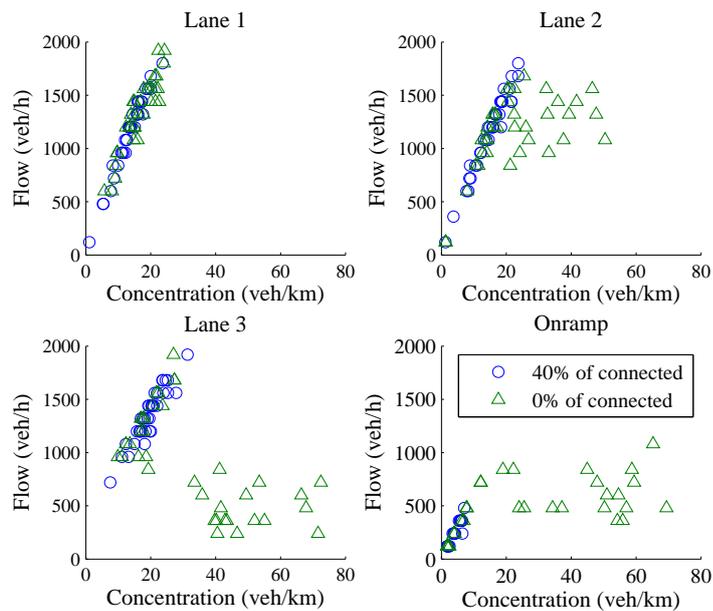


Fig. 3.19: Mesures des débits et concentrations dans le scénario d'insertion. Chaque point correspond à des données agrégées durant 1 min. Les diagrammes fondamentaux représentent alors chacune des voies de gauche à droite : (a-b-c) et la voie d'insertion (d).

Les diagrammes confirment qu'en l'absence de toute forme de contrôle (pas de véhicules connectés), la congestion se propage de la voie d'insertion jusqu'aux voies les plus à gauche. Dans le cas avec 40 % de véhicules connectés, la partie congestionnée (partie droite du diagramme fondamental) n'est jamais observée. Cela confirme l'effet des véhicules connectés, y compris sur le flux des véhicules circulant sur la voie d'insertion.

3.5.2 Fermeture de voie

Ce second scénario reproduit une situation commune génératrice de congestion : une réduction de la capacité du réseau. Une fermeture de voie peut être temporaire ou non, mais produit systématiquement une situation de conflit pour les véhicules. Dans ces conditions de trafic dense, la chute de capacité peut entraîner des ralentissements menant jusqu'à la propagation de congestion sur la section. Le scénario modélisé prévoit une réduction de 3 à 2 voies : la figure 3.20) montre une capture d'écran dans les deux situations à $t = 15 \text{ min } 13 \text{ s}$. La stratégie de contrôle de l'UBR reste identique au scénario précédent, mais la consigne demande désormais un changement de voie vers la droite.

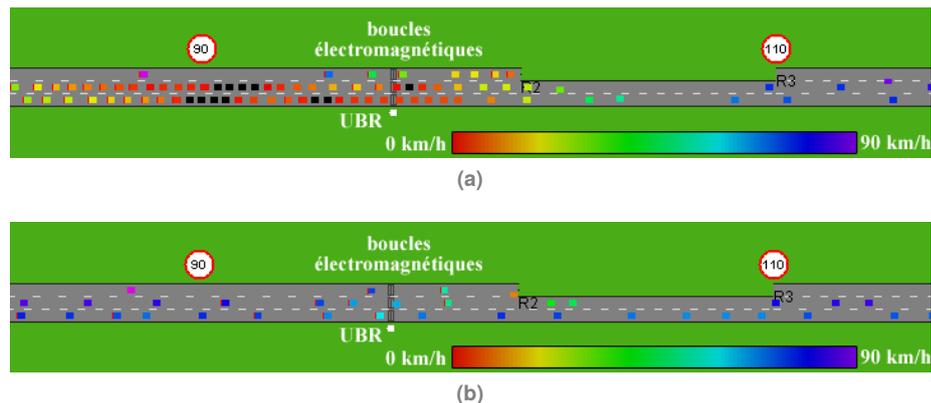


Fig. 3.20: Capture d'écran du simulateur pour le scénario d'insertion. (a) Instabilités du trafic avec 0% de véhicules connectés (b) Même situation avec 50% de véhicules connectés.

L'objectif de la stratégie de contrôle (figure 3.20 (b)) est double : laisser des créneaux d'insertion sur la voie centrale pour les véhicules qui arrivent de la voie de gauche, et anticiper le délestage de la voie de gauche. Cela permet de réduire le nombre de véhicules stockés sur cette voie tout en limitant le ralentissement sur les autres voies. Sans véhicules connectés (figure 3.20 (a)), toute la section est impactée, et le bouchon formé prend beaucoup d'ampleur. Dans une situation réelle (non modélisée dans le simulateur), des comportements égoïstes des conducteurs pourraient encore aggraver la situation (rester sur la voie de gauche jusqu'au dernier moment).

Le tracé des diagrammes fondamentaux par voie (figure 3.21) confirme les observations et montre la transition entre les états de trafic fluide et congestionnés dans le cas 0% de véhicules connectés. Dans ce scénario, plus de véhicules sont nécessaires (50%) pour empêcher la formation de congestion. Les diagrammes montrent bien l'importance de cibler les consignes, puisque les perturbations prennent forme exclusivement sur la voie de gauche.

3.5.3 Discussion

La stratégie de contrôle mise en place dans cette expérimentation est simple. Il s'agit de déterminer quelles consignes sont efficaces, ce qui dépend fortement des conditions de trafic des scénarios modélisés. Une UBR confrontée à des situations réelles fera face à des états de trafic évoluant dynamiquement. Le problème devient alors beaucoup plus complexe : l'UBR, autonome, doit être capable de discriminer ces états de trafic. Sa représentation doit alors intégrer la politique d'action

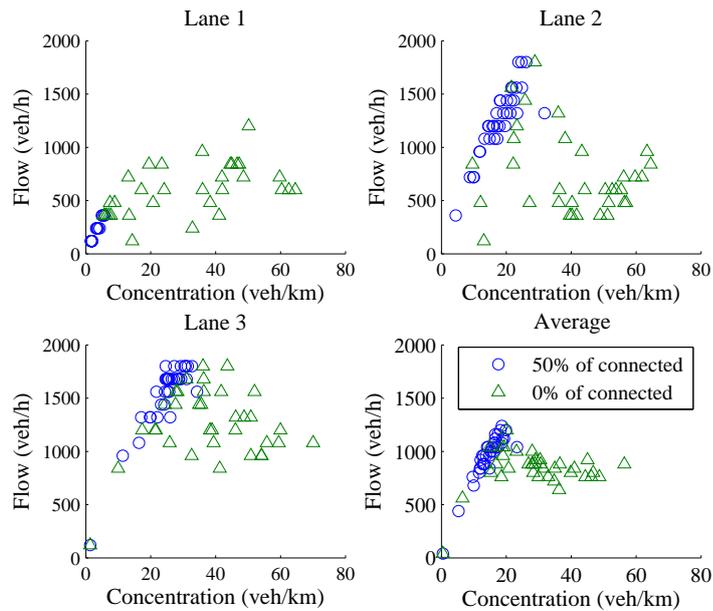


Fig. 3.21: Diagrammes débits-concentration du scénario de fermeture de voie (données boucles agrégées chaque minute). Voies de gauche, milieu, droite (a-b-c) et moyenne des trois voies (d).

optimale associée à chaque état de trafic. L'expérimentation menée sur des situations fixes a montré l'impact positif d'une modification partielle et temporaire des comportements d'un flux de véhicules connectés. Cela confirme le potentiel de l'utilisation des messages de consignes pour contrôler un flux de véhicules équipés.

La prochaine étape sera de concevoir un processus de décision complexe qui permettra de générer dynamiquement une stratégie de contrôle embarquée dans une UBR. Pour nous aider dans cette tâche, nous envisageons d'explorer les possibilités offertes par les systèmes multi-agents. Comme les comportements exhibés par les véhicules connectés et automatisés sont encore peu connus, le système devra être capable d'apprendre à la fois à reconnaître de nouvelles situations de trafic, mais aussi à assimiler de nouvelles stratégies de contrôle. Cette problématique trouve tout son intérêt dans les systèmes coopératifs mais se heurte à des verrous scientifiques plus généraux du domaine de l'IA distribuée. Par exemple, comment éviter d'orienter le comportement du système en incorporant trop de connaissances expertes ?

3.6 Bilan et perspectives

La motivation ayant mené au travail présenté dans ce chapitre prend son origine dans les objectifs fixés au préalable. Nous souhaitons montrer le potentiel des systèmes coopératifs dans une phase de déploiement anticipée. La simulation permet de se projeter dans ce cas de figure en modélisant le futur environnement des C-ITS. Ces systèmes sont cependant très complexes car les entités y interagissent par de multiples biais : capteurs, communication, et stratégies de contrôle. Cette réflexion a mené à l'élaboration d'une plateforme de simulation basée sur un outil existant tiré de la communauté du trafic. Grâce à l'ajout des concepts issus du domaine des multi-agents, nous avons implémenté le couplage proposé dans le chapitre 2. Le simulateur est désormais un outil

d'aide à la décision robuste qui permet de tirer profit des nouveaux véhicules et de l'infrastructure communicante pour proposer des nouvelles stratégies applicables à l'échelle des flux. Les véhicules et les UBR sont modélisés comme des agents autonomes capables de percevoir leur environnement par l'intermédiaire de leurs capteurs. Le processus de décision des véhicules peut être modélisé par une loi de poursuite classique mais il est aussi possible d'intégrer un processus plus complexe. La différence entre les véhicules automatisés ou autonomes et les véhicules classiques (conduits par des humains) se traduit donc par des processus de décision distincts. Dans la même veine, l'infrastructure possède la capacité d'agir en totale autonomie ou d'être asservie par un système automatique. Entre toutes ces entités, la communication est assurée par un protocole basé sur la notion de message. Les considérations techniques exposées dans le chapitre 1 ont permis d'accorder ces échanges avec les capacités actuelles des technologies et standards en place.

Le couplage des dynamiques du système permet d'envisager des stratégies de régulation à différentes échelles. Au niveau local, nous avons proposé une stratégie allant vers l'homogénéisation d'un flux de véhicules semi-automatisés et connectés. Agissant comme un système avancé d'assistance au conducteur, et voué à être utilisé dans des conditions de trafic propices à la congestion (trafic dense sur autoroute), le modèle est implémenté dans le simulateur. Ce modèle multi-anticipatif bilatéral profite de toutes les dynamiques du système, y compris la couche de confiance qui permet de gouverner les échanges d'information. Cette forme de contrôle a montré son potentiel dans les expérimentations en simulation, où l'objectif était de minimiser l'impact d'informations issues de capteurs défaillants. Le modèle complet a quant à lui confirmé le potentiel d'une forme de régulation locale permettant, par émergence, l'homogénéisation d'un flux et la réduction des perturbations. Un des résultats intéressants est que cet effet ne nécessite que la coopération d'une partie du flux ; ce qui confirme que les prochaines phases de déploiement des systèmes coopératifs seront décisives. L'évaluation des bénéfices de ce type d'approche est particulier dans la mesure où les résultats proviennent exclusivement de simulations. L'observation des flux modélisés ne permet pas d'évaluer toutes les variables impactées par les systèmes étudiés. Nous avons proposé un cadre robuste d'évaluation qui se base sur la définition d'indicateurs. En comparant des situations nominales (pas de véhicules connectés) avec des scénarios plus complexes, il est possible de pointer les différences attendues. Même si les différences avec des systèmes réels peuvent pondérer les résultats obtenus, ils pourront tout de même orienter les futures recherches vers l'étude en conditions réelles des modèles proposés.

Les résultats présentés dans ce chapitre montrent que de nouveaux comportements vont se mettre progressivement en place. Ils seront apportés lors du déploiement de nouveaux systèmes embarqués et côté infrastructure. Les problèmes d'interopérabilité entre les systèmes et les interactions possibles avec les usagers rendront leurs effets difficilement prévisibles. Un des enjeux majeurs des gestionnaires de trafic est d'optimiser l'usage de leur infrastructure dans un contexte où le nombre d'usagers ne cesse de croître alors que les capacités arrivent à saturation. Le potentiel d'action grâce à des stratégies intelligentes visant les C-ITS intéressera ces acteurs particuliers. Il est alors difficile de prédire quelle forme prendront ces stratégies de contrôle. Face aux résultats préliminaires obtenus par une stratégie décentralisée, nous avons identifié un axe de travail au potentiel évident. Il s'agit de mettre à profit l'infrastructure dans l'élaboration de stratégies de contrôle. Cette approche peut se combiner avec une approche locale de régulation et présente l'avantage de pouvoir intégrer des objectifs de haut niveau à la stratégie mise en place. L'apparition de nouveaux comportements est aussi à prendre en compte pour le succès à plus long terme de la stratégie. Nous proposons donc d'intégrer une forme d'apprentissage dans le processus de décision des unités d'infrastructure.

Une nouvelle fois, notre réflexion nous oriente vers la discipline de l'IA, qui présente à la fois des solutions mais aussi des problèmes associés. Le contrôle d'un flux de véhicules connectés par les UBR présente aussi ses défis. Comment intégrer des connaissances expertes dans un processus de décision autonome ? Comment ne pas biaiser le comportement d'un système autonome lors de sa conception ? Quels outils permettent à un système de construire une représentation de son environnement ? Comment intégrer des objectifs à l'échelle du système dans une stratégie de contrôle localisée et individualisée ? Toutes ces questions sont abordées dans le chapitre 4. Nous proposons un modèle d'IA capable de construire une stratégie de contrôle. Ce modèle est testé sur le cas d'étude du trafic coopératif grâce à la plateforme de simulation développée.

Vers un contrôle décentralisé et adaptatif du trafic coopératif

” *Instead of controlling the environment for the benefit of the population, perhaps it's time to control the population to allow the survival of the environment.*

— **David Attenborough**
Chercheur naturaliste

Les résultats de l'implémentation (chapitre 3) d'une stratégie d'homogénéisation du flux grâce aux concepts multi-agents et reposant sur les technologies C-ITS (chapitre 2) font entrevoir l'émergence de nouveaux comportements. Les véhicules connectés, voire automatisés, pourront proposer un usage de l'infrastructure plus altruiste qui bénéficiera par son efficacité globale à l'ensemble des usagers. Dans cette perspective, il est nécessaire de repenser les stratégies de contrôle et de régulation des flux. Ces méthodes doivent en effet prendre en compte le caractère évolutif de ces comportements (puisque liés aux différentes phases de déploiement des systèmes coopératifs) et peuvent désormais profiter de nouvelles sources d'informations (recueil V2I et contrôle I2V). La gestion et l'utilisation de ces informations sont centrales et conditionneront le succès de la stratégie de contrôle. À la manière d'un système autonome, les Unités de Bord de Route peuvent jouer un rôle particulier dans l'élaboration de stratégies intelligentes. Grâce à des capteurs côté infrastructure, en combinaison avec les messages provenant des véhicules, les UBR perçoivent une quantité d'information importante. Toutes ces données peuvent permettre de construire une représentation précise de l'état de trafic de la section observée. L'objectif est alors d'apprendre une politique d'action associée, qui permettra de contrôler un flux de véhicules connectés.

En étudiant les approches de contrôle classiques et celles dédiées aux C-ITS (section 4.1), nous mettons en évidence les parallèles existants avec l'Intelligence Artificielle Distribuée (IAD) et plus particulièrement les problématiques rencontrées dans le cadre de l'apprentissage multi-agent. Les verrous soulevés sont en effet très proches de certains problèmes fondamentaux rencontrés en IA. La construction d'une représentation fine par un agent autonome implique des précautions dans la modélisation. Comment garantir l'autonomie d'un système tout en orientant son processus de décision ? Comment intégrer des objectifs de haut niveau dans son processus sans biaiser son comportement ? Un nouveau paradigme émanant des sciences cognitives a orienté nos travaux vers les approches constructivistes (section 4.1.3). Nous étudions en détail les problématiques de ce domaine dans la section 4.2. Cette étude met en évidence des verrous qui émanent de ces problématiques (section 4.2.3) : le problème de l'attribution de récompense et le problème de l'élaboration de représentations concurrentes. En visant ces verrous scientifiques, nous appliquons ce paradigme en proposant un modèle d'IA générique dans la section 4.3. Ce modèle est ensuite implémenté dans deux environnements : un prototype permettant d'évaluer son comportement (section 4.4), et une application pour le contrôle des C-ITS grâce aux UBR (section 4.5). Le prototype permet de placer

le système dans un contexte propice à l'étude de certains traits de son comportement. Le cadre d'application du trafic coopératif permet de valider son fonctionnement dans un cas réel, où les données sont bruitées, et où l'environnement est moins prédictible. Ces deux applications ouvrent des perspectives à la fois dans le domaine de l'IA et aussi dans le futur des C-ITS (section 4.6). Notre réflexion débute avec l'étude des méthodes actuelles de contrôle et de régulation des flux de trafic.

4.1 Contrôle et régulation du trafic dans l'environnement du véhicule connecté

L'essor des technologies de l'information et le développement des capteurs côté infrastructure permet aux gestionnaires de surveiller l'état du réseau en temps réel. Le nombre croissant de véhicules évoluant quotidiennement au sein d'un réseau contraint pose la question de la régulation et du contrôle. Ce besoin est motivé par le caractère variable de la demande au sein du réseau (spatialement) et suivant l'heure de la journée (temporellement : les bien connus pics de trafic du matin et du soir). Les C-ITS offrent de nouvelles possibilités de recueil de données et de remontée d'information. Dans le sens véhicule à gestionnaire, ils permettent une meilleure compréhension de l'état de trafic et peuvent mener à une détection plus précoce des incidents (accidents, perturbations). Dans le sens gestionnaire à usager, les nouvelles technologies permettent de relayer des informations de contrôle (consignes) plus personnalisées et plus précises (impactant par exemple seulement une partie du flux ou un type de véhicule en particulier).

Les approches de régulation ne sont pas récentes, mais les gestionnaires comptent encore aujourd'hui sur des méthodes empiriques qui reposent essentiellement sur l'expertise d'un opérateur. Ces méthodes, présentées dans la section 4.1.1, sont appliquées en utilisant des technologies qui ont évolué au cours du temps : du panneau à message variable (PMV) jusqu'aux limites dynamiques de vitesse (VSL). Plus récemment, de nouvelles approches visent à automatiser les méthodes de régulation (contrôle d'accès, régulation des feux, régulation dynamique des vitesses). Dans la section 4.1.2, nous étudions les solutions qui mettent à profit les technologies des ITS et s'attèlent à proposer des stratégies intelligentes de régulation du trafic. Les nouveaux comportements des véhicules connectés et automatiques nécessiteront de nouvelles formes de contrôle adaptées. L'objectif de notre étude est de déterminer les techniques les plus propices à l'élaboration de telles stratégies.

4.1.1 Approches classiques de contrôle et régulation du trafic

Il existe beaucoup de stratégies de contrôle du trafic routier, qui sont étudiées du côté des gestionnaires d'infrastructure (voir [158] pour une revue). La grande majorité des travaux s'intéresse à un réseau maillé où la cause de la congestion est bien souvent le réseau lui-même (feux de signalisation, priorités, etc.). Les informations proviennent de capteurs (boucles électromagnétiques, caméras) permettant de récolter des informations et d'assurer une surveillance (*monitoring*) du réseau par un opérateur, ou un système intelligent. Le moyen d'application de la stratégie de contrôle choisie passe par la transmission d'information à l'ensemble du flux (PMV) ou en modifiant directement les caractéristiques du réseau (fermeture d'accès, régulation du cycle de feux, etc.). La

tendance récente oriente les approches vers une gestion dynamique, où la stratégie repose sur une représentation proposée par un expert. C'est par exemple le cas pour la régulation dynamique des vitesses (voir [109] qui présente les différentes approches à base de VSL), en place sur l'autoroute du soleil (A7).

La performance de ces approches repose essentiellement sur les connaissances expertes intégrées dans les représentations manipulées par les systèmes de contrôle. Cela permet de proposer des systèmes rapidement opérationnels, mais l'écart observé entre les modèles issus de la théorie du trafic et les comportements hétérogènes des conducteurs ne garantit pas forcément une stratégie de contrôle optimale. De plus, les moyens d'application des stratégies classiques se limitent principalement à des actions ciblées sur l'ensemble d'un flux d'une section contrôlée. Dans certains cas comme pour le contrôle d'accès (où il s'agit de gérer les flux entrant sur une section), on favorise une partie du flux au détriment des usagers circulant sur d'autres parties du réseau. La difficulté réside dans la propagation des perturbations. L'origine d'une perturbation ne peut pas toujours être maîtrisée et le contrôle d'une section peut avoir des incidences néfastes sur une autre portion du réseau. Face à ces limitations, les C-ITS permettent d'agir sur les deux niveaux : en apportant plus d'informations permettant une description plus fine de l'état de trafic, mais aussi en donnant des moyens d'action plus précis permettant de transmettre des consignes plus ciblées.

4.1.2 Contrôle et régulation des C-ITS

Face à la complexité des systèmes où de multiples acteurs partagent un même environnement routier, le contrôle des C-ITS se rapproche clairement des problématiques abordées au niveau de l'intelligence artificielle distribuée (IAD). C'est pourquoi on retrouve beaucoup d'approches qui utilisent des techniques issues du domaine des systèmes multi-agents (Bazzan [17] propose une revue complète des approches existantes). En parfaite analogie avec les problématiques issues des SMA, la question qui se pose relève du type de contrôle à appliquer (approche centralisée, décentralisée, ou complètement locale). Les travaux récents montrent une tendance allant vers la décentralisation des approches classiques [26]. D'autres approches proposent des stratégies de contrôle locales, intégrées à la prise de décision du véhicule [107]. L'approche proposée dans le chapitre 2, autour du concept de confiance, peut être vue comme une solution locale : les informations sont partagées localement et le contrôle, qui s'appuie sur ces données, fait partie de la décision de l'agent. Les approches locales sont cependant plus difficiles à maîtriser dans l'objectif de répondre à des objectifs de haut niveau (*i.e.* à l'échelle du système).

La plupart des travaux existants [55, 203] se focalisent une nouvelle fois sur la gestion des intersections signalisées (feux tricolores). Les concepts multi-agents permettent de reproduire des comportements de coalition permettant de favoriser le passage de pelotons homogènes de véhicules. Le moyen d'action (*i.e.* l'application du contrôle) se limite principalement aux limitations de vitesse dynamiques (VSL) [220]. Il est aussi possible de directement modifier les comportements individuels en fonction de l'état de trafic [180]. Conformément à notre hypothèse, le cas d'étude d'un scénario autoroutier est aussi représenté dans les travaux portant sur le contrôle des véhicules connectés [187] et les approches basées sur une forme d'apprentissage se développent [125]. L'intérêt pour ces approches reste récent mais est en forte croissance depuis quelques années ; suivant le rythme de déploiement des véhicules intelligents.

4.1.3 Vers une approche constructiviste pour le contrôle des C-ITS

Les limitations mises en avant dans notre étude des approches de régulation classiques reposent essentiellement sur la dépendance du système à une représentation experte (souvent couplée avec une base de règles elle-même experte). L'utilisation de formes d'apprentissage semble être une alternative pour donner au système la capacité de construire et pourquoi pas faire évoluer sa représentation. Cette possibilité est en adéquation avec la dynamique de déploiement des systèmes coopératifs, qui nécessite à la fois de nouvelles stratégies de contrôle dès aujourd'hui, mais qui demandera aussi une adaptation future. L'application des stratégies (*i.e.* le type d'action du système) est aussi un axe d'amélioration des approches existantes, qui se limitent souvent à une seule forme de contrôle (VSL, routage) même si les consignes propagées sont de plus en plus personnalisées. Enfin, les stratégies décentralisées, qui permettent d'intégrer une dimension d'autorité (l'unité de contrôle) tout en garantissant de mener à bien des objectifs à grande échelle, semblent la forme de contrôle qui semble la plus adaptée aux systèmes coopératifs.

L'approche présentée dans cette thèse est vouée à être intégrée dans une Unité de Bord de Route, modélisée comme un agent autonome, et dont le processus de décision sera un système décentralisé de contrôle des C-ITS capable, par apprentissage, de construire sa représentation de l'état de trafic courant et d'optimiser les consignes à envoyer aux véhicules. Ces consignes ne se limiteront pas à un seul paramètre (la vitesse) mais permettront aussi aux véhicules d'adapter d'autres variables menant à leur comportement (inter-distance, et changement de voie).

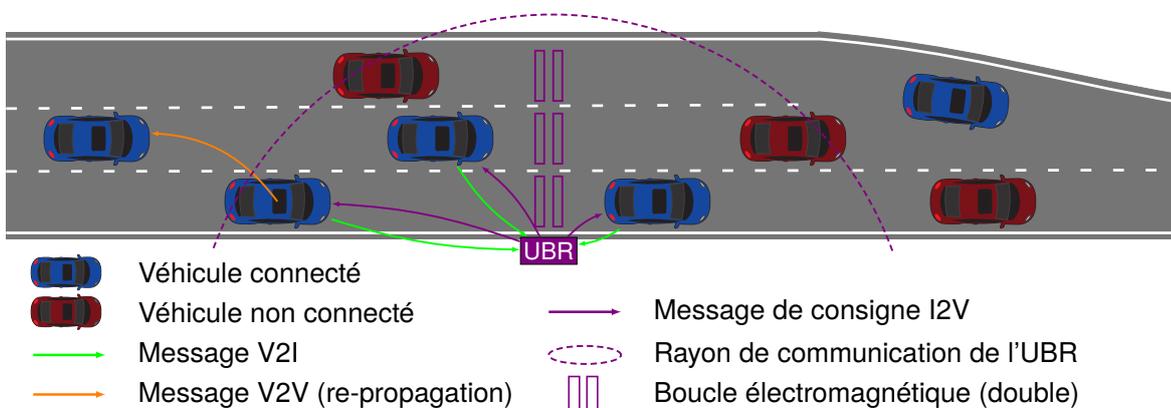


Fig. 4.1: Illustration d'un scénario de contrôle décentralisé des systèmes coopératifs

La figure 4.1 représente un scénario de contrôle des véhicules connectés par une UBR autonome. L'objectif de ce système de contrôle est d'améliorer l'écoulement du flux. Pour cela, il doit construire une représentation de l'état de trafic courant et apprendre quelles consignes sont les plus pertinentes dans ces différents contextes. Les interactions dans l'environnement sont rendues complexes par l'ajout de la communication entre les véhicules connectés. La proportion de véhicules équipés est une nouvelle variable à prendre en compte dans la décision de l'UBR pour obtenir une stratégie de contrôle robuste. Pour cela, l'UBR dispose de plusieurs capteurs (communication – vitesses – et boucles – débit et concentration) lui permettant de percevoir des variables de l'environnement. Les approches classiques ne suffisent plus à décrire précisément l'état de trafic courant : les nou-

veaux comportements des véhicules connectés et autonomes doivent être intégrés dynamiquement. L'impact de nouveaux systèmes (comme un nouveau véhicule autonome) à l'échelle du système ne sera pas évalué *a priori* mais observé lors de son déploiement. L'objectif de notre approche est de permettre au système de construire ces états et les actions associées, qui feront partie de sa représentation de l'environnement. La difficulté de cette tâche est fortement liée à l'environnement particulier du véhicule connecté. Cet environnement présente en effet les caractéristiques suivantes :

- **continuité spatio-temporelle** : l'environnement physique est continu et le réseau routier complexe, tout en étant le support de dynamiques à la fois court-terme (communication) et plus long-terme (congestion). Le système de contrôle doit être capable de gérer un flot continu de données, par exemple en isolant les plus pertinentes. Il doit aussi être capable de percevoir des événements récurrents ou non, quelle que soit leur dimension temporelle.
- **indéterminisme et évolutivité** : l'environnement physique, notamment l'ensemble des états accessibles, ne sont pas connus *a priori*. Le système de contrôle doit être capable de faire évoluer sa représentation pour s'adapter à cette contrainte. De la même manière, les objectifs du système peuvent être amenés à évoluer (cas des phases successives de déploiement où de nouveaux usages et besoins vont apparaître).

Ces caractéristiques de l'environnement visé dans notre travail définissent un type de problème particulier, qui est visé par les approches constructivistes.

4.2 Motivations

La construction de la représentation est un problème fondamental dans l'élaboration de toute stratégie de contrôle. Il s'agit d'un problème cognitif qui fait intervenir toute la chaîne d'interaction de l'agent. Les approches constructivistes, inspirées des sciences cognitives, visent à reproduire ce comportement itératif d'élaboration de la représentation. Dans cette partie, nous proposons de les décrire en étudiant leurs caractéristiques (section 4.2.1). Ces caractéristiques peuvent être intégrées dans une architecture de référence d'un système dont le processus d'apprentissage suit le paradigme constructiviste (section 4.2.2). Les phases successives de déploiement des C-ITS feront apparaître de nouveaux comportements au niveau des véhicules. La problématique principale relative au contrôle de ses systèmes passera alors par les capacités d'adaptation de l'infrastructure. Afin de couvrir le plus de situations possibles, la représentation du système de contrôle devra aussi faire preuve de suffisamment d'expressivité. Ces problématiques sont liés à des verrous scientifiques présentés en section 4.2.3.

4.2.1 Caractéristiques d'un système constructiviste

Reproduire le comportement d'apprentissage et d'adaptation d'un système autonome, tel est l'objectif du constructivisme. En effet, l'une des caractéristiques principales de l'autonomie d'un système réside dans ses capacités d'adaptation. Un tel système doit donc être capable de construire et faire évoluer sa représentation. Les approches constructivistes sont inspirées des travaux précurseurs de Piaget [161], en sciences cognitives. Il propose de décrire le processus d'acquisition de connaissance lors de la phase de développement chez l'enfant. Ces travaux ont orienté la réflexion portant sur les systèmes autonomes et plus particulièrement les conditions menant à l'autonomie [221] :

- le système doit être incarné (*embodiment*) ;
- il doit être situé dans un environnement physique et social ;
- il doit exhiber des mécanismes de développement épigénétiques à travers lesquels des structures plus complexes émergent du résultat des interactions avec l'environnement social et physique.

La notion d'épigénétique a été proposée par Piaget. Elle désigne le développement d'un individu grâce à ses interactions successives avec l'environnement. Cette théorie a donné lieu à la proposition de méthodes permettant de construire une représentation sous la forme de schémas sensorimoteurs [54], qui représentent les interactions par des triplets contexte-action-résultat. Ce concept a été largement repris par la communauté IA (le lecteur intéressé est invité à consulter la revue de Guerin [84]). Les applications de ces concepts se focalisent sur l'intelligence ambiante [145, 131] ou plus majoritairement sur une variante de la robotique : la robotique développementale [124, 132, 7]. L'approche la plus prometteuse dans le domaine de l'IA développementale semble consister à reproduire le processus d'acquisition de la connaissance dans les différentes phases du développement de l'humain [139]. Ce type d'approche se heurte à une difficulté majeure qui est la complexité de l'environnement, car le système d'apprentissage doit être confronté à un monde physique raisonnablement réaliste [84].

Le fonctionnement interne d'un système constructiviste peut être décomposé en deux processus couplés : un processus d'identification du contexte, qui permet au système de discriminer les situations (ou l'état de l'environnement) et un processus de prise de décision, qui permet au système de choisir une action à entreprendre dans le contexte identifié. Ces deux processus sont liés et font partie du processus plus global d'apprentissage du système qui mène à l'élaboration d'une représentation de l'environnement du système.

4.2.1.1 Identification du contexte

Le processus d'identification du contexte permet à un système d'extraire les caractéristiques de sa perception de l'environnement utile à son interaction. Les approches expertes fournissent, grâce à des connaissances *a priori*, une discrétisation de l'environnement sous la forme d'un espace d'états finis. Ce type de partitionnement de l'espace de perception limite l'interaction du système à des environnements faiblement dynamiques, où l'ensemble des configurations peut être connu *a priori*. De même, l'ensemble des capteurs du système doit être connu. Dans le cas d'un système constructiviste, ce processus d'identification du contexte devient dynamique. Il peut reposer également sur une discrétisation de l'espace de la perception (variables des capteurs), mais cette discrétisation est cette fois-ci dynamique. Cela permet d'améliorer les capacités d'adaptation du système dans le cas où il fait face à un nouvel environnement, ou si ces capteurs évoluent dans le temps.

Les problèmes attendus de ce type de processus dynamique reposent essentiellement sur la quantité de données nécessaires, qui va de pair avec le temps nécessaire avant d'obtenir une discrétisation précise des contextes. En effet, dans le cas d'approches expertes, les connaissances nécessaires pour proposer une discrétisation sont issues de l'expérience du concepteur. Il y a donc un apport non négligeable de connaissances acquises *a priori*. Cela permet à un système autonome de disposer d'une discrétisation avant de se concentrer sur l'apprentissage de la politique d'action associée (processus de décision). En contrepartie, cette discrétisation est fixe, ce qui va à l'encontre d'une

forme d'adaptation face à un environnement dynamique, ce qui est un pré-requis à l'autonomie du système. Faire évoluer dynamiquement la discrétisation, comme proposé dans les approches constructivistes, semble permettre de pallier ce problème. Cependant, il est nécessaire de lier cette discrétisation à l'usage, c'est-à-dire la partie décision du processus d'apprentissage du système.

4.2.1.2 Prise de décision

L'objectif du processus de décision du système est de choisir dynamiquement une action à effectuer face au contexte identifié. L'ensemble des actions entreprises par le système forme une politique d'action. La représentation du système peut alors être modélisée comme le couplage de la politique d'action et de la discrétisation associée. Cette représentation peut être donnée au système à partir de connaissances expertes, ou construite itérativement, en suivant le paradigme constructiviste.

Dans le cadre d'approches expertes, la politique d'action est souvent limitée à une base de règles pré-établies, ce qui rend la représentation du système statique. Ce type d'approche peut convenir pour des problèmes bien connus et au périmètre maîtrisable, mais devient rapidement inefficace face aux environnements ouverts, dynamiques et de grande dimension. Les approches classiques d'apprentissage ont pour objectif de générer, par essai erreur, une politique d'action associée à la discrétisation du système. Afin de représenter la performance du système et donc la qualité de la politique construite, il est possible d'attribuer une récompense au système (*feedback*) : on parle alors d'apprentissage par renforcement. Le choix du *feedback* contribue pour beaucoup au résultat de la construction de la politique d'action. Sa performance, ou plutôt sa précision, dépendent en plus de la qualité de la discrétisation associée dans la représentation. Les approches constructivistes proposent de lier les deux processus afin de s'assurer que l'apprentissage de la politique d'action soit lié au processus d'identification du contexte. De telles approches peuvent donc théoriquement permettre de modéliser un processus d'adaptation, où des connaissances acquises précédemment contribuent à l'élaboration d'une nouvelle (partie de la) représentation.

4.2.2 Une architecture de référence

Afin d'identifier les axes de travail permettant de proposer un modèle d'apprentissage constructiviste, le plus simple est de représenter les processus mis en œuvre lors de la construction d'une représentation par un système autonome. La figure 4.2 reprend les deux grandes caractéristiques d'un système constructiviste : l'identification du contexte et la prise de décision.

Ces deux caractéristiques principales, qui, couplées, mènent à l'élaboration de la représentation du système, peuvent être décomposées en différents processus inter-liés. Chacun de ces processus, considéré individuellement, vise à traiter l'une des tâches de l'apprentissage. D'autres caractéristiques, comme la capacité d'abstraction du système, dépendante de la précision de la représentation, sont des propriétés qui émanent de la synergie au sein du système. Le schéma décrit une forme d'apprentissage par renforcement, qui exploite une récompense attribuée par le système. Cette récompense, ou *feedback* permet au système d'évaluer le résultat des actions exécutées dans l'environnement. Il s'agit d'une partie de la perception (ou d'une transformation). L'hypothèse est que les canaux (et/ou le processus de transformation) sont implicites pour le système. Dans la suite, les notions de *feedback* et de récompense sont analogues. Le problème d'attribution de la récompense se réfère alors au problème du *feedback*, critique dans l'apprentissage par renforcement,

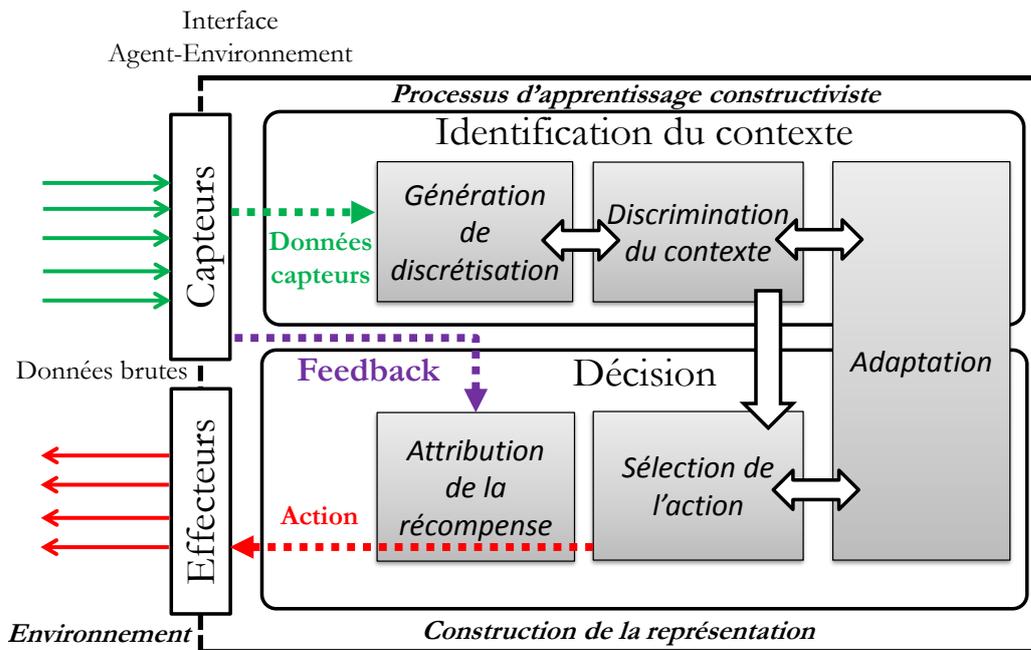


Fig. 4.2: Schéma des processus internes d'un agent en interaction avec son environnement capable de construire sa représentation en suivant un apprentissage constructiviste.

puisqu'il détermine la façon de faire évoluer les processus d'identification du contexte et de prise de décision.

La modélisation de ces processus intégrés dans la construction itérative de la représentation soulève des verrous qui ne sont pas encore tous résolus dans la littérature.

- Le processus de **génération de la discrétisation** peut être délégué à un modèle issu de connaissances *a priori*, ce qui risque de biaiser le comportement du système. Pour éviter cela, une possibilité est de permettre au système de générer une ou plusieurs discrétisations. Il est alors nécessaire d'identifier les méthodes et algorithmes qui conviennent le mieux dans cette tâche.
- La **discrimination du contexte** impose au système d'être capable de percevoir suffisamment finement son environnement. Cela nécessite de proposer des mécanismes permettant d'explorer les différentes combinaisons de variables d'entrée de la perception. En effet, certains phénomènes ne s'observent que dans des sous-espaces précis de la perception.
- Le processus de **sélection de l'action** est central car il est l'un des plus importants de l'aspect décision. C'est ce processus qui mène à l'élaboration de la politique d'action du système. Dans les approches classiques d'apprentissage, l'objectif est principalement de maximiser une fonction de récompense qui ne dépend que du processus de sélection de l'action. Dans les approches constructivistes, le problème est plus complexe puisqu'il s'agit de lier les processus de discrétisation et de sélection de l'action. On ne fait pas l'hypothèse que la discrétisation est juste, puisqu'elle est produite par le système lui-même et est vouée à évoluer.
- C'est le **processus d'attribution de la récompense** qui peut permettre d'agir sur les deux caractéristiques de la représentation. Le problème du *feedback* est cependant un verrou en soi. Il est difficile d'évaluer dans quelle mesure les autres processus (génération, discrimination, sélection) doivent être impactés pour permettre une bonne construction par le système. Une

récompense déséquilibrée favoriserait un processus au détriment des autres, entraînant la construction d'une représentation peu précise.

- Enfin, le processus d'**adaptation** peut être vu comme un processus de plus haut niveau, qui s'ajoute aux processus précédents. La capacité d'adaptation du système est alors modélisée comme un mécanisme capable de réutiliser les connaissances acquises lors des interactions précédentes afin d'accélérer les processus du système (réutilisation de discrétisation, capacité de prévision, généralisation, etc.).

L'objectif de notre travail est de proposer des mécanismes permettant de modéliser ces processus. Face à la complexité de certains d'entre eux (comme l'adaptation), il est nécessaire d'étudier les problématiques sous-jacentes. Il est évident que certains processus resteront limités par nos hypothèses, mais le cadre constructiviste proposé reste la ligne directrice du modèle décrit dans ce chapitre.

4.2.3 Problématique et verrous scientifiques

Notre travail est avant tout motivé par des problématiques émanant du contrôle de systèmes complexes tels que les C-ITS. Nous avons proposé un cadre de modélisation multi-agent permettant de tester des stratégies de régulation locale (homogénéisation), mais le comportement non linéaire des systèmes coopératifs laisse entrevoir un besoin de contrôle plus générique, permettant de répondre à des objectifs à l'échelle du réseau (ou d'un sous-réseau). Nous souhaitons proposer un modèle capable de se substituer à la prise de décision d'un agent autonome (UBR) afin de générer dynamiquement des stratégies de contrôle (représentation et politique d'action associée) permettant de contrôler un flux perçu par l'intermédiaire de capteurs. En faisant preuve d'un peu d'abstraction, notre problème peut être défini de manière plus générique. Nous souhaitons proposer un processus de prise de décision d'un agent autonome, situé dans un environnement physique, capable de percevoir partiellement cet environnement par ses capteurs, et cherchant une stratégie d'action visant à maximiser des objectifs globaux. L'étude des travaux existants nous a menés à nous intéresser au paradigme constructiviste et à ses propres problématiques (section 4.2.3.5). Nous proposons d'abord d'étudier les verrous scientifiques soulevés et relevant du domaine de l'IAD (sections 4.2.3.1 à 4.2.3.4).

4.2.3.1 L'apprentissage par renforcement

Apprentissage et intelligence sont deux notions étroitement liées, ce qui rend souvent difficile la distinction entre IA et apprentissage. Un système capable d'apprendre mérite d'être appelé "intelligent"; et à l'inverse, on attend entre autres d'un système qualifié d'intelligent qu'il soit capable d'apprendre [208]. L'apprentissage en général peut alors se définir comme une auto-amélioration du comportement grâce aux expériences passées. L'apprentissage plus spécifiquement en IA est alors décrit par l'acquisition de nouvelles connaissances et capacités cognitives qui sont, par action du système lui-même, réutilisées dans les activités futures du système, afin d'améliorer ces performances. La notion de performance implique qu'une forme d'évaluation des interactions avec l'environnement est possible. Les formes d'évaluation (*feedback*) décrivent alors les grandes familles d'apprentissage [156] : supervisé, à base de récompense, non supervisé. Pour l'apprentissage supervisé, la sortie du système "idéale" (*i.e.* la meilleure action possible pour la perception courante) est donnée au système dont l'objectif est de s'y conformer. Sa politique (*i.e.* les actions entreprises

par l'agent dans l'environnement) évolue dans le but de se rapprocher des actions idéales connues par l'environnement. En apprenant ainsi à partir d'exemples, le système a pour objectif de généraliser son apprentissage à des configurations inconnues. L'apprentissage non supervisé utilise une fonction d'évaluation (aussi appelée fonction d'utilité) qui est implicite pour le système et permet d'évaluer les activités ou les étapes du processus d'évolution de la représentation de l'agent. La fonction que l'on cherche à maximiser correspond souvent à un *a priori* sur la solution recherchée. Sa politique, *i.e.* la suite d'actions exécutées en réponse aux contextes successivement observés, n'est donc pas directement évaluée. Enfin, les méthodes d'apprentissage par récompense utilisent un *feedback* qui représente une fonction objectif à maximiser. Il est alors possible d'orienter le comportement du système vers un objectif à l'échelle de l'environnement, tout en garantissant l'autonomie du processus de décision. Plusieurs variantes de méthode d'apprentissage par récompense existent ; la plus classique reste l'apprentissage par renforcement. Ce type d'apprentissage reproduit le comportement par essai-erreur dans un environnement dynamique [100] où chaque action est immédiatement évaluée par l'attribution d'une récompense. L'évolution de la performance de la politique d'action du système peut alors être observée en suivant l'évolution des récompenses (quel que soit l'horizon). Cette façon d'évaluer le comportement du système correspond bien aux enjeux haut niveau dans les C-ITS, et s'accorde avec notre modélisation. On peut en effet évaluer par simulation l'efficacité d'une politique sur le flux par l'intermédiaire de différents indicateurs (section 3.3) afin de la renforcer. Cependant, de par la complexité du système simulé et de la politique elle-même, cette relation de récompense n'est pas toujours triviale à identifier. En théorie, la récompense obtenue sur le long terme doit permettre au système de converger vers une représentation stable, malgré les latences et la complexité du système. En pratique, pour les problèmes réels, la vitesse de convergence peut être très faible.

La figure 4.3 résume de manière abstraite le cycle d'un agent qui souhaiterait contrôler son environnement en faisant reposer la construction de sa représentation sur un apprentissage par renforcement.

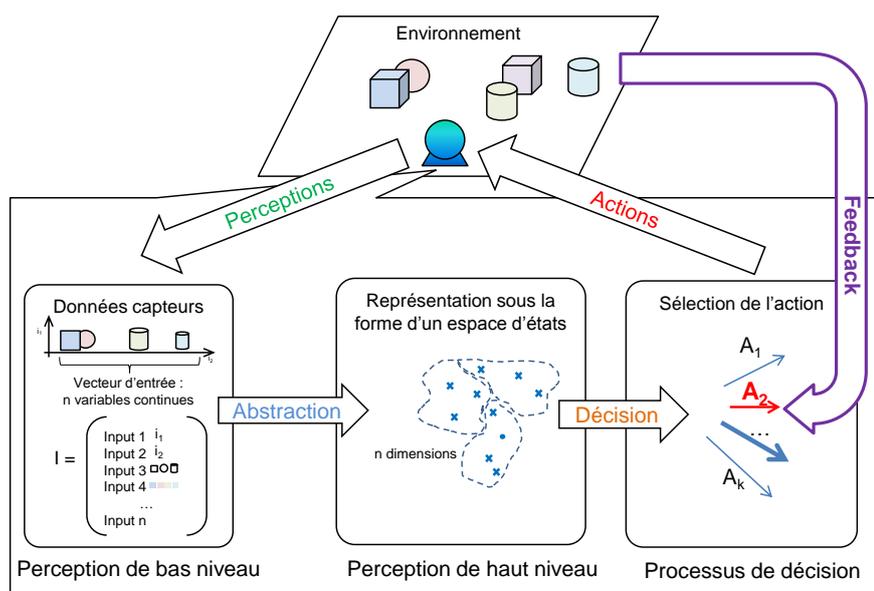


Fig. 4.3: Schéma simplifié des interactions d'un agent faisant évoluer sa représentation sous la forme d'un espace d'états dans un contexte d'apprentissage par renforcement.

Dans les problèmes d'apprentissage par renforcement, la représentation de l'environnement de l'agent est assimilée à un espace d'états discrets. Cet espace d'états constitue une abstraction de l'espace des données des capteurs issus de la perception. Chaque perception est alors assimilée à un des états de la représentation et la prise de décision de l'agent correspond à la sélection d'une action qui sera exécutée dans l'environnement. Chaque action donne lieu à une récompense immédiate qui permet à l'agent d'évaluer le résultat de son action. La représentation étant fixe, l'objectif de l'agent est alors de trouver la politique associée à cette représentation qui lui permettra de maximiser les récompenses obtenues de l'environnement sur le long terme. Le cœur de l'algorithme se situe donc avant tout dans l'utilisation des récompenses pour faire évoluer la politique de l'agent. En réalité, la notion de représentation peut être plus largement étendue. L'abstraction de la perception et la politique d'action font partie des capacités sensorimotrices de l'agent. Le processus d'apprentissage ne se limite donc plus à faire évoluer la politique ; il consiste à mettre à jour la représentation complète (discretisation et politique). Dans ce cas, le problème est plus complexe, puisque les interactions du système reposent sur des processus inter-dépendants. Ce type de problématique est typiquement un des cas où l'usage des SMA permet de simplifier le problème lors de la modélisation. D'ailleurs, l'apprentissage en IA distribuée constitue un champ d'étude à lui seul.

4.2.3.2 L'apprentissage en Intelligence Artificielle Distribuée

Les systèmes multi-agents permettent de reproduire ou d'imaginer des sociétés d'agents autonomes, qui peuvent être dotés de formes d'intelligence, ou mener par interactions à l'émergence de comportements de groupe intelligents [6]. L'introduction de mécanismes d'apprentissage dans les SMA peut donc être réalisée de différentes manières, qui profitent à leur façon des propriétés de décentralisation de l'IAD. Pour simplifier, il est possible de considérer le SMA comme un système d'apprentissage, où chacun des agents est vu comme une entité dotée de capacités d'apprentissage. Dans ce cas, les méthodes classiques d'apprentissage peuvent être directement appliquées dans le but d'améliorer le comportement individuel des agents, catalyser leurs interactions, ou encore faire émerger des organisations. Les agents peuvent par exemple apprendre à coordonner leurs actions [207]. Dans le cas où le système est l'entité qui réalise l'apprentissage, les interactions des agents sont le support de la construction de la représentation du système. Il est possible d'imaginer que les agents eux-mêmes font partie de la représentation du système [131], et que leur organisation a une incidence sur le comportement global du système (et donc sa politique). Tous les agents du système ne sont pas forcément dotés de capacité d'apprentissage, le comportement d'un groupe d'agents peut être appris par un seul agent spécialisé. A l'inverse, chaque agent peut être spécialisé dans une sous-tâche de l'apprentissage du système. Panait et Luke [156] nomment ces types de topologie d'apprentissage pour l'équipe (*team learning*) et apprentissage concurrent (*concurrent learning*).

Au delà de l'aspect organisationnel décrit jusqu'à présent, nous attirons l'attention sur la façon de faire évoluer l'apprentissage de façon concurrente. Dans le cadre de l'apprentissage concurrent, le problème peut être projeté dans autant de sous-espaces que le système comporte d'agents apprenants ; ce qui est un avantage manifeste pour les problèmes très complexes ou décrits dans des espaces à grand nombre de dimensions. Cette forme d'apprentissage offre des possibilités intéressantes, telles que l'instanciation naturelle d'une mécanique d'exploration et d'exploitation au sein de la population d'agents, ou encore la possibilité de maintenir des politiques contradictoires, et de les activer suivant les contextes rencontrés. Le problème principal que rencontre ce type d'approche dans le domaine de l'apprentissage multi-agent est qu'il est difficile de s'assurer de la

performance du comportement des dynamiques système [21]. Ce verrou reste difficile à lever, malgré les avantages possibles à utiliser les méthodes récentes du domaine des SMA. Par exemple, si l'on revient à la notion de représentation, sa construction et son évolution pourraient être déléguées à une population d'agents dont la prise de décision intègre des mécanismes classiques d'apprentissage (comme la classification – *clustering*). L'apport de mécanismes concurrents tels que décrits dans [78] (*stacking, boosting, bagging*) est d'ailleurs un avantage indéniable, qui pourrait être apporté par des solutions multi-agents.

Les méthodes d'apprentissage multi-agent sont séduisantes pour la résolution de problèmes complexes et/ou prenant place dans des environnements avec beaucoup de dimensions. L'apprentissage par renforcement permet de donner un objectif global au système et de guider son processus. Dans le modèle présenté dans ce chapitre, nous proposons de tirer profit d'une forme d'apprentissage concurrent, afin d'améliorer la performance du système. Cela nous confronte donc aux trois problèmes sus-cités, qui se rapportent à des verrous scientifiques en IAD, et de façon plus générale relèvent du problème de l'élaboration de toute représentation cognitive. Nous proposons d'étudier en détail ces problèmes que nous chercherons à résoudre dans le cadre de notre travail : le problème de l'attribution de récompense (section 4.2.3.3) et le problème de l'élaboration de représentations concurrentes (section 4.2.3.4)

4.2.3.3 Le problème du *feedback* : individuel ou collectif ?

Le problème du *feedback* (ou encore d'attribution de la récompense) est un problème classique en apprentissage. En effet, l'objectif d'un système reposant sur une population d'agents est souvent éloigné des objectifs individuels. Pourtant, l'apparition d'une synergie au sein du système repose essentiellement sur des comportements individuels précis, et exécutés au même moment. En d'autres termes, le problème est de s'assurer que les objectifs individuels des agents seront corrélés avec l'objectif global du système. L'apprentissage par renforcement offre la possibilité au système d'évaluer sa politique en temps réel grâce à une récompense instantanée. Le concepteur peut alors proposer une fonction de récompense qui dépend des objectifs de haut niveau définis pour le problème, ce qui est un avantage pour orienter le comportement du système. Si les mécanismes reposent sur des comportements individuels, cette récompense risque de ne pas guider suffisamment rapidement les agents dans leur propre apprentissage. Dans ce cas, il convient de décomposer le problème de l'attribution de récompense en deux sous-problèmes [6] :

- **attribution** de récompenses **inter-agent** : récompenser les agents sur la base de la performance globale du système. Il s'agit d'une tâche difficile car il est nécessaire d'estimer l'apport de chacun des agents dans le fonctionnement global du système.
- **attribution** de récompenses **intra-agent** : récompenser les agents sur la base de leur comportement individuel. Plus simple au premier regard, cela consiste à récompenser les actions individuelles qui peuvent mener à des bonnes performances du système, mais on retrouve les problématiques liées au lien interactions-émergence qui n'est pas toujours explicite [21].

Une approche naïve consisterait à décomposer équitablement les récompenses pour tous les agents. Cela n'est possible que pour des systèmes simples ou les entités sont toutes identiques. Lorsque les agents sont spécialisés, le problème de l'attribution de récompense est crucial. Il faudra donc proposer des formes de *feedback* à la fois interne aux agents, et aussi entre agents afin d'orienter le comportement du système vers l'objectif global. Pour cela, le système devra être capable de décrire précisément le contexte auquel il est confronté.

4.2.3.4 Le problème de l'élaboration de représentations concurrentes

Dans la section 4.2.3.6, nous avons identifié l'adaptation comme une des composantes les plus importantes de l'autonomie d'un agent. Cette capacité permet à un système autonome de modifier dynamiquement sa représentation en fonction des changements de son environnement. Ce processus peut être réalisé en exploitant des représentations concurrentes. Face aux bénéfices que l'on peut attendre de l'apprentissage multi-agent concurrent, plusieurs verrous restent encore non résolus. Fulda et Ventura [69] identifient trois facteurs qui peuvent mener à un comportement peu performant du système (équilibre de la sélection, comportements individuels, masquage des actions et activation).

Le problème d'équilibre de la sélection

Le problème d'équilibre de la sélection (*equilibrium selection problem*) survient lorsque le système fait face à plusieurs solutions optimales. Dans un cadre multi-agent, cela signifie que plusieurs agents proposent des solutions optimales (ou assimilées) pour le système. Le problème est que ces solutions peuvent ne pas couvrir l'intégralité de l'espace de perception du système, qui risque de se contenter d'une solution partielle en préférant un seul agent au détriment des autres. Dans ce cas, le fonctionnement général du système n'est pas forcément sub-optimal, mais les agents peuvent avoir des problèmes dans leur coordination (s'ils convergent vers des solutions différentes). Ce problème émane directement de la définition des récompenses (et leur attribution) dans les problèmes d'apprentissage par renforcement. Ils peuvent donc être résolus en agissant sur le *feedback* du système (section 4.2.3.3), à condition de disposer d'un nombre d'interactions suffisant. Le problème reste ouvert si le système souhaite assurer une performance donnée dans un délai qui lui est accessible. Dans le cadre du contrôle des systèmes coopératifs, une solution pour augmenter le nombre d'expériences sensorimotrices du système de contrôle est de réaliser une partie des interactions avec un simulateur. La complexité de l'environnement implique néanmoins de s'assurer que les mécanismes concurrents soient suffisamment performants pour ne pas altérer la construction de la représentation du système.

Comportements individuels peu performants

Avoir des comportements individuels peu performants (*poor individual behavior*) est possible au sein d'une population d'agents. Il est alors nécessaire de mettre en place des mécanismes permettant de limiter l'apport des agents incriminés ou au contraire d'améliorer leur performance. La définition stricte de l'apprentissage par renforcement se base sur un environnement Markovien où l'état futur du système ne doit dépendre que de son état courant et non pas de la séquence d'événements qui l'ont précédé. Cette propriété n'est pas vérifiée puisqu'il est très difficile de proposer un comportement individuel optimal pour chacun des agents (dont le comportement est basé sur celui des autres) et qui couvre tout l'espace des états. Néanmoins, les propriétés des SMA (robustesse, auto-organisation) peuvent être mises à profit pour combler les faiblesses individuelles. Cela ne garantit pas que l'intégralité du problème sera traitée, car le comportement global peut être performant sur une partie du problème seulement. Le problème restant se situe alors au niveau de l'expressivité de la représentation manipulée par le système (section 4.2.3.7).

Masquage des actions

Le problème du masquage des actions (*action shadowing*) n'apparaît qu'avec les besoins de coordonner un groupe d'agents (avec un seul agent, il n'y a théoriquement pas de conflit). On parle de masquage des actions lorsque le comportement d'un agent (qui propose une action pas forcément optimale) est préféré par le système alors que d'autres actions (proposées par d'autres agents) potentiellement plus performantes sont possibles. Le problème est assez critique puisque dans ce cas, la performance du système est directement impactée sans que le système ne puisse l'améliorer. Une autre façon de voir les choses est d'imaginer que le système fait face à deux problèmes similaires mais dont les politiques optimales sont nécessaires. Garder la même politique entraîne une performance moyenne plus faible. Le système doit donc adapter sa représentation aux deux sous-problèmes. Ce problème peut être résolu en dotant le système de capacités d'adaptation, lui permettant de réutiliser des connaissances pour d'autres problèmes (section 4.2.3.6). Cependant, le nombre de sous-problèmes peut être conséquent, et donc rendre malgré tout l'évaluation des politiques difficiles.

Problème de l'activation

Grâce à l'utilisation de représentations concurrentes, on peut distinguer deux niveaux de représentation : la représentation individuelle de chacun des agents et la représentation globale, à l'échelle du système, qui est une forme émergente des représentations individuelles. Cela permet plus de robustesse et offre plus de possibilités dans la sélection et l'application de la politique. L'apprentissage peut aussi être introduit aux deux niveaux. L'apprentissage des agents leur permet d'assimiler des connaissances adaptées à de nouveaux contextes ; et l'apprentissage au niveau du système permet de sélectionner la bonne représentation (ou la combinaison des représentations individuelles) qui mènera aux récompenses optimales. Conceptuellement, un tel système pourrait évoluer face à un environnement dynamique, et faire évoluer sa représentation de l'environnement tout au long de son cycle de vie, en cherchant tout le temps à s'améliorer en s'adaptant. Malgré ces avantages attendus pour modéliser les mécanismes d'accommodation et d'assimilation (section 4.2.3.6), il reste un problème auquel faire face : quels sont les mécanismes et données minimum nécessaires au système pour réaliser cette forme d'apprentissage ? Comment gérer dynamiquement l'utilisation des représentations concurrentes pour s'adapter au contexte ? Dans le modèle conceptuel présenté dans la section 4.3, nous proposons de répondre à ces verrous à travers un mécanisme d'activation des représentations concurrentes. Le challenge est alors de proposer un processus capable de maintenir la performance de la représentation du système, sans retomber dans les problèmes précédents : masquage des actions qui peut survenir dans le cas de comportements individuels peu performants ou en cas de déséquilibre.

4.2.3.5 Verrous du constructivisme

L'intérêt pour les approches constructivistes est plutôt récent, et seuls quelques travaux ont tenté de proposer la construction d'une représentation à partir des interactions sensorimotrices [141]. Le problème fondamental rencontré par ces approches est connu sous le nom du problème d'amorçage de la représentation (*bootstrap problem*). Ce problème est lié au démarrage de la construction de la représentation de l'environnement et peut par exemple se décrire ainsi : il est nécessaire de percevoir pour mieux percevoir. Ce problème s'inscrit parfaitement dans le cadre décrit précédemment pour

l'élaboration de la représentation à base d'agents : une représentation initiale trop biaisée sera difficile à améliorer. Si l'on suit strictement les concepts du constructivisme, le système ne devrait pas bénéficier de connaissances expertes pour réaliser cette tâche. Sachant que la construction de la représentation doit suivre un processus itératif, les "briques" de base sont cruciales pour assurer le bon démarrage de ce processus. Une façon de contourner ce problème est de fournir au système quelques mécanismes de bas-niveau lui permettant de réaliser la phase d'amorçage [115, 168]. Dans ces approches, les auteurs se focalisent sur la partie discrétisation de la représentation. Grâce à des algorithmes issus des méthodes de classification, il est en effet possible de laisser le système proposer et faire évoluer une discrétisation de l'environnement. La connaissance experte se limite alors à la sélection d'hyper-paramètres. Pour cette tâche, les algorithmes "en-ligne" (*i.e.* capables d'évoluer avec de nouvelles données d'entrée, par itération) et non supervisés, sont favorisés. Dans l'approche de Kuipers *et al.* [115], le modèle utilise le Gaz Neuronal (*Growing Neural Gas* – GNG [67, 68]), qui est une évolution des cartes auto-organisées (ou cartes de Kohonen [111] – *Self-Organizing Maps* – SOM). La différence entre les deux algorithmes se situe au niveau du nombre de classes (états discrets) qui était fixe dans les SOM mais devient dynamique avec le GNG, qui introduit des conditions de génération et de suppression des neurones (représentant chacune des classes). Grâce à l'utilisation d'un GNG, le système est capable de générer lui-même une discrétisation de son environnement. Sa tâche peut donc se focaliser sur l'apprentissage de la politique d'action associée à ce découpage. L'avantage est la réduction de la dimensionalité de l'espace de la perception (données d'entrée), qui permet de faciliter l'apprentissage de la politique, en faisant l'hypothèse d'un nombre d'actions fini. La problématique du *feedback* (section 4.2.3.3) est ici présente. Mugan et Kuipers [140] montrent qu'il est possible de proposer un processus d'évaluation des prédictions et du processus de discrétisation afin d'améliorer la représentation. L'approche se repose néanmoins sur une discrétisation réalisée *a priori*, ce qui ne correspond pas intégralement au paradigme constructiviste, mais le système a la possibilité de le faire évoluer par expérience.

Faire évoluer la représentation est un premier pas vers l'adaptation. Un des travaux en ce sens propose d'intégrer des mécanismes de renforcement directement au niveau de l'algorithme de discrétisation (toujours le GNG), où le déplacement des neurones est réalisé dans l'espace des récompenses [204]. Cette approche est intéressante dans notre étude puisqu'elle montre que l'apprentissage peut être efficace même si la discrétisation sur laquelle repose la représentation évolue dynamiquement. Malgré tout, le problème du masquage des actions (*action shadowing*, section 4.2.3.4) devient très difficile à traiter dans la mesure où les neurones se spécialisent irrémédiablement vers une seule action. De la même manière, certaines zones de l'espace de perception peuvent présenter moins d'intérêt (d'un point de vue des récompenses), donc être moins discrétisées (moins de neurones), et la discrétisation souffrira d'un manque d'expressivité. Dans notre approche, nous proposons de palier ces problèmes en permettant au système de manipuler plusieurs représentations concurrentes. L'expressivité de la représentation du système sera ainsi renforcée par les représentations individuelles. Les algorithmes de classification non supervisée en ligne semblent permettre de contourner le problème de l'amorçage en gardant le caractère adaptatif de la représentation du système. Nous proposons d'intégrer ce type d'algorithme au sein d'une population interne d'agents, permettant au système de construire une discrétisation de sa perception. Ces travaux récents ont montré le potentiel de l'utilisation d'agents pour générer et faire évoluer une représentation dans une forme d'apprentissage constructiviste [131].

4.2.3.6 Accommodation et assimilation

Les processus d'accommodation et d'assimilation sont centraux dans les problématiques d'apprentissage. Ces processus garantissent l'adaptation d'un système dans un environnement dynamique. En biologie, l'accommodation réfère à la modification du comportement d'un groupe ou d'un individu face aux changements de son environnement. L'assimilation consiste à acquérir de nouvelles connaissances sur un domaine inexploré. L'analogie en IA est évidente, puisque les systèmes sont très souvent conçus pour résoudre un problème particulier (ou un type de problème) et deviennent complètement inefficients ou inapplicables face à d'autres problèmes (tout du moins, sans intervention externe comme la modification de paramètres). L'assimilation est prépondérante et vise une tâche donnée, avec peu ou pas de capacité d'adaptation. Ce problème est très connu en *Machine Learning* et même le récent développement des Réseaux de Neurones Profonds (plus connus sous le nom de *Deep Learning*) n'y a pas apporté de solution triviale. Il s'agit du problème du *catastrophic forgetting* [77] : un système qui a appris une tâche initiale et tente d'en apprendre une nouvelle oublie irrémédiablement tout ou partie de ses connaissances sur la première tâche. Malgré la complexité apparente de ce problème, le paradigme multi-agent semble une nouvelle fois apporter des éléments de résolution. Il peut alors se décomposer en trois sous-problèmes [90] : le système doit rapidement identifier le changement de contexte ; le système doit pouvoir s'adapter en peu de temps à ces changements ; et si l'accommodation n'est pas suffisante, il doit être capable d'assimiler de nouvelles connaissances afin d'apprendre une nouvelle représentation qui correspond au nouveau contexte rencontré.

Plusieurs solutions sont envisageables pour modéliser les processus d'accommodation et d'assimilation, et leurs différences se situent principalement dans la nuance donnée à la notion de représentation. Soit la représentation du système est composée par les entités internes (les agents), soit le système est composé de plusieurs représentations, chacune portée par un agent. Les deux approches permettent de représenter plusieurs contextes, mais la transition doit être assurée par des mécanismes dédiés. Il devient alors difficile de faire apprendre une nouvelle politique ou de reconnaître un changement de situation. Dans notre approche, nous faisons reposer le comportement du système sur la mise à jour de représentations concurrentes encapsulées dans des agents. Cette problématique est abordée en détail dans la section 4.2.3.4.

4.2.3.7 Expressivité de la représentation

Dans cette partie, nous discutons de deux formes de représentation : une représentation par classification couplée à l'environnement, ou une représentation statique. L'objectif d'un système doté d'une capacité d'apprentissage est de faire évoluer une représentation de son environnement afin qu'il puisse maximiser la somme des récompenses obtenues grâce à ses actions sur le long terme. La représentation est constituée d'un ensemble d'états discrets (permettant de discrétiser les situations) et d'une politique d'action associée. Cette modélisation, propre aux méthodes d'apprentissage par renforcement, permet de plus facilement distinguer les possibilités offertes au système en les limitant à un nombre fini. Une fois la représentation apprise, le système peut exhiber des compétences de prédiction, qui lui permettent de choisir la meilleure action en exploitant les compétences acquises par interaction, et intégrées dynamiquement dans sa représentation.

Représentation statique

L'approche la plus classique est d'utiliser des méthodes capables de proposer des discrétisations d'un ensemble de données. Ce sont les méthodes dédiées à la classification (*clustering*), qui peut être supervisée ou non. Elles reposent malgré tout sur des paramètres (ou hyper-paramètres) qui permettent de s'adapter à un problème donné, mais dont le choix est laissé à la discrétion du concepteur. Ces méthodes peuvent donner de très bons résultats en fournissant au système un découpage très proche de la réalité, au prix de la simplification apportée par la discrétisation (toujours discutable pour un problème réel, mais beaucoup plus compatible avec les approches statistiques). De même, la politique d'action dépend d'un ensemble d'actions données au système. Ces actions reposent sur l'utilisation d'effecteurs, dont le nombre est limité. Souvent en robotique, afin de réduire ce nombre et la combinatoire associée, des actions de haut niveau sont préférées à un contrôle très fin. La représentation ainsi proposée au système, dépend de ces limitations et peut donc souffrir d'un manque d'expressivité. C'est-à-dire que les possibilités offertes au système peuvent être bridées par conception, puisqu'il ne dispose pas de mécanismes lui permettant de les modifier.

Représentation couplée à l'environnement

Une autre approche consiste à utiliser le moins d'*a priori* possibles sur la pondération des différentes dimensions. L'ensemble doit résulter d'un apprentissage libre couplé à l'environnement. Dans ce travail, nous optons pour ce type de démarche, tout en contenant la complexité associée grâce à l'apport des mécanismes d'apprentissage concurrents. Nous proposons au système de faire évoluer sa représentation grâce à des mécanismes dédiés. Cette représentation sera portée par des agents internes, dont le comportement visera à améliorer les performances du système, et donc la précision et la capacité d'expression de sa représentation. La politique d'action du système consiste à attribuer la meilleure action possible à chacun des états de la discrétisation. Les approches classiques délèguent ce travail de discrétisation à des méthodes ou des connaissances expertes : un concepteur est chargé de proposer un découpage à partir de ses propres connaissances. Cela permet au système d'accélérer son apprentissage, mais limite grandement sa tâche. La difficulté réside dans la bonne utilisation des mécanismes d'attribution de récompense pour permettre aux agents d'évoluer dans le sens d'une meilleure représentation. Il est alors possible d'envisager que le système déclenche lui-même des mécanismes lui permettant d'adapter sa représentation [23] à de nouveaux problèmes. La multiplication des discrétisations dans une population d'agents permet plus de robustesse.

4.3 Modèle conceptuel

L'objectif de notre travail est de proposer un modèle de prise de décision capable de s'adapter dynamiquement à un problème. Le modèle est voué à être implémenté comme le processus de décision d'un système autonome. Le but du système est alors de construire itérativement une représentation de son environnement (discrétisation et politique d'action associée) et de la faire évoluer. Pour cela, nous souhaitons mettre à profit les bénéfices des systèmes multi-agents à travers l'utilisation de représentations concurrentes dans le cadre d'un apprentissage constructiviste. Le système de contrôle peut alors construire une représentation de haut niveau, composée par les

représentations partielles individuelles de sa population interne d'agents. Grâce aux mécanismes d'interactions, la représentation du système peut s'adapter à d'autres contextes. Ce travail fait face à des verrous scientifiques propres à l'élaboration d'un système cognitif dans un cadre distribué, identifiés dans la section précédente (4.2.3). L'approche se définit comme un problème d'apprentissage par renforcement, ce qui permet d'intégrer des objectifs globaux. Le modèle est présenté d'un point de vue conceptuel (sections 4.3.1 à 4.3.6), et est adaptable à des problématiques variées lors de la phase d'implémentation (sections 4.4 et 4.5).

Afin de mener à bien les objectifs de ce travail, le comportement du modèle repose sur des processus modélisés sous la forme d'interactions entre les agents internes du système. Les mécanismes mis en place visent à répondre aux problèmes identifiés : expressivité de la représentation, masquage des actions et attribution des récompenses. Pour cela, il est possible de définir différents types d'agents. Ce sont les interactions entre les agents qui mènent à l'élaboration de la représentation du système. Parmi les challenges les plus intéressants à relever, on retrouve le problème d'amorçage. Nous proposons de déléguer la phase de discrétisation à une population d'agents qui embarquent une méthode de classification. La particularité du modèle est d'utiliser plusieurs représentations concurrentes. Les agents proposent donc des hypothèses qui sont utilisées dans le processus de décision du système. Le challenge réside alors dans l'attribution des récompenses aux différents agents. Un processus d'accommodation de la représentation est proposé grâce à l'utilisation d'un autre type d'agent utilisant le concept d'activation. Enfin, les interactions entre les agents visent à améliorer la performance globale, ce qui permet une meilleure expressivité de la représentation de haut niveau du système.

4.3.1 Vue d'ensemble

En adoptant un point de vue très générique et général, le modèle décrit le processus de décision d'un système autonome capable de percevoir des données par l'intermédiaire de ces capteurs et d'agir par action de ses effecteurs dans son environnement. Une partie de sa perception est identifiée comme une récompense et sert de *feedback* à son processus d'interaction. L'objectif du système est de maximiser les récompenses obtenues sur le long terme. Son processus peut alors se décomposer en deux phases : la perception puis la décision, comme illustrées dans la figure 4.4.

Les données d'entrée, récoltées par les capteurs, constituent des observations de l'environnement. Ces observations représentent la perception de bas niveau du système. Comme dans la majorité des problèmes rencontrés par des agents situés, l'hypothèse retenue est que ces données issues de la perception de bas niveau sont continues et non identifiées. Cette volonté tient aussi de l'effort de fournir un minimum de connaissances expertes au système. Les capteurs peuvent agréger certaines variables de l'environnement, mais cette information est implicite pour le système. L'importance ou la pertinence des données ne sont donc pas connues *a priori*. Lors de la phase de perception, le système exploite une population interne d'agents "discrétiseurs". Ces agents sont en charge de proposer et de mettre à jour une abstraction des informations qu'ils traitent. Ces abstractions peuvent être vues comme des façons différentes d'intercepter le signal perçu. Chacun des agents applique une stratégie interne de partitionnement menant à la génération d'une discrétisation sous la forme d'états discrets. Ce processus peut être dynamique (suivant la méthode choisie), afin de permettre à la discrétisation d'évoluer en ajustant le périmètre des états (classes). Le processus de

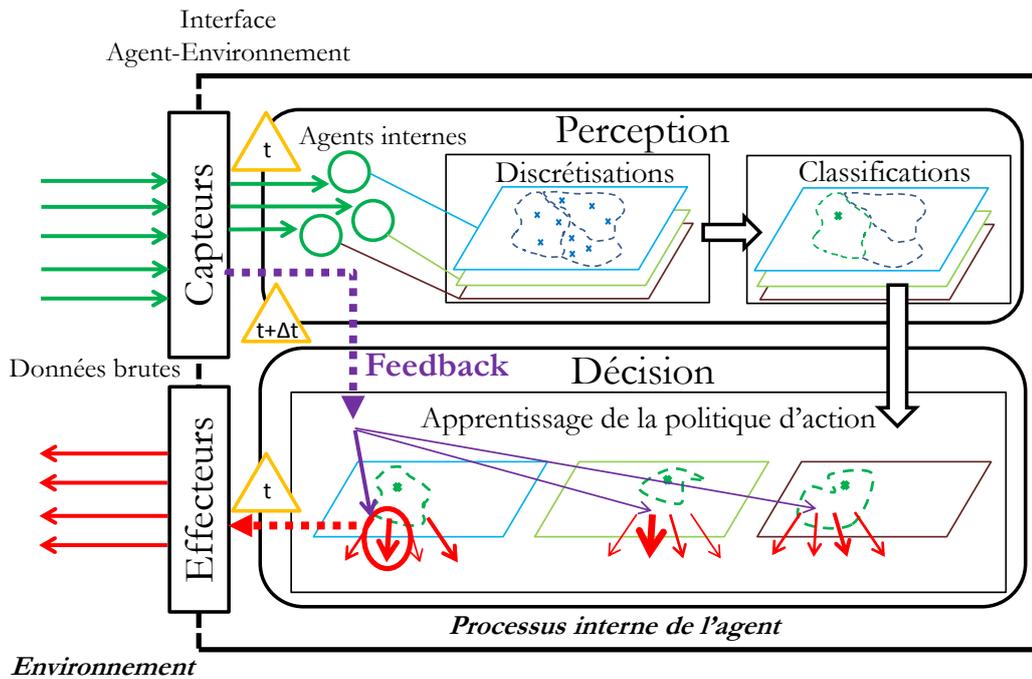


Fig. 4.4: Aperçu général du modèle d'élaboration de représentation par apprentissage.

perception multi-agent permet d'obtenir un ensemble de partitions concurrentes qui constituent des hypothèses de représentation pour le système. La sélection d'une action repose sur un processus d'apprentissage par renforcement qui utilise toutes les hypothèses d'état proposées par les agents. Le processus complet du modèle, décrit dans l'algorithme 1, exploite les états perception-actions précédemment appris afin de construire une représentation de haut niveau de l'environnement toujours plus précise.

Après une phase d'initialisation (section 4.3.2), qui permet au concepteur d'instancier les agents discrétiseurs, le processus principal de prise de décision est lancé. Le système réagit aux perceptions de ses capteurs en les partageant avec les agents. En réponse, chaque agent fournit au système une hypothèse d'état discret. Cela mène à l'élaboration d'une perception de haut niveau (section 4.3.3). Puis, le système doit décider d'une action à entreprendre dans le contexte défini par les agents discrétiseurs. Cette décision mène à l'application d'une action permettant le contrôle de l'environnement (section 4.3.4). Cette phase intègre aussi le renforcement des agents qui ont mené à la sélection de l'action.

4.3.2 Initialisation

Le modèle, présenté ici de manière générique, peut être implémenté pour différents problèmes. Le pré-requis est de pouvoir les modéliser sous la forme d'un problème d'apprentissage par renforcement. C'est-à-dire que la récompense donnée au système doit être directement liée à une fonction objectif à maximiser. Le concepteur doit alors proposer des agents discrétiseurs capables de construire une représentation précise de l'environnement, qui modélise le problème. L'approche la plus évidente est sans doute la plus efficace. Dans le cas où le domaine étudié dispose déjà de méthodes efficaces, l'idéal est d'appliquer plusieurs méthodes simultanément grâce au modèle

Algorithme 1 Processus général du modèle

```
Instancier l'ensemble  $A$  des actions possibles
Instancier l'ensemble  $D$  des agents discrétiseurs
Instancier l'ensemble  $S$  des états sélectionnés concurrents (vide)
répéter                                     # boucle qui représente le processus de décision du système
  si  $S$  n'est pas vide alors                 # condition pour la première itération
     $F = feedback(P)$                          # seule une sous partie des variables d'entrée sert de feedback
    pour tout  $s_i \in S$  faire                 # récupérer le lien  $L$  entre l'état  $s_i$  et l'action  $a^*$ 
       $L = getPerceptionActionLink(s_i, a^*)$ 
       $incrementReward(L, F)$                  # mise à jour des récompenses
    fin pour
       $reinforce(D, a^*, F)$                  # renforcement des agents ayant mené à l'action sélectionnée
  fin si
   $P = getPerception()$                        # début du processus de sélection de l'action
   $S.removeAll()$                              # la liste des hypothèses d'état est vide
  pour tout  $d_i \in D$  faire
     $p_i = discretizerPerception(P)$          # les agents peuvent percevoir différentes variables
     $d_i.discretize(p_i)$                      # chaque agent met à jour sa discrétisation
     $s_i = D_i.classify(p_i)$                  # chaque agent propose une hypothèse d'état
     $S.add(s_i)$                              # ajout de l'hypothèse d'état  $s_i$  à l'ensemble  $S$ 
  fin pour
   $a^* = explorationExploitationActionSelection(S)$ 
   $execute(a^*)$                              # exécution de l'action  $a^*$  dans l'environnement
fin répéter                               # fin du processus de décision avant la prochaine perception (et feedback)
```

en les intégrant dans des agents discrétiseurs. Ces méthodes doivent alors être basées sur un algorithme de classification capable de proposer un découpage de l'espace (ou d'un sous-espace de la perception) sous la forme d'un ensemble d'états discrets. La méthode peut être dynamique (*online* – i.e. évoluer après chaque donnée d'entrée) ou statique (i.e. baser la discrétisation sur un jeu de données préalable). Dans les deux cas, il est nécessaire de choisir quelles variables sont perçues par chacun des agents. Si un problème physique est modélisé, une bonne pratique peut être d'affecter les données d'un capteur pour chaque agent ; ce qui permet de réduire la combinatoire. Mais dans certaines configurations (peu de variables), il est possible de fournir toutes les données à plusieurs agents qui proposent des discrétisations différentes et donc concurrentes. Les méthodes employées peuvent être expertes, même si l'approche vise un apprentissage constructiviste, qui oriente plutôt vers l'utilisation d'algorithmes de classification *online* et non supervisée comme les cartes auto-organisées [111] et ses évolutions [67].

4.3.2.1 Choix du *feedback*

Conformément à notre étude des verrous, le problème du *feedback* (section 4.2.3.3) est très présent dans l'approche. Il est donc important d'apporter un soin particulier dans la sélection des variables utiles à la perception. Dans une volonté de généralité, la récompense, transmise au système et qui permet d'évaluer le résultat d'une action, fait partie de la perception. Cela permet d'utiliser toutes formes de récompenses (y compris binaires, discrètes et continues) en toute transparence. Le modèle intègre un algorithme qui gère l'attribution des récompenses. Le type de récompense peut s'adapter au type de problème rencontré. La récompense traduite doit correspondre à l'évolution d'un indicateur choisi, soit au niveau du système lui-même, ou au niveau de l'environnement. Compte-tenu de la complexité de certains problèmes, un mauvais *feedback* pourrait entraîner des faibles performances pour le système (et inversement).

4.3.2.2 Définition des actions

Si le problème visé correspond au cadre de l'apprentissage par renforcement, le nombre d'actions à disposition du système doit être fini et connu *a priori*. Cette définition stricte n'est pas obligatoire pour le modèle qui pourrait gérer des actions dynamiquement (ajout ou suppression). En revanche, l'hypothèse d'un nombre fini impose de simplifier certains problèmes (en discrétisant les actions continues). La définition des actions dans le modèle passe par la notion d'effecteur. Chaque effecteur du système disposant d'un nombre fini d'actions possibles est défini par le concepteur. Le modèle ne requiert pas une définition fine de ces actions car il manipule des actions abstraites. Tout type d'action (de la plus simple à la plus complexe) est alors envisageable. Pour des besoins de modélisation, il est par exemple possible de créer une action "ne rien faire" qui n'a aucune incidence sur l'environnement. La récompense obtenue permet au système d'évaluer le résultat de ses actions. Pour un bon fonctionnement, les actions doivent avoir une incidence suffisante pour être perçues dans la récompense.

La phase d'initialisation du modèle permet au système de générer la population d'agents discrétiseurs (et leur correspondance avec la perception de bas niveau ainsi que les actions) à partir des paramètres du concepteur. Cette initialisation n'a lieu que lors de la première mise en œuvre du système, qui évolue par la suite à la manière d'un processus de perception-action d'un agent autonome.

4.3.3 Construction d'une perception de haut niveau

Au temps t , le système perçoit son environnement par l'intermédiaire de différents capteurs. Soit Ω l'ensemble des signaux perçus par le système à travers l'ensemble de ses capteurs. Une partie de ces signaux forme la perception du système. Cette perception est représentée sous la forme d'un vecteur constitué de n valeurs continues et notée $P \in \mathbb{R}^n$. Le système cherche à construire une représentation de haut niveau de son environnement à partir des données perçues, puis décide d'une action à effectuer en réponse à cette perception, parmi l'ensemble A des k actions $\{a_1, a_2, \dots, a_k\}$ disponibles. Pour cela, il exploite différentes représentations partielles modélisées sous la forme de m agents discrétiseurs D . Chaque agent D_i dispose de sa propre méthode de classification appliquée dans son espace de perception (qui peut varier d'un agent à un autre).

Durant la phase de perception, le système a pour but de convertir le vecteur de perception P (perception de bas niveau) en une perception de plus haut niveau, sous la forme d'un état discret. Afin d'explorer l'espace de perception, une population interne d'agents discrétiseurs est en charge de proposer des hypothèses d'états. Chaque agent travaille sur des dimensions propres (*i.e.* une partition du vecteur de perception P). On note p_i la perception de l'agent D_i avec $i \in [1, m]$ tel que $p_i \subseteq P$. Cette étape de mise en correspondance de la perception fait référence à l'étape 1 de la figure 4.5.

La seconde étape permet de générer et d'identifier un contexte à partir d'une perception de haut niveau. Elle peut se décomposer en deux phases successives (ou simultanées, suivant l'algorithme utilisé) : une phase de discrétisation, et une phase de classification.

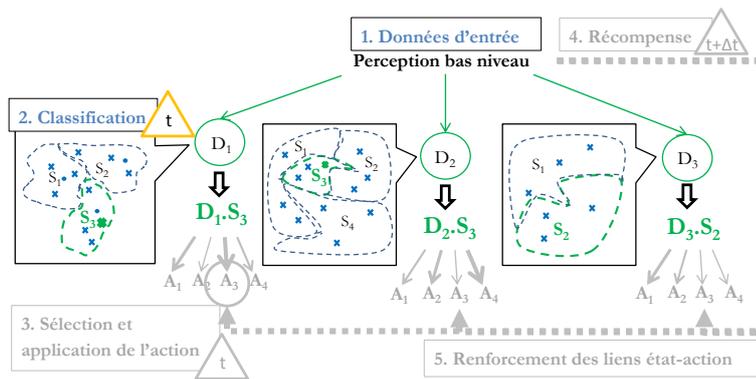


Fig. 4.5: Processus de perception visant à proposer une discrétisation de l'environnement à partir de représentations concurrentes.

4.3.3.1 Discrétisation : mise à jour des représentations individuelles

La phase de discrétisation correspond à l'actualisation des représentations individuelles (partitions) des agents discrétiseurs. Pour cela, tous les agents D mettent à jour en parallèle leur propre espace d'états potentiels (hypotheses), grâce à la méthode de classification (*clustering*) qu'ils utilisent (K-moyennes, SOM, GNG, ...). Par exemple, la discrétisation d'un agent D_i peut être composée de l'ensemble des hypothèses d'état : $\{s_1, s_2, \dots, s_j\}$. Selon cette modélisation, des configurations similaires de l'environnement correspondront aux mêmes hypothèses d'état. Le nombre d'hypothèses d'état est variable d'un agent discrétiseur à un autre et dépend uniquement des paramètres de la méthode de classification exécutée par l'agent discrétiseur considéré (ce nombre peut d'ailleurs varier au cours du temps pour certains algorithmes "en ligne"). Pour les algorithmes qui nécessitent un jeu d'apprentissage, les données utilisées peuvent être préalables à l'instanciation du modèle. Dans ce cas, la phase de discrétisation ne sera pas exploitée par l'agent en question. Seule la phase suivante de classification sera déterminante.

4.3.3.2 Classification : proposition d'hypothèses d'état concurrentes

Pour chaque perception du système, tous les agents discrétiseurs perçoivent leurs variables et classifient cette perception en un état de leur discrétisation. Le processus de classification des discrétiseurs mène à la sélection d'un seul état (un par agent). Les états ainsi classifiés par tous les agents forment l'ensemble (noté S dans l'algorithme 1) des hypothèses d'état. Chaque hypothèse représente l'état le plus probable qui représente la donnée perçue par l'agent considéré. Pour les algorithmes qui se basent sur la définition de centres de classes (c'est le cas des K-moyennes, SOM et GNG), la sélection de la classe est limitée à un calcul de distance euclidienne et au choix du centre de classe le plus proche. Ce calcul est généralement réalisé dans l'espace de perception de l'agent. Bien que cette méthode soit la plus intuitive, elle peut limiter l'expressivité de la discrétisation proposée (*e.g.* cas de données perçues sur les "bords" de l'espace de perception). Il est possible de définir d'autres méthodes de mise en correspondance des classes. Le seul impératif pour le modèle est que chacun des agents dispose d'une méthode permettant de faire correspondre une valeur d'entrée dans son espace de perception à un état discret de son découpage de l'espace. À l'issue de cette étape, l'ensemble des états S forme l'ensemble des hypothèses permettant au système de décrire le contexte auquel il est confronté. Son objectif est alors, en se basant sur son expérience, de déterminer l'action qui maximisera sa récompense immédiate dans ce contexte.

4.3.4 Contrôle de l'environnement et renforcement de la représentation

Durant la phase de décision, le système choisit une action à exécuter dans le contexte défini par la perception de haut niveau construite lors de la phase de perception. Cette action est choisie parmi l'ensemble A des k actions $\{a_1, a_2, \dots, a_k\}$ du système. Pour accomplir cette tâche, le système exploite une nouvelle fois les éléments disponibles au niveau des agents discrétiseurs, qui décrivent ses expériences sensorimotrices. Le choix de l'action est alors déterminé par une politique d'exploration et d'exploitation des représentations individuelles des agents discrétiseurs.

4.3.4.1 Représentation des expériences sensorimotrices

Chaque hypothèse d'état générée par les agents discrétiseurs est liée à toutes les actions du système afin de former des liens perception-action. Ces liens représentent les différentes expériences sensorimotrices du système. L'ensemble des interactions sensorimotrices possibles est alors défini comme l'ensemble $L = S \times A$, représentant tous les couples $(s_{i,i=1,j}, a_{j,j=1,k})$ qui combinent sous la forme d'un lien perception-action les hypothèses d'état $(s_{i,i=1,j} \in S)$ produites par les m agents aux k actions possibles $(a_{i,i=1,k} \in A)$. Afin de garder l'approche générique, chaque lien perception-action peut être enrichi de paramètres : $Param_l = p_1, p_2, \dots, p_q$; un ensemble de q paramètres associé à chaque lien $(l \in L)$. Ces paramètres permettent aux agents discrétiseurs de mémoriser les expériences des interactions du système réalisées en utilisant leur représentation.

4.3.4.2 Un dilemme exploration-exploitation

L'objectif du système est d'explorer l'espace des actions (liens états-action) afin d'exploiter celles qui donnent les meilleures évaluations dans le contexte perçu. Ce processus correspond à l'étape 3 de la figure 4.6.

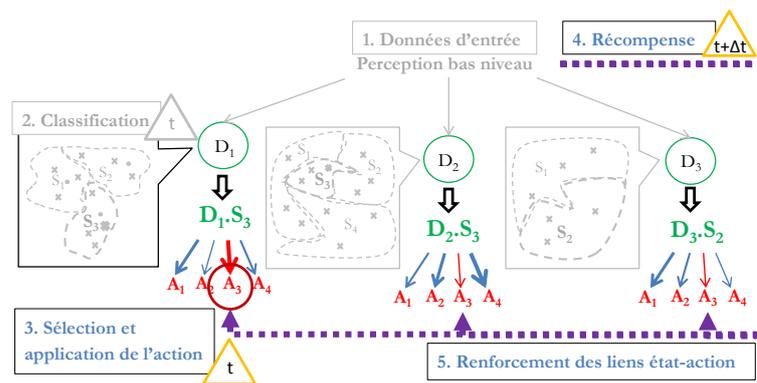


Fig. 4.6: Processus de sélection de l'action et renforcement des représentations concurrentes.

Le processus de sélection mène au choix d'une action notée a^* . La particularité de notre approche est qu'après la phase de perception, le système doit choisir parmi des liens (couples associant un état à une action). L'ensemble de ces liens est dynamique : à chaque itération, les agents discrétiseurs fournissent des hypothèses différentes si le contexte change. Afin d'affiner sa représentation, le

système doit explorer l'ensemble des possibilités disponibles tout en exploitant les liens qui vont mener vers une bonne évaluation moyenne des actions exécutées. Ce problème est très comparable au dilemme d'exploration-exploitation, aussi connu sous le nom du problème du bandit à n bras [9].

La définition originale du problème du bandit à n bras se base sur un exemple concret. Un joueur dans un casino est confronté à un ensemble fini de machines à sous qui ont toutes des récompenses moyennes (espérance) inconnues. L'objectif du joueur est de choisir une stratégie lui permettant de maximiser les récompenses (ses gains) à plus ou moins long terme. Il est donc nécessaire d'essayer les différentes machines (exploration) avant de choisir la (ou les) machine(s) qui rapporteront le plus (exploitation). Les propriétés de convergence de ce type d'algorithme ont été largement étudiées : la performance des modèles classiques est garantie après un nombre maximum fini d'essais [60]. Bien que les algorithmes s'expriment très bien dans un contexte d'apprentissage par renforcement, une question persiste au niveau de leurs propriétés de convergence. Elle fait référence aux problèmes de sur-apprentissage : à quel moment un agent peut-il déclarer que la politique d'action découverte est (presque) optimale [61], et ainsi stopper son processus d'apprentissage [30] ? Dans notre travail, le système est en interaction permanente avec l'environnement, ce qui limite les effets de ce problème. L'important est de s'assurer que sa politique d'action puisse converger, mais de permettre que l'apprentissage continue afin que la politique s'adapte aux nouvelles données d'entrée.

Dans le cadre de notre approche, les machines sont ici les possibilités offertes au système à chaque étape de décision. Il s'agit des liens sensorimoteurs (l'ensemble L). Plutôt que de choisir directement une action, le système choisit en réalité un lien noté $l^* = (s^*, a^*)$ où s^* et a^* correspondent respectivement à l'hypothèse d'état et à l'action sélectionnés. Toujours dans un souci de généralité, le choix de l'algorithme d'exploration-exploitation est laissé à la discrétion du concepteur lors de l'implémentation du modèle. Les valeurs associées au lien l^* ($Param_{l^*}$) permettent d'ajouter des paramètres relatifs à l'algorithme choisi (e.g. un nombre de sélections, ou une récompense moyenne).

Dans la suite de notre travail, nous avons sélectionné l'algorithme nommé *Upper Confidence Bound* [8] (UCB). Cette méthode garantit des propriétés de convergence d'un point de vue théorique [102]. Lors de la phase de décision du système, le lien l^* sélectionné est celui qui maximise :

$$UCB(l) = \bar{x}_l + \sqrt{\frac{2 \ln n}{n_l}} \quad (4.1)$$

où :

- \bar{x}_l est la moyenne estimée des récompenses du lien perception-action l ;
- n_l est le nombre total de sélections du lien l ;
- n est le nombre total d'interactions sensorimotrices du système.

Les paramètres de chaque lien sont ainsi les paramètres de l'algorithme UCB. Les liens portent donc un nombre de sélections (n_l), une somme des récompenses (r_l) et le système comptabilise le nombre total d'interactions (n). Ces paramètres sont mis à jour à chaque pas de décision. D'autres algorithmes d'exploration-exploitation plus complexes pourraient être utilisés (comme le *Thompson sampling* par exemple). Notre choix de UCB est motivé par la garantie de convergence en un temps limité. La récompense utilisée dans l'algorithme se base sur le *feedback* fourni par l'environnement. Cette récompense participe au renforcement de la représentation en étant propagée aux liens ayant participé à la sélection de l'action a^* .

4.3.4.3 Renforcement de la représentation

Afin d'augmenter sa performance (ce qui est lié à l'optimisation des récompenses reçues), le système doit améliorer sa représentation. Pour cela, chaque action décidée est évaluée par une fonction de l'environnement sous la forme d'une récompense. Grâce à l'utilisation d'un algorithme d'exploration-exploitation, le système est capable de converger vers une politique d'action efficace. La représentation du système ne se limite cependant pas à la seule politique d'action. Dans le cas contraire, chacun des agents discrétiseurs pourrait optimiser sa propre politique en fonction de sa représentation, et ce indépendamment des autres. L'originalité de notre approche se situe dans le mécanisme d'attribution de récompense qui ne limite pas l'effet du *feedback* à la partie apprentissage de la politique. En effet, nous cherchons aussi à proposer un renforcement de la discrétisation ayant mené à l'application de l'action α^* choisie par le système.

Au temps $t + \Delta t$ (*i.e.* Δt après l'exécution de l'action), le système reçoit, par l'intermédiaire de sa perception, une évaluation (*feedback*) du résultat de l'action précédemment exécutée. Cette évaluation, notée F , est une sélection de variables de l'ensemble de perception Ω au temps $t + \Delta t$ (ce sont toujours les mêmes valeurs qui sont sélectionnées). On a ainsi $P \cup F \subseteq \Omega$ (étape 4 sur la figure 4.6). Afin de renforcer la représentation du système grâce à cette récompense, une possibilité est de renforcer directement le lien l^* sélectionné par l'algorithme d'exploration-exploitation. Le problème est que cela risque d'entraîner une différence de traitement entre les agents (et mener au masquage de certaines actions, section 4.2.3.7). Certains pouvant être plus souvent sélectionnés que d'autres. Comme le nombre d'états entre les discrétisations des discrétiseurs peut varier, ce problème est d'autant plus présent. Un agent avec moins de classes serait alors plus exploré qu'un autre. Nous faisons l'hypothèse que chacun des agents doit avoir sa chance dans le processus de sélection. Le nombre de possibilités offertes au système à un pas de temps de décision donné est donc constant. Il s'agit de l'ensemble L composé du nombre de liens (hypothèses) émanant de chacun des agents (un par agent) fois le nombre d'actions disponibles.

Le processus d'attribution de la récompense se base sur l'action α^* choisie par le système (étape 5 de la figure 4.6). Cette action est finalement évaluée par l'intermédiaire du *feedback*. Tous les agents ayant participé à l'élaboration de la perception de haut niveau et qui auraient choisi la même action sont récompensés. La récompense est partagée (sans être divisée), à tous les agents qui ont sélectionné un lien dont l'action la plus probable était la même action que celle choisie par le système. L'évaluation de l'action la plus probable ne dépend que des liens d'un agent donné émanant de l'état sélectionné pendant la classification par l'agent en question. En d'autres termes, il s'agit d'évaluer, en faisant l'hypothèse que l'agent était seul dans le système, le résultat de l'algorithme d'exploration-exploitation appliqué sur les k liens disponibles (exactement autant que d'actions). Cette probabilité de sélection dépend encore une fois de l'algorithme utilisé et des récompenses (paramètres des liens) manipulées. Un bon exemple est d'utiliser la récompense moyenne de chacun des liens (somme des récompenses divisée par le nombre de sélection).

Grâce à ce mécanisme, l'espace de recherche (phase d'exploration) est réduit. Le modèle profite en réalité des avantages de l'apprentissage concurrent [78] sans introduire de biais au niveau de l'algorithme d'exploration-exploitation qui utilise toujours un nombre constant de possibilités. La différence avec l'algorithme original est que les "bras" disponibles sont variables d'une itération à l'autre. Cela permet d'optimiser un parcours des états discrets tout en permettant un apprentissage de la politique d'action. L'autre originalité de l'approche est d'attribuer des récompenses pour

une interaction sensorimotrice complète. Cela veut dire que les deux processus d'apprentissage (discrétisation et politique d'action) sont couplés. Il n'y a pas de distinction entre les deux processus. La représentation à l'échelle du système peut alors converger vers une combinaison dynamique des meilleures discrétisations et des meilleures politiques, apprises simultanément. Grâce à la robustesse offerte par l'utilisation de plusieurs agents composant et faisant évoluer la représentation, celle-ci est aussi dotée de capacité d'adaptation. Toutes ces propriétés sont validées dans un cadre expérimental simple (section 4.4) avant d'être exploitées dans un cadre applicatif plus complexe (section 4.5). Ces bénéfices sont toutefois contrebalancés par la précision des discrétisations manipulées par les agents discrétiseurs. Le modèle peut alors être étendu pour améliorer cette étape du processus d'apprentissage.

4.3.5 Processus d'amélioration de la discrétisation par associations

Dans notre modèle, la notion de représentation est disponible à deux niveaux. La représentation du système est composée d'une combinaison dynamique des représentations individuelles de la population d'agents discrétiseurs. Grâce aux mécanismes d'apprentissage concurrent et au processus d'attribution de récompense, les performances du système peuvent surpasser les performances individuelles. Cependant, le problème d'amorçage de la représentation impose au système de disposer de représentations individuelles suffisamment performantes. De même, à l'issue de la convergence de l'algorithme d'exploration-exploitation, la performance du système ne peut plus évoluer si la représentation reste stable. Pour pallier ce problème, nous avons conçu un processus permettant d'agir sur l'étape de découpage proposée par les discrétiseurs. Le challenge est d'atteindre un équilibre entre les ressources mises en œuvre pour trouver de nouveaux états et le temps nécessaire au système pour les explorer. Si la création de nouveaux états mène à l'amélioration des performances du système, il faut les conserver ; dans le cas contraire, il faut minimiser leur impact sans pour autant risquer de réitérer les mêmes essais.

L'approche repose sur la proposition d'une seconde population d'agents internes. Ces agents sont nommés "associateurs" : il sont en charge de générer de nouveaux états en associant les états disponibles des agents discrétiseurs. Afin de maintenir le cadre défini par le modèle, et conserver ses propriétés, les associateurs sont eux-même considérés comme des discrétiseurs. Ils peuvent en effet proposer des hypothèses lors de la phase de classification. À la différence des discrétiseurs, le nombre d'hypothèses proposées à chaque itération n'est pas limité. Enfin, les différences entre agents associateurs se situent au niveau de la méthode d'association et du niveau d'association.

4.3.5.1 Les agents associateurs

Le niveau d'association permet de générer des états en cascade. Un agent associateur de niveau 1 crée des états en associant les états des agents discrétiseurs (considérés comme le niveau 0). Un agent associateur de niveau n crée des combinaisons d'états des niveaux inférieurs (de 0 à $n - 1$). La méthode d'association définit la manière avec laquelle un agent associateur génère un nouvel état à partir de deux états provenant des découpages d'autres agents. Dans une première étape, nous considérons seulement les associations binaires, et plus particulièrement l'intersection (l'union est aussi facilement envisageable). Par exemple, si un état composé C caractérise l'union d'un état

A d'un discrétiseur D_1 et d'un état B de D_2 , alors l'état C sera sélectionné par l'agent associateur si D_1 classe sa perception en l'état A et si l'agent D_2 choisit B comme état représentant ses données d'entrée dans son espace. Même sur cet exemple basique, on se rend compte qu'en procédant aussi simplement, l'agent associateur est capable de générer de nouveaux états et de les intégrer dans la représentation du système. La particularité est que l'agent associateur va travailler sur un nouvel espace de perception qui sera une combinaison des espaces manipulés par les agents discrétiseurs sources des états utilisés. Cela ouvre la voie vers une amélioration de l'expressivité de la représentation. En effet, il est possible que certains phénomènes survenant dans l'environnement ne soient observables que dans certaines combinaisons de dimensions (variables) ; ils peuvent être occultés par d'autres variables. Le coût de l'ajout de nouveaux états est très faible pour le système. En revanche, la tâche d'exploration risque d'être grandement allongée si le nombre d'états disponibles croît. Bien que le mécanisme d'attribution de récompense limite cette augmentation, il est nécessaire de mettre en place des mécanismes visant à limiter l'usage d'états issus d'associations (états agrégés).

4.3.5.2 Activation des états agrégés

Pour réguler le nombre d'états disponibles à chaque itération du système, nous proposons d'introduire la notion d'activation des états. Ainsi, les états des agents discrétiseurs (niveau 0) sont systématiquement activés au démarrage du système et restent toujours actifs. Ils servent en effet de base d'amorçage de la représentation du système. Seuls les états activés sont disponibles au moment des phases de classification et de sélection de l'action. Toujours lors de la phase d'initialisation, les agents associateurs créent, en suivant leur méthode d'association, tous les états agrégés possibles (compte tenu de leur niveau). Par défaut, les états ainsi générés sont tous activés. Par la suite, leur activation (donc intégration dans le processus d'apprentissage) est sujet à leur utilisation. Les états menant à la même action sont récompensés de la même manière lors de l'attribution de la récompense. Afin de choisir des états agrégés dont l'usage est susceptible d'améliorer la représentation du système, leur activation dépend de la notion de conflit. Un conflit se définit entre deux états distincts. Un conflit entre un état A et un état B survient lorsque l'action de A est sélectionnée par le système et qu'elle diffère de l'action la plus probable pour B . Ce cas de figure est intéressant car il indique, après convergence, des zones (dans les espaces respectifs des agents discrétiseurs), où une amélioration est possible. En effet, l'objectif du système est de définir finement le périmètre des actions dans l'environnement. Sa représentation idéale devrait alors être composée d'un nombre fini d'états qui prédisent exactement les meilleures actions, pour l'ensemble de la perception du système. Afin de se rapprocher de cet objectif, le système génère des états agrégés et gère leur activation en maintenant une durée d'activation basée sur les conflits entre les états. La définition sous forme d'agents distribués autonomes rend difficile la présentation de ce processus. Pour simplifier, imaginons que les agents (associateurs et discrétiseurs) partagent une matrice des conflits, comme représentée dans la figure 4.7.

Les conflits ne sont pas symétriques pour deux raisons. D'abord, l'évolution de la politique d'action peut faire varier significativement cette valeur. Et les agents travaillent dans des sous-espaces de la perception possiblement différents. Une donnée perçue par un agent pourra correspondre à une position similaire de la perception alors qu'il s'agira de deux états complètement distincts pour un autre agent. En réalité, la notion de conflit n'existe que dans les états agrégés (la notion de matrice est totalement conceptuelle). Il stockent le nombre d'interactions où l'action la plus probable a différé des autres états sélectionnés (parmi l'ensemble des hypothèses L).

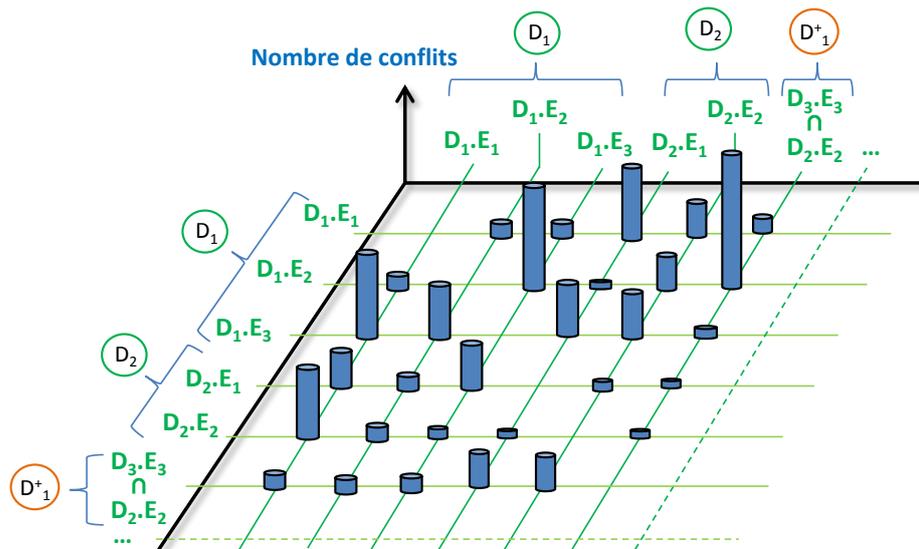


Fig. 4.7: Illustration simplifiée des conflits entre les états sous la forme d'une matrice de conflits. Les agents associés sont notés D^+ .

Chaque état agrégé dispose des propriétés suivantes :

- une référence vers les états sources : qui permet à l'associateur de déterminer si l'état agrégé est sélectionné, suivant sa méthode d'association ;
- une durée d'activation : qui est une réserve d'interactions (nombre) durant lesquelles l'état restera actif.

La durée d'activation est modulée dynamiquement grâce à trois paramètres du modèle :

- une valeur de décroissance : qui correspond à la consommation du temps d'activation pour chaque interaction du système où l'état agrégé a été sélectionné ;
- un bonus d'usage : qui correspond à un gain de temps d'activation dans le cas où l'état agrégé a été directement sélectionné par l'algorithme d'exploration-exploitation ;
- un bonus d'activation : qui garantit un nombre minimal d'interactions lors de la première activation de l'état agrégé.

Lors de la phase d'implémentation, plusieurs stratégies d'activation sont envisageables. Les plus efficaces sont de concentrer les efforts du système sur les états agrégés qui proviennent d'états sources fortement en conflit. L'apport du nouvel état doit pouvoir réduire ces conflits. Un état agrégé plus précis améliore la représentation à l'échelle du système. Il est donc plus souvent sollicité, et reste actif si sa performance (indirecte) est plus importante que celle produite par ses états sources. Ce comportement permet d'améliorer les performances du système en limitant l'augmentation du temps d'exploration lors de la phase d'apprentissage (section 4.4).

4.3.6 Perspectives

Comme l'a montré l'étude des verrous scientifiques des approches constructivistes et de l'apprentissage à base de représentations concurrentes (section 4.2.3), le problème visé par le modèle

est très complexe. Nous avons proposé à travers l'utilisation d'une population d'agents internes portant des représentations individuelles concurrentes, un modèle capable de générer un processus de construction itératif d'une représentation de haut niveau de l'environnement. Le modèle est directement confronté à une problématique d'attribution de la récompense à laquelle nous proposons de répondre par l'utilisation d'un processus d'exploration-exploitation mettant à profit des techniques d'apprentissage concurrent. En plus de ce fonctionnement nominal, nous avons proposé, via la notion d'association, des améliorations visant certains des verrous présentés : adaptation, expressivité, ou encore activation (section 4.3.5). Tous ces mécanismes sont étudiés dans les expérimentations en simulations présentées dans la section 4.4, qui expose la mise en place d'un simulateur dédié. A ce stade de maturité du modèle, beaucoup d'idées restent à intégrer pour améliorer le comportement global du système. Ces propositions s'inscrivent pleinement dans la dynamique constructiviste explorée.

4.3.6.1 Vers une catalyse de l'adaptation pour la discrimination de concepts

Nos propositions pour améliorer (affiner) la représentation du système passent par la notion d'activation d'états discrets encapsulés dans les représentations concurrentes (les agents discrétiseurs). Dans la version actuelle du modèle, tous les états (proposés par les discrétiseurs ou par les associeurs) se trouvent au même niveau dans la représentation. C'est-à-dire que leur poids dans la construction est identique quelle que soit leur taille par exemple. Le fonctionnement de la discrimination des contextes paraît un peu différent : les contextes peuvent être hiérarchisés : du moins précis au plus précis. Le système qui fait face à une nouvelle perception va donc d'abord dégrossir la discrimination de sa perception de haut niveau en catégorisant grossièrement. Ensuite, si le système continue à être confronté à un contexte similaire, il pourra se concentrer sur une discrétisation encore plus précise, en distinguant de nouveaux sous-contextes par exemple. Le modèle actuel est capable de discriminer différents contextes (section 4.4.3) mais cette notion de hiérarchie n'existe pas. Nous proposons donc de la modéliser en intégrant la chaîne d'activation des états agrégés dans le processus de sélection de l'action du système. Grâce à des mécanismes similaires à ceux en place dans les agents associeurs, notre proposition repose sur l'intégration de la suite d'activation d'états (suivant les niveaux des agents associeurs) pour catalyser l'exploration de certaines zones de la perception. Il faudra alors proposer des paramètres permettant de déclencher une chaîne d'activation afin que la transition entre différents contextes soit plus rapide et permette au système apprenant de se focaliser plus rapidement sur l'assimilation des contextes inédits, tout en exploitant les contextes appris.

4.3.6.2 Vers une forme de meta-apprentissage

Le meta-apprentissage pourrait se résumer par "apprendre à apprendre". Dans le cadre d'un apprentissage d'un agent autonome, l'objectif est de reproduire la capacité des systèmes biologiques à améliorer non pas seulement leur représentation, mais aussi leur capacité d'assimilation. Pour cela, l'apprenant fait preuve de capacités de généralisation qui vont au delà de la comparaison de contextes. C'est bien le processus complet d'apprentissage d'une tâche qui est intégré en tant que connaissance pour améliorer l'assimilation future de nouvelles tâches. Reproduire ce comportement semble difficile dans la mesure où la notion de représentation proposée dans notre modèle ne permet pas de modéliser ce type de connaissance. La clé de ce verrou se situe probablement dans la généralisation de la notion de représentation. Une façon de généraliser la représentation

est d'utiliser une nouvelle fois les agents associateurs dans le processus de discrimination du contexte. Par exemple, nous proposons de créer une nouvelle forme d'agents capables d'observer le comportement des associateurs mais aussi d'évaluer le processus général d'apprentissage du système (donc d'accéder aux représentations concurrentes des agents discrétiseurs). Le but de ces nouveaux agents est d'exploiter les méthodes d'agrégation portées par les agents associateurs qui sont les plus prometteuses. Pour cela, il est nécessaire de proposer un indicateur qui permette de déterminer les associateurs qui contribuent le plus à améliorer la représentation. De plus une forme d'exploration (par exemple, un algorithme génétique) pourra faire varier les méthodes d'agrégation (donc les associateurs) afin de trouver de meilleurs candidats qui amélioreront encore la représentation. Cette proposition, bien que prometteuse, se heurte une nouvelle fois au problème d'attribution de la récompense et à l'explosion combinatoire attendue qui ne sera peut être pas compatible avec beaucoup de problèmes réels.

4.4 Étude du modèle

Le modèle proposé a pour objectif d'être implémenté en tant que processus de décision d'un système autonome en interaction avec son environnement. L'apprentissage multi-agent lui permet de construire une représentation, renforcée par une récompense immédiate, et assurant un meilleur contrôle. Afin d'évaluer le comportement du modèle, qui dépend des mécanismes proposés, nous souhaitons étudier ses réactions dans différents contextes applicatifs. Dans cette section, nous proposons un cadre expérimental permettant de maîtriser le comportement d'un environnement complètement virtuel. Ce prototype expérimental (section 4.4.1) permet de générer différents types d'environnements tout en comparant différentes implémentations du modèle. La définition générique des agents permet d'utiliser différentes méthodes de classification, qui s'adaptent au problème rencontré. Lors de la description du modèle, des mécanismes ont été mis en place afin de répondre aux différents problèmes identifiés. Ces mécanismes mettent en avant l'apport de l'apprentissage concurrent, de la capacité d'adaptation ou encore d'une méthode d'affinage de la représentation. Tous ces mécanismes sont étudiés grâce à des simulations dédiées (sections 4.4.2 à 4.4.4). Il est important de vérifier que ces mécanismes fonctionnent et ne s'accompagnent pas d'effets de bord indésirables avant de proposer une implémentation dans le cadre applicatif du contrôle intelligent d'un flux de véhicules connectés.

4.4.1 Un prototype expérimental

Pour mener l'étude du comportement du modèle, nous proposons un cadre applicatif expérimental. Ce prototype se présente comme un outil de simulation dans lequel il est possible d'instancier un agent autonome (le système) qui a pour objectif d'interagir avec un environnement virtuel dans le but de maximiser les récompenses obtenues. Le prototype permet de définir différents types d'environnements continus composés d'un nombre variable de dimensions. Le système est modélisé comme un agent autonome dont le cycle de vie est régulé par ses interactions avec l'environnement. Le processus interne du système repose sur une implémentation du modèle présenté dans notre travail.

4.4.1.1 Simulation de l'environnement

L'apport principal du prototype se situe au niveau de la définition d'un environnement virtuel. Le système, dont le processus de décision est gouverné par une instanciation de notre modèle, interagit avec cet environnement simulé. Le processus d'apprentissage vise à améliorer les récompenses obtenues de l'environnement lors de l'application de la politique d'action intégrée à la représentation construite. L'objectif de la simulation est de montrer, dans un premier temps, les bénéfices du modèle en faisant abstraction des biais possibles d'un cadre applicatif plus complexe.

Le contrôle d'un environnement complexe : un problème exigeant

Dans tous les problèmes "réels", une des difficultés rencontrées par les systèmes autonomes est qu'il n'existe pas de représentation absolue du monde. Même dans les problèmes plus théoriques, cette représentation absolue n'est pas accessible au système. Grâce à un environnement virtuel, il est possible de définir une représentation absolue. Bien sûr, celle-ci reste inaccessible au système que l'on cherche à évaluer, mais il devient possible de comparer la représentation construite par le système avec cette représentation parfaite. Il faut avoir conscience qu'un écart important risque d'exister entre l'environnement virtuel et la définition de problèmes plus complexes. Pour limiter ce *reality gap* [96], lorsque le système subira le passage du monde virtuel au monde réel, il est nécessaire de proposer un environnement relativement réaliste. Cela peut passer par l'adjonction de bruit dans les signaux manipulés par le système. Le cadre d'un environnement parfait garde néanmoins son intérêt pour vérifier le bon fonctionnement de certains mécanismes sans subir de biais.

Modèle de l'environnement

L'environnement virtuel simulé se décompose en deux parties :

- une partition fixe (discrétisation) de l'espace d'entrée (sous la forme d'états discrets) ;
- une politique d'action fixe associée (chaque état discret correspondant à une action).

Le modèle de la partition de l'espace se base sur la définition de la position de centres de classes. Lorsqu'une nouvelle donnée de perception (dans l'espace complet) est générée, cette valeur est assimilée par l'environnement à l'un des états discrets. L'état choisi est celui qui est le plus proche de la position de la donnée (au sens de la distance euclidienne avec les centres de classes). Même si cette définition est simple à mettre en place, elle permet toutefois de générer rapidement des environnements relativement complexes. Le périmètre des états de l'environnement est linéaire par morceau ; l'environnement est défini comme un pavage à la manière d'un diagramme de Voronoï (comme illustré dans la figure 4.8).

Trois variantes de génération de l'environnement virtuel

Parmi les paramètres permettant de générer l'environnement, il est possible de choisir ses dimensions (*i.e.* le nombre de variables de l'espace d'entrée), le nombre d'actions possibles, et la méthode de génération de la discrétisation. Actuellement, trois types d'environnement peuvent être générés :

- un **environnement "grille"** où les centres de classes de la discrétisation sont organisés en grille, chaque action correspondant à une classe (autant de classes que d'actions). Les obser-

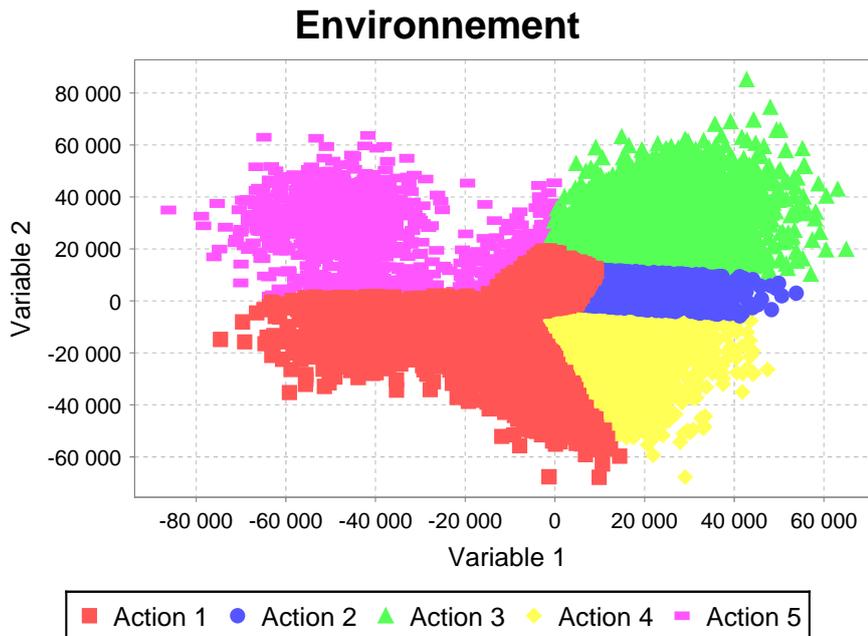


Fig. 4.8: Exemple d'environnement virtuel en deux dimensions généré dans le prototype expérimental. La discrétisation est composée de 6 états associés à une politique de 5 actions possibles (représentées en couleurs). Chaque point correspond à une observation générée par l'environnement.

vations (données d'entrée du système) sont générées de manière uniforme sur l'ensemble de l'espace (borné) ;

- un **environnement aléatoire** où les positions des centres de classes suivent une distribution gaussienne (centrées uniformément). Les observations suivent aussi une distribution gaussienne.
- un **environnement interpolé** linéairement qui permet de créer un contexte mélangé entre deux environnements aléatoires. Les positions des classes sont interpolées linéairement entre les deux discrétisations choisies (constituées du même nombre de classes). La politique d'action ne change pas (elle reste dépendante des nouvelles positions des classes).

4.4.1.2 Cycle d'exécution du simulateur

L'objectif du simulateur est de reproduire un cadre d'interaction possible du système avec un environnement virtuel. Comme il s'agit d'une simulation, le nombre d'itérations est fixé à l'avance. Chaque itération correspond à un cycle complet d'interaction du système avec l'environnement. L'environnement génère une observation : un vecteur de valeurs de taille égale au nombre de dimensions de l'environnement. Cette observation est ensuite transmise au système qui l'utilise comme sa perception de bas niveau. Le système, dont le processus de décision repose sur une implémentation du modèle conceptuel, doit décider d'une action à appliquer dans cet environnement. A l'issue de cette phase de décision, l'action est transmise à l'environnement qui doit l'évaluer.

La figure 4.9 montre le résultat de la politique d'action d'un système composé d'une implémentation du modèle avec deux agents (en bas à gauche). L'objectif du système est de construire une représentation qui soit la plus proche possible de la représentation absolue de l'environnement (en haut à gauche) : c'est ce qui lui permettra de maximiser les récompenses obtenues lors de

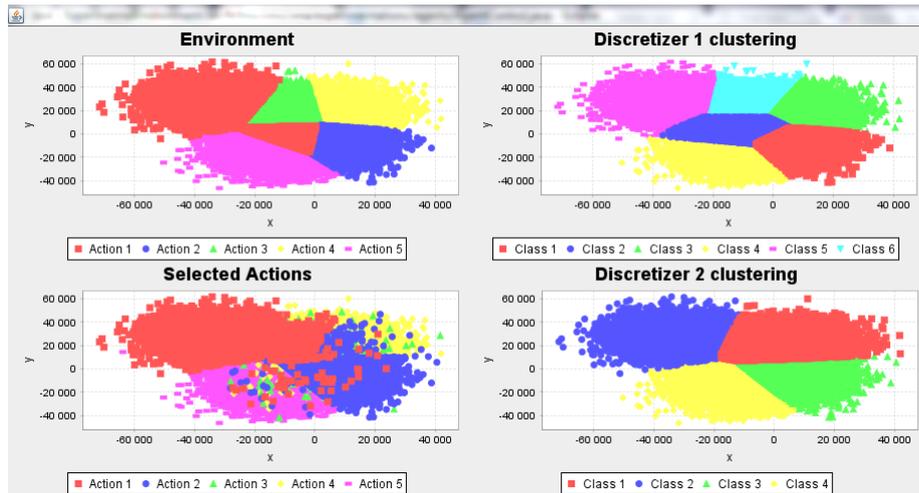


Fig. 4.9: Capture d'écran du prototype expérimental montrant le résultat d'une simulation. Par colonne et de haut en bas : représentation de l'environnement, actions effectuées par le système ; discrétisation proposée par les deux agents discrétiseurs implémentés dans le modèle. Les couleurs entre les graphes de gauche et de droite ne sont pas corrélées.



Fig. 4.10: Capture d'écran du prototype montrant les actions résultant de la politique du système comparées aux actions attendues (même simulation que dans la figure 4.9).

l'évaluation du résultat des actions. Pour cela, il met à profit des représentations individuelles concurrentes des agents discrétiseurs (graphes de droite dans la figure 4.9).

4.4.1.3 Évaluation de la performance du système

L'environnement compare l'action choisie par le système à l'action de sa représentation absolue (*i.e.* celle associée dans la politique d'action à l'état correspondant à l'observation précédemment générée). La récompense associée à une bonne action (celle attendue) ou une mauvaise action (une autre) est binaire (respectivement 1 ou 0). Nous faisons l'hypothèse que cette fonction de *feedback* est parfaite (sans bruit), dans la mesure où la représentation de l'environnement virtuel suit la même hypothèse. Comme la représentation absolue de l'environnement est connue, il est possible d'estimer la performance du système. D'un point de vue extérieur, la performance du système peut être observée en suivant l'évolution de la proportion du nombre de bonnes actions au fil des interactions du système.

La figure 4.10 montre les actions attendues par l'environnement et celles données par la politique d'action de la représentation d'un système simulé. Dans cet exemple, la représentation du système décrit bien l'action n°5 mais échoue à déterminer le périmètre de l'action n°3. Sa performance moyenne ne sera donc pas maximale. Un système qui disposera d'une représentation parfaite pourra maximiser cet indicateur. La vitesse de l'apprentissage est aussi observable en traçant l'évolution de la performance au cours du temps : il est alors possible de visualiser le degré de convergence de la représentation du système.

4.4.2 L'apport des représentations concurrentes

Le modèle proposé dans notre travail utilise des discrétisations existantes ainsi que des politiques d'action associées afin de construire une représentation des expériences sensorimotrices d'un système autonome. Le bénéfice attendu de la combinaison de représentations concurrentes est une amélioration de l'expressivité de la représentation du système. Nous souhaitons évaluer l'apport de représentations partielles dans la précision et la rapidité de construction d'une représentation de haut niveau. Les expérimentations qui suivent se focalisent sur la performance atteinte par le système et le temps de convergence nécessaire lors de l'apprentissage. L'environnement utilisé est généré suivant le modèle grille (comme illustré dans la figure 4.11 (a)).

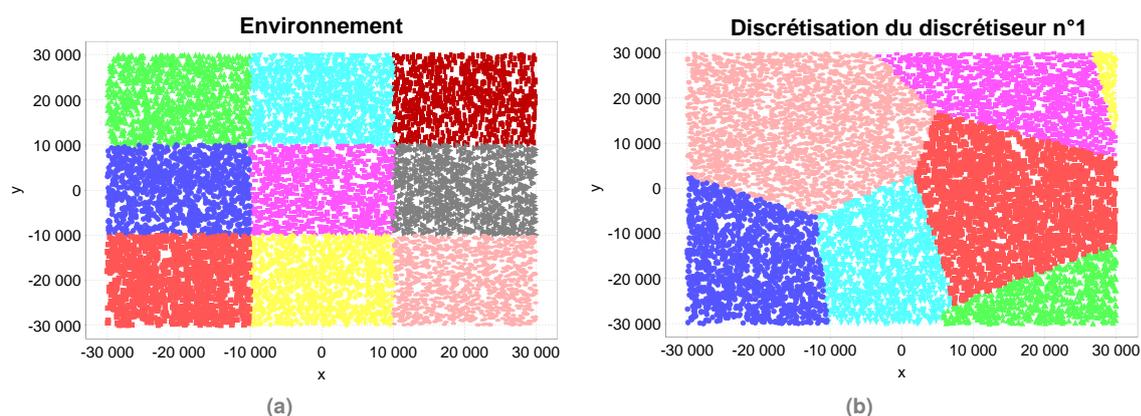


Fig. 4.11: (a) Aperçu de l'environnement. Chaque couleur représente une action attendue différente.
 (b) Exemple de découpage d'un agent discrétiseur utilisant 7 classes. Chaque couleur associe une observation à un état discret.
 Les couleurs entre les deux graphiques ne sont pas corrélées.

Les observations sont générées depuis une distribution uniforme dans un espace à deux dimensions. Les 9 actions disponibles sont disposées pour former une grille 3 par 3 où chaque état correspond à une action distincte. Les agents discrétiseurs implémentés dans le modèle de contrôle utilisent des discrétisations aléatoirement générées (les centres de classes sont placés suivant la même distribution). Tous les agents discrétiseurs perçoivent toutes les dimensions de l'environnement. Chaque représentation individuelle est donc directement en concurrence avec les autres. Un exemple de partition générée avec 7 classes est illustré dans la figure 4.11 (b). Cette topologie particulière de l'environnement permet de s'assurer que la fréquence des observations n'est pas corrélée à la fonction de récompense. De même, les discrétisations des agents restent fixes, ce qui permet d'évaluer leur précision directement. La performance d'une implémentation donnée du modèle

dépend alors seulement des partitions individuelles proposées par la population d'agent et le temps d'apprentissage dépend de la vitesse de convergence du modèle.

4.4.2.1 Amélioration de la précision de la représentation

Dans cette expérimentation le système utilise des représentations partielles (discrétisations et politique d'action associée) pour construire une représentation plus précise de l'environnement. La performance des représentations individuelles est comparée à la performance de la représentation combinée. Ces deux performances sont obtenues en proposant différentes implémentations de notre modèle (un agent contre plusieurs agents). Les discrétisations des représentations individuelles sont générées aléatoirement (avec un nombre fixe d'états) et les politiques d'action associées sont construites grâce au processus d'apprentissage du modèle.

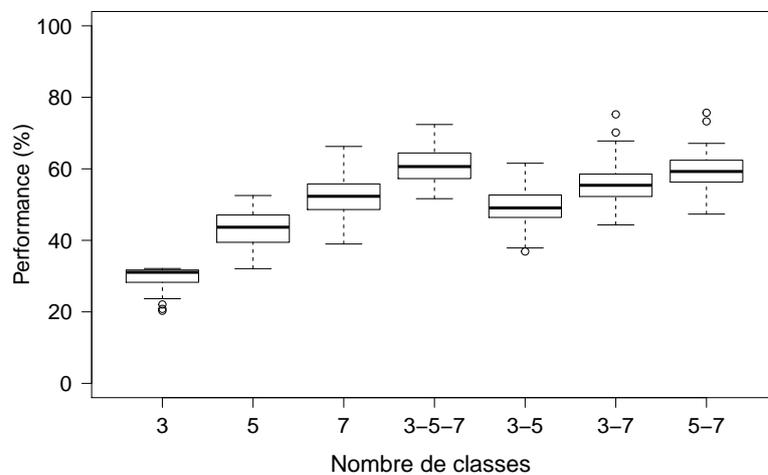


Fig. 4.12: Comparaison de la performance de 7 implémentations du modèle utilisant jusqu'à 3 agents discrétiseurs. Les nombres de classes seuls représentent les cas avec un seul agent alors que les nombres multiples donnent les classes de chacun des agents. Exemple : 3-5-7 signifie une implémentation avec 3 agents discrétiseurs ayant chacun 3, 5 et 7 classes dans leur discrétisations.

La figure 4.12 montre les résultats obtenus avec différentes implémentations du modèle. Les trois distributions de gauche correspondent aux performances des représentations individuelles (avec respectivement 3, 5 et 7 classes). Les autres distributions sont les résultats obtenus avec des implémentations du modèle utilisant toutes les combinaisons possibles des agents discrétiseurs portant les représentations individuelles. Dans toutes les simulations, les discrétisations sont fixes mais l'apprentissage permet de générer la politique d'action associée à la (aux) discrétisation(s). Les résultats présentés ont été produits en répétant 100 fois chaque scénario, qui consiste en 100000 itérations. Les résultats sont complètement reproductibles dans la mesure où chaque discrétisation aléatoire produite est exactement identique pour chacune des 100 répétitions.

Les résultats obtenus confirment que le modèle est capable de construire une représentation plus précise que celles originales (les représentations individuelles), lorsqu'elles sont utilisées séparément. Les valeurs de performance obtenues par les implémentations du modèle utilisant plusieurs représentations concurrentes sont systématiquement supérieures aux implémentations

avec un seul agent. Ces résultats montrent que le modèle est capable de produire, sous une forme implicite, une nouvelle discrétisation et sa politique d'action associée. Cette représentation de haut niveau est en fait la meilleure combinaison possible des éléments de représentation proposés par les agents discrétiseurs. La performance atteinte semble toutefois bornée. On peut exprimer deux raisons à cela. D'abord, la représentation absolue est constituée de 9 actions (et états) alors que les agents discrétiseurs manipulent au plus 7 classes. Ensuite, les 100000 itérations incluent la phase d'apprentissage de la politique. D'ailleurs, les performances atteintes par un modèle utilisant plusieurs discrétiseurs sont obtenues avec un nombre fixe d'itérations. Ce qui suggère que le temps de convergence probablement augmenté n'est pas décisif pour la performance du système à plus long terme. Ce résultat nous oriente tout de même vers une étude approfondie des temps de convergence.

4.4.2.2 Intégration de représentations expertes contradictoires

De meilleures performances sont souvent synonymes d'un temps de convergence accru pour les méthodes par apprentissage. Le processus d'apprentissage du modèle conceptuel se base sur un algorithme d'exploration-exploitation. Ces algorithmes sont supposés converger vers une sélection optimale des actions au bout d'un nombre d'interactions limité. Cette expérimentation a pour but de montrer l'apport de représentations individuelles expertes ou construites préalablement à l'instanciation du système sur la réduction du temps de convergence. La zone de validité (ou d'efficacité) de représentations construites *a priori* est souvent limitée à un sous-espace restreint de la perception. Si le système n'est pas capable de discriminer deux situations (considérées comme le même état dans sa représentation), cela peut signifier qu'il lui manque des informations sur le problème traité. Par l'ajout de représentations expertes contradictoires, il est possible d'améliorer la capacité de discrimination du système. La vitesse de l'apprentissage, qui n'est plus contrainte par l'assimilation initiale doit alors être substantiellement améliorée.

Nous avons sélectionné un des scénarios (le cas avec les 3 agents) de l'expérimentation précédente. La performance du système est alors représentée sous la forme d'une moyenne glissante pour deux implémentations (figure 4.13).

L'apprentissage d'une politique d'action (courbe bleue) nécessite un grand nombre d'itérations pour converger. Lorsque la politique d'action est fournie au système (courbe rouge), le temps de convergence est considérablement réduit. L'algorithme se focalise seulement sur l'apprentissage de la combinaison des discrétisations. La performance atteinte par l'algorithme est aussi améliorée puisque le système dispose de plus d'expérience pour ajuster les récompenses moyennes. Cette expérimentation montre que la tâche la plus coûteuse dans le processus d'apprentissage du système est l'élaboration de la politique d'action. Nous proposons donc d'utiliser des politiques d'actions (expertes ou apprises *a priori*) afin d'amorcer plus rapidement la construction de la représentation du système. L'apprentissage se concentrera alors sur la combinaison dynamique des discrétisations des agents. Il est alors possible de fournir au système des représentations qui sont adaptées à différents contextes qu'il pourrait rencontrer. Nous proposons par la suite d'étudier le comportement d'un système qui ferait face à un environnement dynamique. Les résultats obtenus au niveau du temps de convergence laissent espérer que le système pourrait rapidement s'adapter aux différents contextes, en recombinaison des discrétisations grâce au processus d'apprentissage accéléré.

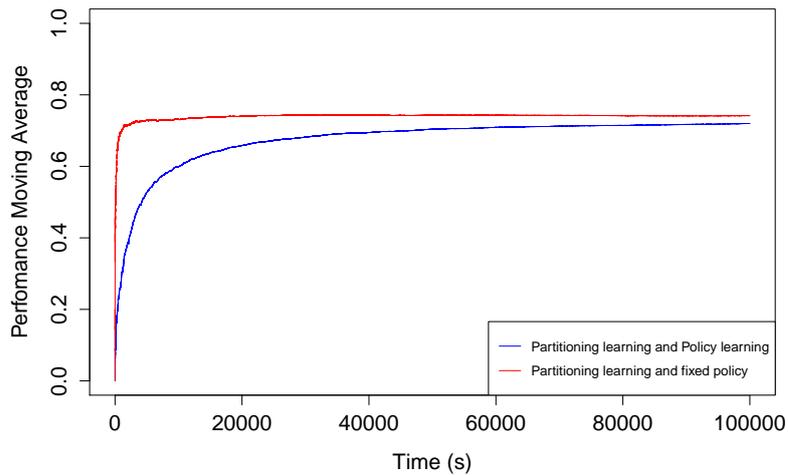


Fig. 4.13: Comparaison des temps de convergence de l'apprentissage pour une implémentation du modèle avec une politique d'action donnée et pour une implémentation où seule la discrétisation est donnée.

4.4.3 Capacité d'adaptation

Dans cette partie, nous proposons d'étudier la capacité d'adaptation du système, *i.e.* sa faculté à réutiliser des connaissances apprises dans des contextes inédits. Un agent autonome en interaction avec son environnement tente de construire une représentation aussi proche que possible de la représentation absolue du monde, pourtant inaccessible. Le problème majeur rencontré par cet agent est que l'environnement est souvent dynamique et peut rapidement changer d'une situation à une autre. Puisque l'agent n'a pas la capacité de percevoir dans son intégralité toutes les variables de l'environnement, il ne sera pas capable de comprendre pourquoi une politique d'action apprise dans un contexte donné échouera dans une situation différente.

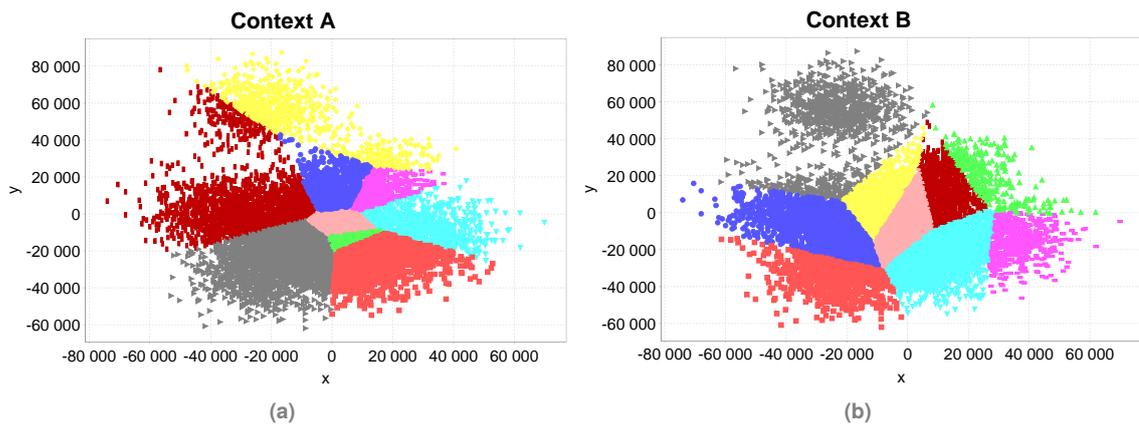


Fig. 4.14: Illustration des environnements générés (deux contextes). Chaque couleur représente une action attendue différente.

Dans cette expérimentation, nous avons défini deux représentations absolues de l'environnement (désignées comme le contexte A et le contexte B dans la figure 4.14). Ces représentations sont

composées d'une discrétisation (9 classes) en deux dimensions et d'une politique d'action associée (une action distincte par classe). L'écart type de la distribution Gaussienne ayant servi à la génération de la discrétisation est fixé à 20000, avec une moyenne centrée en (0,0).

4.4.3.1 Transition entre différents contextes

Chaque contexte a été appris par un ensemble spécifique de trois agents discrétiseurs utilisant une discrétisation composée de 7, 9 et 11 classes (générées aléatoirement). La politique d'action est construite à partir d'un apprentissage dans un premier scénario de 100000 interactions.

Trois implémentations du modèle sont comparées :

1. L'ensemble des agents discrétiseurs spécialisés dans le contexte A ;
2. L'ensemble des discrétiseurs spécialisés dans le contexte B ;
3. L'ensemble composé des agents spécialisés dans le contexte A et dans le contexte B (combinés grâce au modèle).

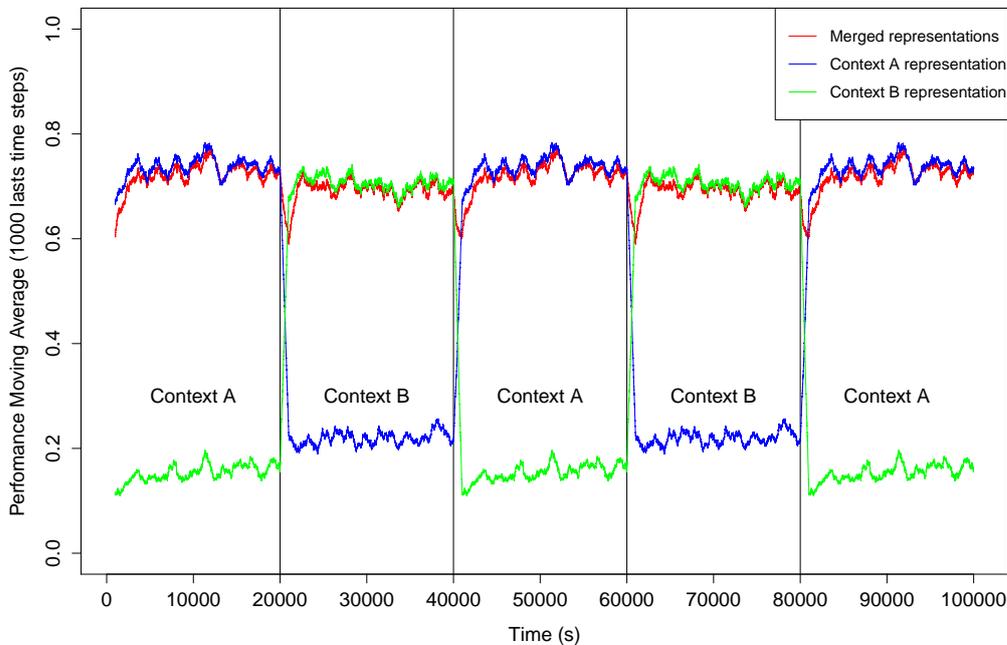


Fig. 4.15: Comparaison de la performance de trois implémentations face à des changements de contexte.

Le scénario consiste à utiliser la représentation de chaque contexte par alternance. Le changement entre les deux contextes, immédiat, s'opère toutes les 20000 itérations. Ces changements sont visibles dans le tracé des performances (moyennes glissantes) de la figure 4.15 : ils sont représentés par les lignes verticales. Les changements de l'environnement sont modélisés comme instantanés. Il n'y a donc pas de transition entre les deux contextes. Les observations sont répétées entre deux occurrences du même contexte, ce qui permet de mieux observer le comportement du modèle (par exemple, le système perçoit la même chose au temps t et au temps $t + 40000$). La figure montre que les stratégies apprises dans un contexte spécifique (courbes bleues et vertes) donnent de mauvaises performances lorsque les agents sont confrontés à un autre contexte. Ce résultat est

logique puisque l'adaptation de la politique d'action ne suffit pas à elle seule lorsque la discrétisation n'est pas adaptée. La représentation issue de la combinaison des deux agents spécialisés dans leurs contextes respectifs (courbe rouge) montre un comportement très intéressant du modèle : le système est capable de s'adapter dynamiquement lors d'un changement de contexte. Comme l'a laissé pressentir l'expérience précédente, ce processus nécessite un temps d'adaptation, qui correspond à la re-convergence de l'algorithme d'exploration-exploitation. Un résultat intéressant est que les changements de contexte successifs n'ont pas d'impact sur le délai d'adaptation de la représentation. Cette expérimentation confirme que le système peut facilement passer d'un contexte connu à un autre (*i.e.* lorsqu'il a préalablement appris une représentation associée). Dans les problèmes réels, une distinction aussi nette entre les contextes n'est toutefois pas évidente. Le système devra alors adapter sa représentation à des contextes proches de ceux qu'il a rencontré.

4.4.3.2 Adaptation à différents contextes

Au long de son existence, un système autonome est confronté à plusieurs contextes (ou plusieurs environnements). La tâche d'apprentissage ne se limite pas à la discrimination de ces contextes et à l'apprentissage d'une représentation associée à chaque contexte. Le comportement d'adaptation permet en fait au système de réutiliser les connaissances acquises précédemment dans un contexte variable. Notre modèle utilise une combinaison dynamique de représentations individuelles qui peuvent être fixes (données *a priori*) ou dynamiques (intégrées dans le processus global d'apprentissage). La représentation du système peut alors se re-construire dynamiquement en exploitant des éléments re-combinés à partir des représentations individuelles. Dans cette expérimentation, nous réutilisons les deux contextes (A et B) avec 9 actions définies dans l'expérimentation précédente (figure 4.14). Les discrétisations générées dans chacun des contextes ne changent pas (elles sont toujours modélisées par 3 agents avec 7, 9 et 11 classes fixes). Et les politiques d'action des deux implémentations sont apprises séparément dans les deux contextes. En utilisant le modèle, nous proposons une implémentation utilisant tous les agents discrétiseurs, soit 6 représentations concurrentes. L'objectif de l'expérimentation est de montrer le comportement du modèle lorsqu'il fait face à des situations qui diffèrent quelque peu des contextes originaux qu'il connaît (par l'intermédiaire des représentations de ses agents discrétiseurs).

La figure 4.16 décrit les performances moyennes obtenues par les différentes implémentations dans différents contextes. Les valeurs extrêmes (droite et gauche) correspondent aux contextes A et B originaux. Les valeurs intermédiaires correspondent à une interpolation linéaire des positions originales des centres de classes entre les contextes A et B (pour les états affectés à la même action). Ces performances sont obtenues après les 100000 itérations. Les agents discrétiseurs utilisés dans les trois implémentations embarquent leurs expériences acquises dans leurs contextes respectifs. Chaque point représente le résultat d'une simulation. Chacune des trois implémentations est testée dans tous les contextes générés (11 contextes, y compris les contextes A et B originaux). Par exemple, "20% A - 80% B" décrit un contexte où les positions des centroïdes des états correspondant à chacune des actions sont calculées comme la moyenne pondérée des positions du contexte A (0,2) et du contexte B (0,8). Les contextes A et B ont été sélectionnés pour qu'ils soient suffisamment différents et mènent à des politiques d'actions incompatibles. Les résultats obtenus montrent que le modèle obtient de meilleures performances (ou au moins égales) qu'avec les représentations initiales dans les contextes interpolés. Cela confirme que le modèle est capable d'adapter dynamiquement sa représentation à des contextes proches de ceux précédemment rencontrés. Comme prévu, les représentations apprises dans un contexte spécifique donnent des performances décroissantes à

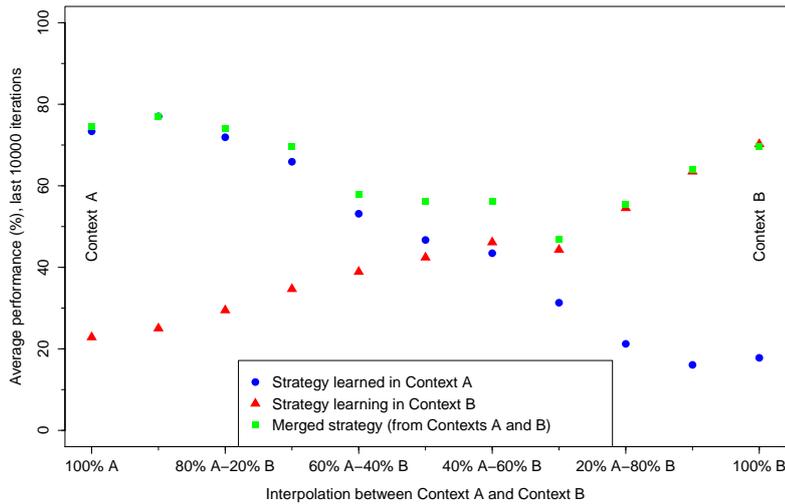


Fig. 4.16: Performance moyenne des trois implémentations face à des interpolations entre les contextes A et B.

mesure que l'on s'éloigne du contexte original. Un autre résultat inattendu correspond au scénario "90% A - 10% B", où la performance obtenue par les agents ayant appris dans le contexte A est meilleure que dans le cas "100% A". L'explication se situe dans la précision de la discrétisation : elle est plus proche de la représentation absolue de l'environnement dans ce cas.

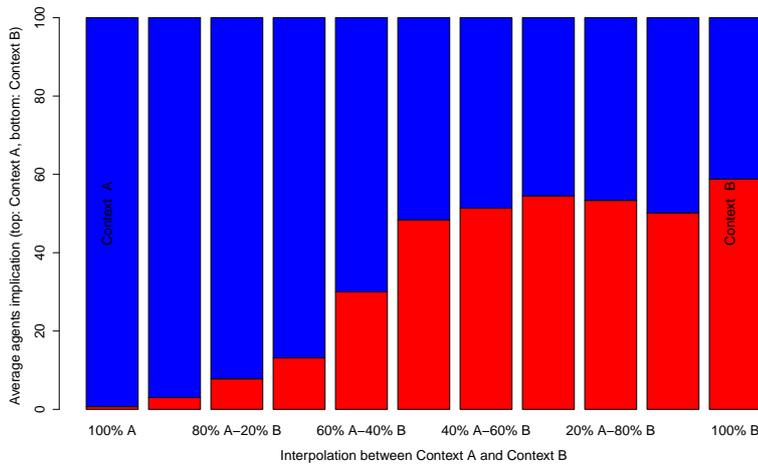


Fig. 4.17: Pourcentage de participation des agents spécialisés dans les contextes A et B face à des contextes interpolés.

La figure 4.17 permet de visualiser la participation de chacun des agents dans la construction de la représentation du système pour chaque contexte simulé. Cette participation est estimée en comptant le nombre de sélections de lien perception-action émanant d'un agent en particulier. La partie gauche du diagramme montre que les agents spécialisés dans le contexte A contribuent quasiment exclusivement à la construction de la représentation du système. Le mélange des agents est moins prévisible lorsque le contexte se rapproche de B. On aurait d'ailleurs pu s'attendre à ce que les agents du contexte B atteignent quasiment 100% d'implication dans le cas du contexte B. En réalité, certains éléments des politiques d'action respectives des agents peuvent être les mêmes. Dans ce cas, le

modèle choisit la politique qui, en moyenne, donne les meilleures récompenses. Cela passe par l'adéquation de la discrétisation à l'environnement rencontré. Il y a en effet deux façons d'améliorer la performance d'un système : se rapprocher d'un contexte proche de la discrétisation disponible, ou permettre au modèle d'affiner sa représentation en générant de nouvelles discrétisations.

4.4.4 Affinage de la représentation

La construction de la représentation d'un agent autonome est itérative et son apprentissage s'effectue en réalité tout au long de son existence. La phase d'amorçage de cette construction est cruciale pour permettre au système de démarrer sur de bonnes bases. Nous avons proposé des mécanismes concurrents permettant de réaliser cette tâche, en profitant des expériences individuelles des agents discrétiseurs. Le système est donc capable d'interagir précisément dans les contextes qu'il a déjà rencontrés. Sa capacité d'adaptation (section précédente) lui permet de faire dynamiquement et temporairement évoluer sa représentation pour des contextes proches. En revanche, la performance atteinte par le système semble bornée. Cela est dû au modèle de discrétisation qui repose sur un découpage statique. On peut bien sûr imaginer d'utiliser des algorithmes de discrétisation en-ligne (type GNG [68]) qui convergeront quand même au bout d'un certain temps et limiteront donc l'amélioration du découpage (même dans le cas d'un couplage de la discrétisation et des récompenses [204]). Notre solution passe par la proposition d'un nouveau type d'agent nommé "associateur" et présenté en détail dans la partie 4.3.5. Le principe de base est de considérer l'ajout des intersections entre les états des agents discrétiseurs.

L'objectif de cette expérimentation est de montrer l'effet des agents associeurs sur l'évolution de la performance. Pour cela, nous proposons de reprendre le cadre de l'expérimentation menée en section 4.4.2.1 et la figure 4.12 associée. Les résultats ont donc aussi été produits en répétant 100 fois des scénarios de 100000 itérations avec un environnement aléatoire différent. Trois implémentations du modèle sont comparées : une implémentation avec 3 discrétiseurs (ayant respectivement 3, 5 et 7 classes ; même cas que dans la figure 4.12) et deux implémentations avec un agent associeur. La différence entre ces deux dernières implémentations réside dans la méthode d'activation des états intersection : dans le premier cas ils sont tous activés dès le début (en vert sur la figure 4.18) et dans le second cas (en bleu), la stratégie d'activation présentée dans la section 4.3.5 est en place. Les paramètres choisis sont :

- niveau de l'associeur : 1
- bonus d'activation : 500 itérations
- bonus d'usage : 2 itérations
- valeur de décroissance : 1 itération

Ces paramètres (voir section 4.3.5) permettent de s'assurer qu'un état activé dispose de suffisamment de temps pour qu'il puisse prendre le dessus sur d'autres états (temps minimum d'assimilation garanti) tout en limitant ses effets néfastes s'il détériore la représentation. Le bonus d'usage est logiquement supérieur à la valeur de décroissance pour encourager l'utilisation des états agrégés. Ces paramètres peuvent être ajustés pour limiter la baisse de performance ou au contraire favoriser l'exploration des nouveaux états créés par les associeurs.

Les résultats obtenus dans la figure 4.18 sont présentés à différents stades des simulations. Le comportement du modèle montre en effet deux phases relativement distinctes : une phase d'exploration

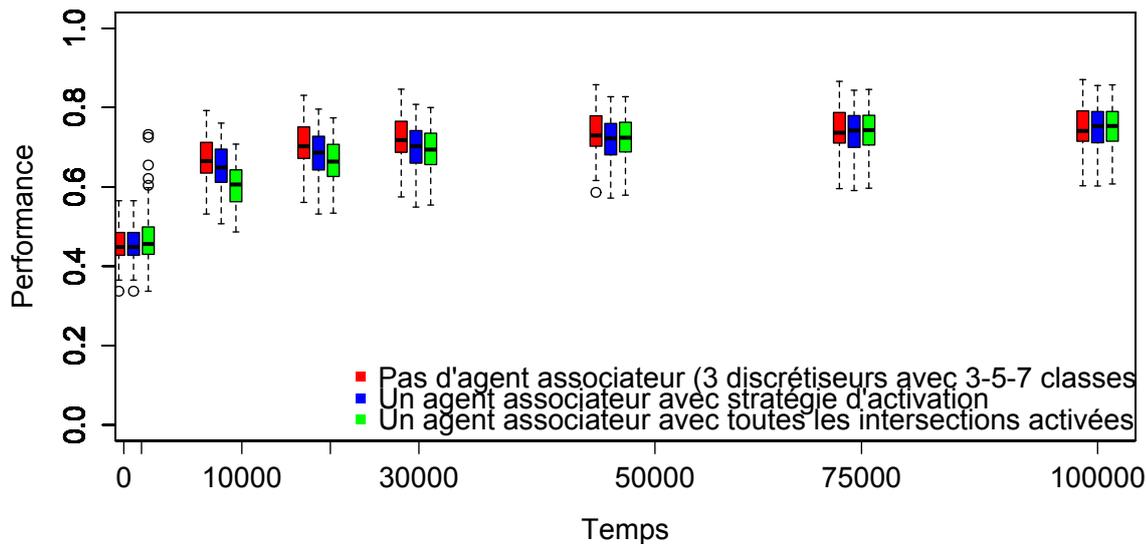


Fig. 4.18: Comparaison de trois implémentations du modèle impliquant ou non un agent associateur.

qui se traduit par des performances plus faibles mais croissantes, et une phase d'exploitation qui se traduit par des performances plus élevées et stables. Dans la figure, on peut observer que le témoin (sans associateur – courbe rouge) converge plus rapidement que les autres implémentations, mais atteint une performance plus faible. En effet, l'ajout de nouveaux états agrégés (intersections des états des 3 discrétiseurs) permet de produire un découpage plus précis, et donc plus proche de la représentation de l'environnement. Les performances obtenues avec les agents associateurs convergent donc vers des valeurs plus élevées. L'autre paramètre intéressant est la rapidité de convergence observée entre les deux stratégies d'activation. La figure montre que la stratégie d'activation (courbe bleue) limite beaucoup le temps de convergence et permet une meilleure performance plus rapidement que dans le cas naïf ou tous les états agrégés sont activés (courbe verte). Logiquement, lorsque tous les états sont activés, et au bout d'un temps important, la performance est plus importante (fin de la simulation ; abscisse à 100000) mais l'écart avec la stratégie d'activation n'est pas flagrant. Cela confirme donc que la stratégie proposée est capable de limiter l'utilisation d'intersections entre états non pertinentes tout en assurant une exploration des candidats les plus prometteurs, qui sont intégrés dans la représentation du système grâce aux mécanismes de base du modèle. Des expérimentations plus poussées devraient confirmer l'apport de plusieurs agents associateurs simultanément, dans un cadre applicatif plus complexe. Le prototype expérimental s'avère en effet quelque peu limité notamment pour l'exploration des différents niveaux d'agrégation qui produisent ici des états trop petits pour être exploités par le modèle. Une perspective à plus long terme sera d'explorer les différentes manières d'agréger les états, en continuant les associations binaires (unions, intersections, ...) ou en proposant de nouveaux mécanismes plus complexes. Le concepteur devra alors faire face à une augmentation substantielle de la combinatoire lors de l'utilisation de l'algorithme d'exploration exploitation qu'il faudra réduire pour garantir des résultats dans un temps restreint.

4.5 Application : stratégie I2V de contrôle des C-ITS

Le déploiement des véhicules connectés va être accompagné par de nouveaux comportements, tant au niveau des usagers que des gestionnaires. En effet, grâce aux technologies ITS, la majorité des acteurs pourra bénéficier d'un partage d'information propice à un meilleur usage de l'infrastructure. La tendance vers l'automatisation permettra d'imaginer des stratégies de conduite individuelles allant dans le sens de l'optimisation du flux (chapitre 3). Ce nouveau type de comportement et ses effets ne sont pas encore connus ou maîtrisés par les gestionnaires. Il est alors nécessaire d'imaginer des stratégies de contrôle afin de réguler les flux d'information (chapitre 2) ou visant à contrôler les véhicules. L'hétérogénéité des comportements est un frein au développement de stratégies globales. Appliquées à l'ensemble du flux, elles risquent de souffrir d'un effet "on/off" qui n'est pas compatible avec les dynamiques des flux de trafic (section 4.1). Comment imaginer alors des stratégies de contrôle plus adaptatives, capables de créer et de mettre à jour une représentation fine de l'état de trafic sur une section de route ? Les approches complètement locales sont possibles mais limitent l'apport d'objectifs de haut niveau. Rappelons que les systèmes coopératifs (e.g. en France) n'en sont qu'à leur phase pilote. Un des enjeux (chapitre 1) passe aussi par la proposition d'outils d'aide à la décision montrant le potentiel de ces systèmes à plus long terme.

Dans ce cadre, nous proposons un modèle conceptuel (section 4.3) dont les besoins découlent de ce problème applicatif concret. Les analogies entre les problèmes rencontrés et les verrous scientifiques de l'IAD nous ont poussés à concevoir des mécanismes basés sur des représentations concurrentes et les systèmes multi-agents dans une approche d'apprentissage constructiviste. Bien que le chemin parcouru puisse paraître important face au problème du contrôle du trafic coopératif, les bénéfices possibles d'une telle approche paraissent évidents. Les capacités d'adaptation (changement de contexte) sont cruciales pour que le système de contrôle modélise précisément différentes situations de trafic (typiquement, les pics du matin et du soir). Les Unités de Bord de Route, véritables systèmes autonomes de contrôle, ne pourront pas être mises à jour aussi régulièrement que le flux évolue. Leur stratégie de contrôle devra alors être capable d'apprendre de nouvelles régularités au cours du temps. Un autre avantage du modèle est l'apport des systèmes multi-agents dans la construction de la représentation. Les agents discrétiseurs peuvent embarquer des connaissances, modélisées sous la forme d'une représentation individuelle (discrétisation et politique d'action associée). Le système mis en place devra en effet rapidement être opérationnel. Deux options sont alors possibles : enrichir les agents avec des connaissances expertes (une stratégie existante et éprouvée), ou utiliser un outil de simulation pour générer un grand nombre de situations (pas forcément maîtrisables) et ainsi faire apprendre les agents avant leur mise en œuvre dans une UBR. Quelle que soit l'option retenue, la phase d'implémentation du modèle reste cruciale. Dans la suite de cette section, nous proposons un scénario en simulation permettant d'illustrer une utilisation du modèle dans le contexte d'un flux de véhicules connectés sur une section d'autoroute.

4.5.1 Cadre expérimental

Notre étude se focalise sur un scénario relativement simple d'un point de vue modélisation, mais qui présente un convergent introduisant une grande part d'incertitude dans l'écoulement du flux. L'impact des changements de voies est en effet difficile à mesurer, bien que la communauté s'accorde sur l'effet néfaste sur le flux, observé directement [72] ou en simulation [117]. Cette chute de capacité,

bien que modélisable, dépend de beaucoup de paramètres individuels auxquels il est difficile voire impossible d'accéder [219] : anticipation, phénomène de relaxation, etc. La section d'étude est parcourue par un flux mixte composé de véhicules connectés et non connectés. Les véhicules connectés peuvent échanger des informations avec les autres véhicules et l'infrastructure (V2V et V2I), mais notre scénario se concentre surtout sur les interactions de l'infrastructure (messages I2V). L'infrastructure est composée d'unités de bord de route (UBR) en charge du contrôle de leur section dédiée. Une UBR est capable de percevoir des informations afin d'estimer l'état de trafic actuel par l'intermédiaire de ses capteurs. La chaussée est par exemple équipée de boucles électromagnétiques qui agrègent les informations de débits et de vitesse au cours du temps. L'objectif d'une UBR est de propager des messages de recommandations aux véhicules connectés à portée. Cet envoi de message est l'application d'une stratégie de contrôle autonome qui tend à réduire l'ampleur ou retarder l'apparition de congestion.

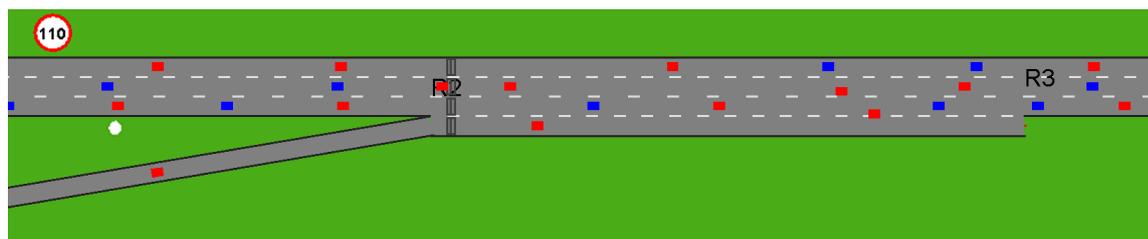


Fig. 4.19: Capture d'écran du scénario de contrôle modélisé dans le simulateur. Le flux est composé de véhicules connectés (bleu) et non connectés (rouge). L'unité de bord de route (disque blanc) récolte les informations des boucles électromagnétiques (rectangles gris) pour construire par apprentissage, sa représentation de cet environnement autoroutier.

Comme illustré dans la figure 4.19, toutes les expérimentations ont lieu sur le même scénario, qui se déroule sur une portion rectiligne de 3 km d'une autoroute à trois voies avec une voie d'insertion à 2,5 km. L'Unité de Bord de Route perçoit des données agrégées de flux (débits) et de vitesses grâce aux boucles électromagnétiques présentes sur les trois voies. Les véhicules connectés partagent à l'UBR leurs vitesses dans son rayon de communication (150 mètres). Le processus de décision de l'UBR se base sur une implémentation de notre modèle de contrôle. Son objectif est, en envoyant des messages de consigne, d'éviter la formation de congestion, ou de réduire son ampleur. Ces messages de consigne sont transmis par communication I2V. La voie d'insertion provoque une chute des vitesses à partir de la voie de droite, qui peut mener à la formation de congestion se propageant à toute la section en condition de trafic dense. Nous proposons d'étudier le comportement du modèle intégré dans une UBR en faisant l'hypothèse d'un taux de pénétration fixe : 30 % de véhicules connectés, mélangés de manière homogène aux autres véhicules.

Afin de générer des contextes de l'environnement différents pour l'UBR, nous avons utilisé des données réelles issues de boucles électromagnétiques. Les débits observés sont reproduits pour générer les flux d'entrée sur la section principale et la voie d'insertion (débits dynamiques). Le jeu de données correspond à plusieurs capteurs situés sur l'autoroute A6, avec un recueil portant sur la période 8h – 9h du mois de juin 2001. Deux capteurs ont été sélectionnés (un sur l'autoroute, l'autre sur une voie d'insertion) et le scénario proposé (figure 4.19) reproduit la topologie des lieux. Un ensemble de 30 simulations ont ainsi été exécutées ; chacune se déroulant pendant 60 minutes. La diversité des conditions de trafic (pic du matin, week-ends et jours fériés) permet d'étudier le comportement du modèle pour des contextes très variables (trafic fluide, régime critique ou congestion).

4.5.2 Implémentation

Afin de décrire précisément les paramètres utilisés dans cette expérimentation, il est nécessaire de décomposer les deux parties des simulations. D'une part, les paramètres du simulateur multi-agent de trafic (présenté dans le chapitre 3) et d'autre part l'implémentation du modèle conceptuel (section 4.3).

4.5.2.1 Véhicules et infrastructure

L'architecture multi-agent du simulateur permet de définir des comportements complexes pour les UBR et les véhicules. Nous proposons d'implémenter notre modèle comme le processus de décision de l'UBR. Afin de reproduire les dynamiques du trafic, le comportement des véhicules est modélisé par une loi de poursuite. Les consignes envoyées par l'UBR aux véhicules sont intégrées dans leur processus de décision : ils surchargent ce comportement afin de s'adapter aux consignes reçues.

L'Unité de Bord de Route

Le résultat de la stratégie de contrôle de l'UBR est la sélection d'une action, à chaque pas de temps de décision (fixé à 120 secondes). Ces actions prennent la forme d'un envoi de message de consigne aux véhicules connectés à portée. Les consignes disponibles sont les suivantes :

- A_1 : pas de consigne envoyée.
- A_2 : consigne de changement de voie (de droite à gauche).
- A_3 : consigne de changement de voie (de gauche à droite).
- A_4 : consigne d'inter-distance (1,8 s).
- A_5 : consigne d'inter-distance (1,2 s).
- A_6 : limite de vitesse (110 km/h).
- A_7 : limite de vitesse (50 km/h).

Les messages de consignes envoyés intègrent les informations relatives aux paramètres des actions de l'UBR (voir sections 3.2.3.2 et 3.2.5.1). La zone de pertinence des messages de consigne est fixée à la section entière, et leur délai d'expiration est limité à 120 secondes. La consigne encapsulée dans un message est appliquée par les véhicules qui l'ont reçue tant qu'ils se trouvent dans la zone de pertinence du message et que celui-ci n'a pas expiré.

Les véhicules

Le comportement physique de base des véhicules (connectés ou non) est régi par un modèle microscopique (IDM [192], équation 2.4) couplé à un modèle de changement de voie (stratégie MOBIL [104]). Les véhicules connectés sont équipés d'un dispositif de communication leur permettant de percevoir les messages dans un rayon de 150 mètres. Les messages reçus de l'UBR (I2V) sont re-propagés par communication V2V et couvrent ainsi l'intégralité des 3 km de la section. Les véhicules connectés affichent le contenu des messages de consigne sur une interface embarquée qui permet au conducteur d'adapter son comportement à la consigne courante. Cette modification du comportement est une interprétation du message par les véhicules. A des fins de simplification, tous les conducteurs réagissent de la même manière aux consignes :

- l'action A_1 n'a aucune incidence puisqu'aucun message n'est envoyé par l'UBR.
- les consignes de changement de voie entraînent une modification temporaire de la stratégie de changement de voie, décrite en détail en section 3.2.5.2.
- pour les consignes d'inter-distance et les limitations de vitesses, il s'agit d'une modification temporaire des paramètres de la loi de poursuite du véhicule (voir section 3.2.5.2).

Toutes ces modifications en réaction aux consignes sont maintenues durant la durée de pertinence du message de consigne reçu. Les messages (et leurs effets) sont automatiquement effacés si le message expire ou si le véhicule quitte la zone de pertinence.

4.5.2.2 Modèle de contrôle

Le modèle est implémenté en lieu et place du processus de décision de l'UBR. Ce système est capable de percevoir les informations des boucles électromagnétiques qui calculent le débit moyen, la concentration et la vitesse de chacune des voies (y compris la voie d'insertion). Il s'agit de données agrégées toutes les 2 minutes (120 secondes). En plus, les véhicules connectés partagent leur vitesse respective lorsqu'ils se trouvent dans le rayon de communication de l'UBR (ce message V2I n'est pas re-propagé). Toutes ces données servent de perception de bas niveau au système. Il dispose donc de deux sources d'informations concurrentes lui permettant de caractériser l'état courant de trafic. La construction de la représentation du système repose sur plusieurs agents discrétiseurs qui manipulent des hypothèses sensorimotrices. Dans cette expérimentation, la perception et la décision du système sont discrétisées temporellement par période de 120 secondes (*i.e.* le cycle d'exécution du système est rythmé par ses capteurs). L'UBR perçoit les données agrégées de ses capteurs durant le dernier pas discret, et choisit immédiatement une action à exécuter (*i.e.* envoie une consigne). Après un nouveau pas de temps de 120 secondes, la consigne expire et le système reçoit un *feedback* pour renforcer sa représentation. Au même moment, l'UBR reçoit sa nouvelle perception par ses capteurs.

Les agents discrétiseurs

Afin de montrer les bénéfices apportés par le modèle, nous proposons une implémentation reposant sur l'instanciation de deux agents discrétiseurs. Le but de chaque agent est de générer une représentation individuelle à partir des données qu'il perçoit. Nous avons choisi d'affecter les variables d'un capteur à chacun des agents D_1 et D_2 . La différence entre les deux agents se situe donc avant tout dans leurs perceptions. D_1 est lié aux boucles électromagnétiques. Il perçoit donc les données de débit et de concentration pour les 4 voies observées (8 dimensions) et agrégées toutes les 120 secondes. D_2 exploite les données reçues dans les messages V2I et construit un modèle (moyenne et écart-type) de la distribution des vitesses reçues dans les messages provenant des véhicules et qui sont passés devant l'UBR durant le dernier pas de temps de décision. Les discrétisations des agents résultent d'une classification par la méthode des K-moyennes (stabilisé avec 10000 itérations pour des initialisations aléatoires) utilisant respectivement 4 et 3 classes (pour D_1 et D_2). Le résultat de la classification de chacun des agents sur les 15 premières simulations est représenté dans la figure 4.20.

L'apprentissage des liens perception-action est réalisé sur les 15 mêmes simulations, grâce à l'exploitation de la récompense fournie au système.

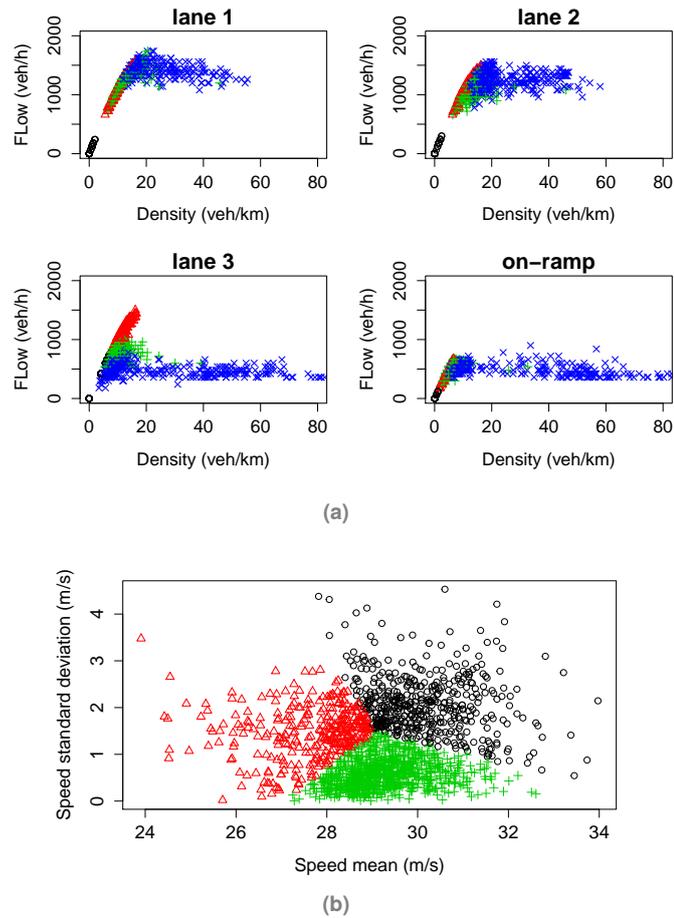


Fig. 4.20: Discretisations générées par les agents discrétiseurs (a) D_1 [8 dimensions] et (b) D_2 [2 dimensions] sur le jeu d'apprentissage. Les couleurs (indépendantes entre les deux discrétisations) correspondent à des classes distinctes pour les valeurs observées.

La récompense du système

Les récompenses mémorisées par les agents discrétiseurs dans les liens perception-action sont calculées au moment de l'élaboration de la perception du système par l'environnement. Dans le scénario étudié, nous avons choisi la vitesse moyenne minimale (sur les 4 capteurs) en tant que *feedback* transmis au système. Cette fonction de récompense correspond à l'objectif de haut niveau : améliorer le flux. Cette valeur externe est utilisée comme une récompense pour le système dans son processus d'exploration-exploitation. La stratégie de contrôle et les discrétisations peuvent être renforcées à partir de l'évaluation du résultat des actions. Dans cette expérimentation, l'algorithme UCB utilisé (équation 4.1) a été légèrement modifié. Les récompenses sont toujours données à tous les liens ayant pour action la plus probable l'action effectivement sélectionnée par le système. En revanche, le nombre de sélections de chaque lien (n_l) n'est incrémenté que pour le lien ayant été sélectionné. Cette stratégie de récompense permet d'accélérer le processus d'exploration tout en réduisant les effets d'un *feedback* bruité.

4.5.3 Résultats

L'expérimentation est menée en trois étapes principales. D'abord, un premier jeu de données d'entrée (perception des agents D) est nécessaire pour générer les classifications à partir de l'algorithme dédié. Durant cette phase, l'apprentissage n'est pas réalisé avec le modèle. Ensuite, une partie des simulations est dédiée à l'apprentissage de la représentation par le système, grâce au processus complet du modèle (les 15 premières simulations). Enfin, le système est prêt à être confronté à des scénarios alternatifs, où sa représentation précédemment construite devrait aider le système dans sa stratégie de contrôle. Cette phase a été menée sur la seconde moitié des simulations. Le résultat témoin de chaque simulation est obtenu en observant différents indicateurs dans un scénario dépourvu d'unités d'infrastructure. Cela signifie qu'il n'y a aucune communication entre les véhicules, et qu'aucune consigne n'est propagée. Les autres résultats proviennent de trois implémentations du modèle : le cas avec un seul agent (D_1 et D_2 séparément) et une combinaison dynamique des représentations de D_1 et D_2 .

Afin d'évaluer les effets des stratégies de contrôle produites par les différentes implémentations du modèle, nous avons choisi trois indicateurs. Le temps total passé (TTS), la vitesse moyenne sur la section, et le pourcentage de congestion (ces indicateurs sont décrits en détail dans la section 3.3). Des valeurs de TTS faibles signifient un meilleur écoulement du flux. La vitesse moyenne permet d'observer l'effet sur l'homogénéisation. Et le pourcentage de congestion donne un aperçu de l'état de trafic et de l'ampleur des perturbations. Le comportement du flux contrôlé par le modèle devrait être directement visible grâce à ces indicateurs.

Parmi les 15 simulations, nous avons sélectionné trois résultats qui permettent d'illustrer différentes situations de trafic. Comme les données des débits d'entrée proviennent de cas réels, les simulations se répètent pour des jours de même profil. La figure 4.21 représente un tracé des trois indicateurs pour les 3 implémentations du modèle (et le scénario témoin).

Les indicateurs permettent d'étudier les bénéfices de chaque implémentation sur l'amélioration du flux, en les comparant avec les scénarios témoins (courbes rouges). Dans les simulations (a) et (c), sans aucun contrôle, les instabilités produites par la voie d'insertion et les changements de voie associés se propagent à l'ensemble du flux. Les trois simulations montrent des comportements différents du modèle. Dans la simulation (a), la stratégie de contrôle de l'agent D_1 permet d'éviter l'apparition de congestion sur la section, alors que celle de D_2 donne de moins bonnes performances que le témoin. Ces résultats sont clairement visibles en observant l'évolution des valeurs de TTS. Une baisse signifie que le temps de trajet des véhicules est plus faible. La combinaison dynamique des représentations individuelles de D_1 et D_2 produit ici une stratégie équilibrée qui résulte en une amélioration moyenne. La simulation (b) illustre un cas spécifique où l'une des représentations utilisées par le modèle est fautive (celle de D_2). Aucune congestion n'est observée dans la simulation témoin mais la stratégie de D_2 provoque des perturbations. Néanmoins, le modèle est capable de converger vers la meilleure représentation individuelle (D_1) dans ce contexte. Cela ne permet toutefois pas d'améliorer l'état de trafic (résultat équivalent au témoin). Dans la figure 4.21 (c), l'état de trafic obtenu par les débits du scénario est proche de l'équilibre du flux. Dans le témoin, la congestion est limitée et la vitesse moyenne ne chute que peu (en comparaison à la simulation (a)). Dans ce contexte, les deux stratégies issues des représentations apprises par les discrétiseurs ont un impact positif sur le flux. L'apport du modèle lors de la construction de sa représentation est aussi

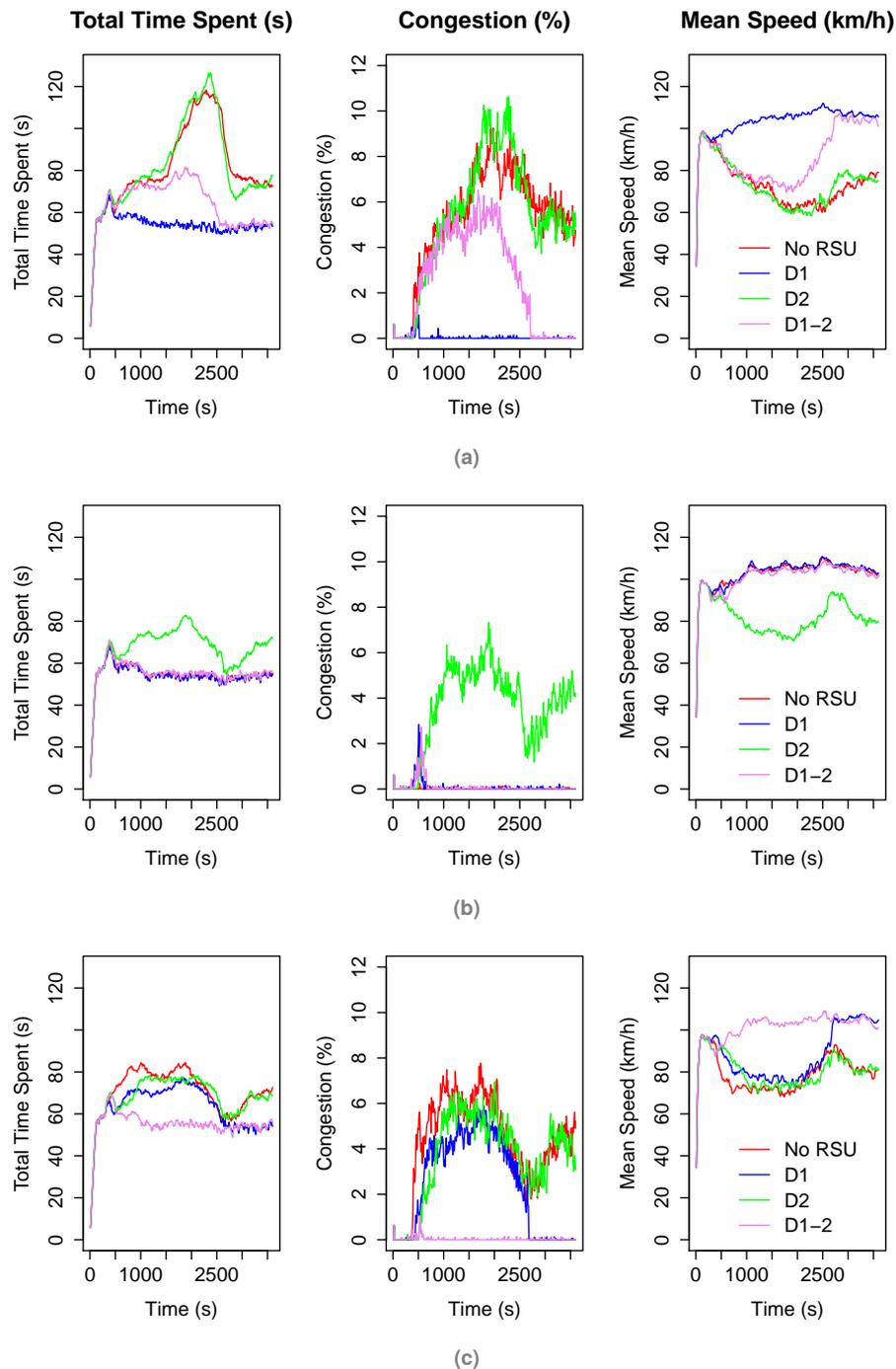


Fig. 4.21: Résultat de trois simulations : (a) Lundi 18, (b) Jeudi 21 et (c) Mardi 28.

clairement visible. Le système a su utiliser les récompenses pour combiner les représentations individuelles, ce qui permet d'éviter toute propagation de congestion sur la section.

La figure 4.22 est une représentation simplifiée de la contribution de chacun des agents discrétiseurs (en reprenant le même code couleurs que celui de la figure 4.21). Comme le laissent présager les résultats obtenus au niveau des indicateurs, on retrouve l'apport des deux agents dans la politique d'action menée par le système. L'action la plus représentée (A_2) est tout à fait logique puisqu'elle

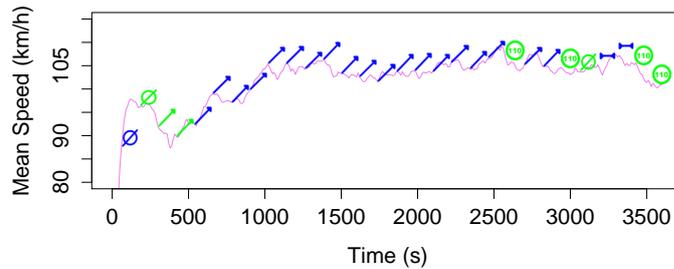


Fig. 4.22: Actions des agents internes au système dans l'implémentation du modèle lors de la simulation (c).

permet de faciliter l'insertion des véhicules arrivant sur le convergent et souhaitant rejoindre la section principale. Cela confirme une nouvelle fois la capacité du modèle à adapter dynamiquement plusieurs représentations concurrentes pour construire une représentation de haut niveau plus précise.

4.5.4 Discussion

La phase d'implémentation est cruciale lorsque le modèle est exposé à un cadre applicatif concret et complexe. Les deux éléments les plus importants sont le choix de la récompense et la précision des discrétisations utilisées. La difficulté du choix de la perception des agents fait référence à un problème de fusion de capteur : certains phénomènes ne s'expriment que dans certaines combinaisons de dimensions. Grâce à l'apprentissage mis en place, le système est capable de produire de nouvelles combinaisons de dimensions dynamiquement et d'exploiter celles qui lui permettent d'augmenter l'expressivité et la précision de sa représentation. Néanmoins, les expérimentations ont montré que le concepteur doit fournir des représentations individuelles relativement précises afin d'aider le système dans l'amorçage de sa construction. Ce problème pourrait être traité en donnant au modèle la possibilité de générer dynamiquement (ou de faire évoluer) les agents discrétiseurs, en exploitant les expériences d'interactions précédentes. Il s'agit du rôle des agents associateurs proposés dans l'extension du modèle (section 4.3.5).

4.6 Bilan et perspectives

Le besoin de proposer des nouvelles stratégies de contrôle des C-ITS découle de l'apparition de nouveaux comportements au sein des flux qui seront progressivement automatisés et/ou connectés. En effet, les stratégies devront pouvoir s'adapter à des changements progressifs, qui rendront les stratégies actuelles inadaptées. Dans ce chapitre, nous avons proposé une approche constructiviste d'élaboration de représentation qui permet à la fois de concilier l'apport de représentations existantes, mais aussi de les faire évoluer pour s'adapter à des contextes encore indisponibles. Notre approche se heurte à des verrous scientifiques du domaine de l'IA (problème d'attribution de la récompense), et notamment en IAD (problème d'élaboration de représentations concurrentes). Pour solutionner une partie de ces verrous, plusieurs mécanismes ont été proposés. Ils mettent à profit les bénéfices de l'apprentissage concurrent : capacité d'adaptation, affinage de la représentation. Afin d'étudier le comportement du modèle face à différents environnements, nous avons proposé un prototype expérimental permettant de les simuler. Les résultats en simulation permettent de valider leur

fonctionnement mais pointent aussi leurs limitations : précision des discrétisations, biais causés par des représentations *a priori*. Enfin, le modèle a été appliqué au trafic : l'objectif est de proposer une forme de contrôle décentralisé où les Unités de Bord de Route sont capables de discriminer l'état de trafic courant pour disséminer des consignes au sein d'un flux de véhicules connectés. La représentation de l'UBR en charge du contrôle d'une section intègre la discrétisation des données de ses capteurs et la politique d'envoi de consignes.

Face à la complexité du problème abordé dans notre travail, mais aussi la particularité du domaine d'étude des C-ITS, le travail présenté ouvre la voie à de multiples perspectives. Du côté du cadre applicatif, la première proposition est de donner la capacité à l'UBR de percevoir le taux de pénétration pour que la politique s'adapte. Cela permettra d'assister la construction de la représentation lors des phases successives de déploiement. Dans le même ordre d'idée, l'outil de simulation de trafic développé et présenté dans le cadre de notre travail pourrait trouver une seconde utilisation. Nous proposons d'utiliser la simulation pour catalyser l'apprentissage d'une UBR avant de mettre le système en exploitation. L'idée est de faire apprendre le système offline avant de lui permettre d'affiner sa représentation en situation réelle ; il serait donc plus rapidement opérationnel. Une seconde voie d'amélioration de la politique serait d'étendre le contrôle à une zone plus étendue en proposant un mécanisme de coordination inter-UBR. Par exemple, les entités pourraient partager leur représentation et choisir dynamiquement d'intégrer les données des capteurs de leurs voisins (si elles apportent un intérêt pour la précision de la représentation). Enfin, comme évoqué dans le chapitre 1, les véhicules peuvent jouer le rôle d'UBR mobile. Il pourrait être utile de proposer d'embarquer le modèle au niveau d'un véhicule et donc d'assurer un contrôle encore plus dynamique. Ces deux pistes d'explorations posent une nouvelle fois le problème d'attribution de la récompense : doit-on privilégier l'intérêt d'un peloton de véhicule ou viser des objectifs plus globaux ?

Plus généralement, comme évoqué dans la section 4.3.6, le modèle peut être amélioré suivant des directions très différentes. Le sujet traité dans ce chapitre est en réalité très vaste. Les verrous scientifiques identifiés, bien qu'abordés sous plusieurs aspects, ne sont encore pas intégralement traités.

Ce chapitre 4 est la concrétisation de la démarche itérative générale du manuscrit présentée dans le chapitre 1. Les résultats présentés n'auraient pas été possibles sans les contributions relatives à la modélisation et la simulation (chapitres 2 et 3). Nous espérons que le travail permettra d'autres études plus poussées ou plus ciblées, qui suivront le rythme des phases futures de déploiement des systèmes coopératifs.

Conclusions et perspectives

Compte tenu de la tendance actuelle, les véhicules autonomes remplaceront les véhicules d'aujourd'hui à l'horizon 2030-2050, et il sera alors nécessaire de repenser intégralement la mobilité. En effet, le conducteur sera remplacé par un pilote automatique et autonome, capable d'agir plus rapidement, tout en ayant une représentation plus précise et fiable de son environnement de navigation. Pourtant, il reste encore du chemin à parcourir avant d'arriver à un tel stade de maturité et de déploiement : l'environnement du véhicule est complexe, imprévisible et conflictuel, car partagé avec d'autres acteurs. La phase la plus importante, et qui a motivé notre travail, se situe entre cette projection dans laquelle la tâche de conduite sera complètement automatisée et la situation actuelle. Cette phase de transition voit déjà l'apparition de systèmes embarqués, communicants, partiellement ou complètement autonomes. La place du conducteur, ou tout du moins sa tâche, évoluera en conséquence et l'usage de l'infrastructure deviendra de plus en plus optimal. Cela signifie que ces systèmes intelligents devront être acceptés. Côté usagers, il sera nécessaire que ces systèmes soient compatibles avec leurs habitudes et leur procurent plus d'avantages que d'inconvénients. Côté gestionnaires et décideurs, il faudra que ces systèmes permettent un meilleur usage ou puissent par exemple améliorer la sécurité routière et réduire les externalités environnementales. Le déploiement futur des systèmes coopératifs est donc conditionné par leur potentiel, pour le moment difficile à évaluer (chapitre 1). Ce travail de thèse a tenté de proposer un cadre de modélisation capable de montrer le potentiel et les limitations de ces systèmes du point de vue du trafic et notamment en termes de stratégies de contrôle.

Synthèse des contributions et limitations

La première étape de notre travail a consisté à définir un cadre de modélisation des systèmes coopératifs (chapitre 2). L'échelle de modélisation était l'un des paramètres le plus décisif car dimensionnant le périmètre de notre étude. La complexité des interactions entre les véhicules et l'infrastructure nous ont orienté vers une modélisation microscopique du trafic. Grâce aux modèles dérivés de la théorie du trafic, la couche physique de ces systèmes peut être modélisée finement. Ensuite, les échanges d'information apportés par les technologies de la communication ont suscité l'intérêt d'utiliser les techniques de modélisation des systèmes distribués, notamment avec l'apport des systèmes multi-agents. Ce paradigme permet d'intégrer la partie prise de décision des véhicules (et de l'infrastructure) dans la modélisation. La distinction entre un système automatisé ou conduit par un humain est alors possible. Rapidement, nous avons identifié l'aspect critique de la fiabilité de l'information dans ces systèmes. Afin de garantir la robustesse du groupe de véhicules connectés, nous avons proposé d'introduire le concept de confiance dans les échanges entre les véhicules.

Ces trois dynamiques, modèles physiques, communication et confiance, font partie intégrante d'une architecture distribuée de modélisation par couplage. Afin d'illustrer les bénéfices de ce couplage, nous avons conçu un modèle bilatéral multi-anticipatif visant à homogénéiser un flux de véhicules partiellement automatisés et connectés. Cette modélisation permet également d'obtenir des bénéfices en termes de sécurité et d'émissions de polluants. Cette loi de poursuite multi-anticipative agit comme une stratégie de régulation locale du trafic, qui s'appuie sur le couplage des trois dynamiques pour adapter le comportement des véhicules et le rendre plus robuste aux perturbations sur autoroute.

L'étape suivante a été de développer un cadre de simulation propice à l'implémentation de nos propositions de modélisation (chapitre 3). Le choix d'une plateforme existante a été motivé par les compatibilités avec notre approche de modélisation. Le simulateur MASCAT (*Multi-Agent Simulator for Connected and Automated Traffic*), construit comme une extension d'une plateforme existante propose plusieurs contributions pour la simulation des C-ITS. Les flux d'informations entre les véhicules et avec l'infrastructure prennent la forme d'échanges de messages. Toutes les informations du système proviennent de capteurs, et toutes les entités, modélisées comme des agents, sont autonomes dans leur prise de décision. De nouvelles formes de contrôle sont désormais envisageables en utilisant des consignes transmises par l'infrastructure communicante. Le couplage des dynamiques assure la cohérence et l'inter-dépendance des différents modèles dans le simulateur. Afin d'évaluer les bénéfices et limitations apportées par les systèmes coopératifs, nous avons proposé un ensemble d'indicateurs dédiés. Les premiers résultats en simulation visent l'évaluation de la stratégie locale d'homogénéisation. Grâce à l'intégration d'informations supplémentaires (anticipation) via la communication, les véhicules sont capables de réduire l'effet de perturbations propagées au sein d'un flux. Un effet intéressant observé est la percolation du système avec contamination influençant les véhicules non-connectés, qui survient, suivant la configuration, autour d'un taux de pénétration de 20% à 30%. Ce résultat préfigure donc une amélioration substantielle de la qualité de l'écoulement des flux grâce aux C-ITS sans attendre les phases plus lointaines de déploiement (jusqu'à 100%). D'autre part, l'apparition de nouveaux comportements de ce type, ou d'une synergie au sein du système, nous oriente à proposer de nouvelles stratégies de contrôle.

Dans le chapitre 4, nous abordons la thématique du contrôle adaptée aux C-ITS. La stratégie locale de régulation (modèle multi-anticipatif bilatéral), bien que prometteuse, est limitée à un cadre relativement strict, et ne couvre donc pas des situations souvent génératrices de congestion. Lorsque le réseau est la cause de l'apparition de ralentissements (la chute de capacité), l'anticipation par les véhicules seuls n'est plus suffisante. Nous proposons donc d'étudier des stratégies plus globales, qui mettent à profit l'infrastructure. Une des problématiques principales est de garantir une forme de contrôle qui s'adaptera aux différentes phases de déploiement des C-ITS. L'analogie avec des problèmes de l'Intelligence Artificielle (problème de cognition) nous a menés à traiter le problème de manière plus abstraite : qui consiste à examiner comment permettre à un système autonome de contrôler son environnement. Pour cela, le système a besoin de construire une représentation précise de son environnement. Son comportement est dépendant des *a priori* donnés par le concepteur. Les approches constructivistes, que nous avons retenues, visent à limiter ce biais en modélisant le processus de cognition comme un phénomène de construction itératif. Pour le trafic coopératif, l'avantage est de disposer d'un système capable de générer ses propres stratégies, en utilisant ou non des connaissances expertes, et de les faire évoluer au cours du temps pour s'adapter aux véhicules composant le flux. Notre contribution consiste dans un premier temps à proposer un

modèle conceptuel générique. Notre approche constructiviste s'appuie sur des mécanismes faisant intervenir des représentations concurrentes encapsulées dans une population interne d'agents. Les rôles des agents permettent d'imaginer d'autres extensions comme l'ajout d'associations menant à un affinage de la représentation du système. Les résultats de notre approche sont présentés dans deux cadres de simulation. Le premier est un prototype visant à illustrer les comportements de bas niveau dans un environnement simplifié. Nous avons montré que le modèle est capable dans ce cadre de combiner différentes représentations individuelles pour construire une représentation et d'adapter à différents contextes en les recombinaison dynamiquement. La seconde étape de validation met à profit le cadre de simulation du trafic coopératif présenté dans notre travail. Nous avons choisi un scénario illustrant une situation complexe où le contrôle par Unité de Bord de Route des véhicules connectés peut avoir un effet bénéfique sur le flux. Les premiers résultats présentés laissent entrevoir le potentiel de notre approche dans de telles applications. Les études en simulation ont aussi pointé des axes d'amélioration possibles qui font partie des perspectives de nos travaux.

Les résultats présentés dans les différents chapitres, bien que prometteurs, mériteraient d'être comparés avec d'autres modèles de la littérature. Le modèle de confiance est comparé à une version simplifiée d'un autre modèle, dont le comportement complet devrait être étudié. Du côté des modèles de trafic (couche physique), notre étude s'est focalisée sur deux modèles précis (loi de poursuite et modèle de changement de voie). Les propriétés du modèle multi-anticipatif bilatéral devraient être valides en le couplant avec d'autres modèles, mais cette hypothèse reste à confirmer en simulation. L'idéal serait de choisir un nouveau modèle de changement de voie couplé à une loi de poursuite pour mener cette étude. L'application du modèle de contrôle à la régulation du trafic coopératif n'est qu'une première étape de validation. Il pourrait être intéressant de comparer dans différents scénarios les résultats obtenus avec notre approche face aux approches classiques (contrôle d'accès), qui devraient éprouver des difficultés à s'adapter aux flux de véhicules connectés. D'autres travaux plus récents méritent aussi d'être étudiés, ils portent sur des extensions des approches de régulation classiques dédiées aux nouveaux véhicules connectés (à base de VSL notamment). Une des perspectives à plus long terme de notre travail est de porter les modèles proposés dans un cadre applicatif réel, *i.e.* au niveau d'un véhicule équipé. Avant ce stade de maturité, plusieurs pistes pourraient significativement améliorer les résultats obtenus.

Perspectives

Le sujet traité dans notre thèse porte sur les phases successives du déploiement des C-ITS. Nous avons identifié plusieurs volets importants qui mériteraient de faire l'objet d'études spécifiques et avancées.

Vers l'auto-organisation du flux

La modélisation par couplage de dynamique proposée dans notre travail a été illustrée par l'intermédiaire d'un exemple faisant intervenir un flux mixte partiellement composé de véhicules connectés et semi-autonomes. L'architecture permet d'étendre la prise de décision des véhicules à d'autres applications plus complexes. Par exemple, nous envisageons d'intégrer davantage d'information dans la perception des véhicules. Ces informations pourraient provenir d'autres capteurs, donnant la possibilité au véhicule de voir les véhicules situés tout autour de sa carrosserie. Nous proposons

d'étendre le modèle bilatéral multi-anticipatif pour qu'il permette de contrôler, en plus du comportement longitudinal, le comportement latéral du véhicule. À la manière des approches bio-inspirées, il est possible de faire reposer le comportement des individus sur le comportement moyen observé de ses voisins afin de générer des comportements émergents (auto-organisation). Pour cela, il est nécessaire d'envisager un contrôle complètement autonome du véhicule. Plusieurs problèmes restent à résoudre. D'abord, les véhicules devront se coordonner de manière plus fine. Le cadre réglementaire doit lui aussi évoluer dans la mesure où les règles du code de la route ne seront pas adaptées à de tels comportements de groupe. La question est d'autant plus présente si le flux n'est pas composé que de véhicules autonomes. La communication pourra alors être étendue pour partager les intentions de tous les usagers (notamment disponibles sur les systèmes de navigation : itinéraires, voies). L'architecture de modélisation par couplage et le simulateur qui en est le fruit permettent de concevoir une telle stratégie et d'en évaluer les impacts.

Vers une stratégie de contrôle étendue

Dans le cadre d'une phase de déploiement plus proche, nous envisageons d'étendre la stratégie de contrôle décentralisée proposée. En effet, l'implémentation actuelle du modèle d'élaboration d'une représentation a été testée lors d'une implémentation au sein d'une Unité de Bord de Route. Les résultats étant encourageants, nous pensons que ce cadre pourrait être largement étendu. Dans un premier temps, l'idée est de permettre une coordination entre différentes Unités de Bord de Route. En effet, l'optimum local, possiblement atteint par adaptation de la représentation d'une UBR risque de se répercuter négativement sur d'autres parties du réseau. Par exemple, si une stratégie sur une section consiste à ralentir le flux, les systèmes en amont pourraient se voir propager des perturbations. De même, certains phénomènes peuvent être produits à des endroits du réseau qui ne sont originellement pas intégrés à la perception d'une UBR qui pourrait le traiter. Dans ce cas, le modèle devrait avoir la capacité de partager les données de la perception de manière étendue. La sélection (*i.e.* l'intégration des variables dans la représentation) de ces canaux de perception pourra être gérée dynamiquement. Enfin, une autre façon d'imaginer cette extension de perception est de mettre en place un mécanisme d'échange de la représentation, permettant à plusieurs UBR de partager les ressources mises en œuvre dans la construction. Notre modèle de confiance pourrait être intégré pour gérer ces échanges. Le contrôle deviendrait encore plus adaptatif et suivrait rapidement les phases de déploiement des systèmes coopératifs.

Bibliographie

- [1] Abdul-Rahman, A., and Hailes, S. Supporting trust in virtual communities. In *Proceedings of the 33rd Annual Hawaii International Conference on Systems Sciences* (2000), vol. 1, p. 9. [Page 41]
- [2] Adler, J. L., and Blue, V. J. A cooperative multi-agent transportation management and route guidance system. *Transportation Research Part C : Emerging Technologies 10* (2002), 433 – 454. [Page 37]
- [3] Adler, J. L., Satapathy, G., Manikonda, V., Bowles, B., and Blue, V. J. A multi-agent approach to cooperative traffic management and route guidance. *Transportation Research Part B : Methodological 39, 4* (2005), 297 – 318. [Page 37]
- [4] Ahn, S., and Cassidy, M. J. Freeway traffic oscillations and vehicle lane-change maneuvers. In *International Symposium on Transportation & Traffic Theory* (2007), M. B. R. Alsop and B. Heydecker, Eds., vol. 1, pp. 691–710. [Page 30]
- [5] AIRPARIF. Surveillance & information sur la qualité de l'air en île-de-france – bilan 2015. Tech. rep., 2016. [Page 1]
- [6] Alonso, E., D'inverno, M., Kudenko, D., Luck, M., and Noble, J. Learning in multi-agent systems. *The Knowledge Engineering Review 16, 03* (2001), 277–284. [Pages 97, 98]
- [7] Asada, M., Hosoda, K., Kuniyoshi, Y., Ishiguro, H., Inui, T., Yoshikawa, Y., Ogino, M., and Yoshida, C. Cognitive developmental robotics : a survey. *IEEE Transactions on Autonomous Mental Development 1, 1* (2009), 12–34. [Page 92]
- [8] Auer, P., Cesa-Bianchi, N., and Fischer, P. Finite-time analysis of the multiarmed bandit problem. *Machine learning 47, 2-3* (2002), 235–256. [Page 110]
- [9] Auer, P., Cesa-Bianchi, N., Freund, Y., and Schapire, R. E. Gambling in a rigged casino : The adversarial multi-armed bandit problem. In *Foundations of Computer Science, 1995. Proceedings., 36th Annual Symposium on* (1995), IEEE, pp. 322–331. [Page 110]
- [10] Aw, A., and Rascle, M. Resurrection of "second order" models of traffic flow? *SIAM journal on applied mathematics 60* (2000), 916–938. [Page 24]
- [11] Backfrieder, C., Mecklenbrauker, C., and Ostermayer, G. Traffsim - a traffic simulator for investigating benefits ensuing from intelligent traffic management. In *Modelling Symposium (EMS), 2013 European* (Nov 2013), pp. 451–456. [Page 52]
- [12] Balmer, M., Rieser, M., Meister, K., Charypar, D., Lefebvre, N., Nagel, K., and Axhausen, K. Matsim-t : Architecture and simulation times. *Multi-agent systems for traffic and transportation engineering* (2009), 57–78. [Page 52]

- [13] Bando, M., Hasebe, K., Nakayama, A., Shibata, A., and Sugiyama, Y. Dynamical model of traffic congestion and numerical simulation. *Phys. Rev. E* 51 (1995), 1035–1042. [Pages 26, 55]
- [14] Barceló, J., and Casas, J. Dynamic network simulation with aimsun. In *Simulation approaches in transportation analysis*. Springer, 2005, pp. 57–98. [Page 52]
- [15] Barrachina, J., Garrido, P., Fogue, M., Martinez, F. J., Cano, J.-C., Calafate, C. T., and Manzoni, P. Road side unit deployment : A density-based approach. *IEEE Intelligent Transportation Systems Magazine* 5, 3 (2013), 30–39. [Page 8]
- [16] Bazzan, A. L., and Klügl, F. A review on agent-based technology for traffic and transportation. *The Knowledge Engineering Review* 29, 03 (2014), 375–403. [Pages 36, 52]
- [17] Bazzan, A. L. C. Opportunities for multiagent systems and multiagent reinforcement learning in traffic control. *Autonomous Agents and Multi-Agent Systems* 18, 3 (2009), 342–375. [Page 89]
- [18] Bazzan, A. L. C., Wahle, J., and Klügl, F. Agents in traffic modelling – from reactive to social behaviour. In *Annual Conference on Artificial Intelligence* (1999), pp. 303–306. [Page 52]
- [19] Beckers, R., Holland, O., and Deneubourg, J.-L. From local actions to global tasks : Stigmergy and collective robotics. In *Artificial life IV* (1994), vol. 181, p. 189. [Page 34]
- [20] Billot, R., El Faouzi, N.-E., Guériau, M., and Monteil, J. Can c-its lead to the emergence of traffic management 2.0 ? In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)* (2014), IEEE, pp. 483–488. [Page 7]
- [21] Bloembergen, D., Tuyls, K., Hennes, D., and Kaisers, M. Evolutionary dynamics of multi-agent learning : a survey. *Journal of Artificial Intelligence Research* 53 (2015), 659–697. [Page 98]
- [22] Bourjot, C., Desor, D., and Chevrier, V. Stigmergy. In *Self-organising Software*. Springer, 2011, pp. 123–138. [Page 34]
- [23] Brooks, R. A. Intelligence without representation. *Artificial intelligence* 47, 1 (1991), 139–159. [Page 103]
- [24] Burmeister, B., Haddadi, A., and Matylis, G. Application of multi-agent systems in traffic and transportation. In *Software Engineering. IEE Proceedings* (1997), vol. 144, IET, pp. 51–60. [Page 36]
- [25] Carbo, J., Molina, J. M., and Davila, J. Trust management through fuzzy reputation. *International Journal of Cooperative Information Systems* 12, 01 (2003), 135–155. [Page 41]
- [26] Carlson, R. C., Papamichail, I., and Papageorgiou, M. Local feedback-based mainstream traffic flow control on motorways using variable speed limits. *IEEE Transactions on Intelligent Transportation Systems* 12, 4 (2011), 1261–1276. [Page 89]
- [27] Carter, J., Bitting, E., and Ghorbani, A. A. Reputation formalization for an information-sharing multi-agent system. *Computational Intelligence* 18, 4 (2002), 515–534. [Page 41]
- [28] Castelfranchi, C. Modelling social action for ai agents. *Artificial Intelligence* 103, 1 (1998), 157–182. [Page 34]
- [29] Castelfranchi, C., and Falcone, R. Social trust : Cognitive anatomy, social importance, quantification and dynamics. In *Proceedings of the First International Workshop on Trust* (1998), pp. 35–49. [Pages 37, 41]
- [30] Chapelle, O., and Li, L. An empirical evaluation of thompson sampling. In *Advances in neural information processing systems* (2011), pp. 2249–2257. [Page 110]
- [31] Chen, B., and Cheng, H. H. A review of the applications of agent technology in traffic and transportation systems. *IEEE Transactions on Intelligent Transportation Systems* 11 (2010), 485–497. [Page 36]

- [32] Chen, B., Cheng, H. H., and Palen, J. Integrating mobile agent technology with multi-agent systems for distributed traffic detection and management systems. *Transportation Research Part C : Emerging Technologies* 17, 1 (2009), 1 – 10. [Page 37]
- [33] Chen, R.-S., Chen, D.-K., and Lin, S.-Y. Actam : Cooperative multi-agent system architecture for urban traffic signal control. *IEICE Transactions 88-D*, 1 (2005), 119–126. [Page 36]
- [34] Choi, J., Oh, S., and Horowitz, R. Distributed learning and cooperative control for multi-agent systems. *Automatica* 45, 12 (2009), 2802–2814. [Page 70]
- [35] Corkill, D. D. Collaborating software : Blackboard and multi-agent systems & the future. In *Proceedings of the International Lisp Conference* (2003), vol. 10. [Pages 33, 34]
- [36] Cremer, M., and Ludwig, J. A fast simulation model for traffic flow on the basis of boolean operations. *Mathematics and Computers in Simulation* 28, 4 (1986), 297–303. [Page 25]
- [37] Daganzo, C. F. Requiem for second-order fluid approximations of traffic flow. *Transportation Research Part B : Methodological* 29, 4 (1995), 277 – 286. [Page 24]
- [38] Daganzo, C. F., Gayah, V. V., and Gonzales, E. J. Macroscopic relations of urban traffic variables : Bifurcations, multivaluedness and instability. *Transportation Research Part B : Methodological* 45, 1 (2011), 278 – 288. [Pages 23, 24]
- [39] Daganzo, C. F., and Geroliminis, N. An analytical approximation for the macroscopic fundamental diagram of urban traffic. *Transportation Research Part B : Methodological* 42, 9 (2008), 771 – 781. [Page 23]
- [40] Davidsson, P., Henesey, L., Ramstedt, L., Törnquist, J., and Wernstedt, F. An analysis of agent-based approaches to transport logistics. *Transportation Research Part C : Emerging Technologies* 13, 4 (2005), 255 – 271. [Page 36]
- [41] Davis, L. Effect of adaptive cruise control systems on mixed traffic flow near an on-ramp. *Physica A : Statistical Mechanics and its Applications* 379, 1 (2007), 274–290. [Page 29]
- [42] de Bruin, D., Kroon, J., van Klaveren, R., and Nelisse, M. Design and test of a cooperative adaptive cruise control system. In *Intelligent Vehicles Symposium, 2004 IEEE* (2004), IEEE, pp. 392–396. [Page 29]
- [43] De Oliveira, L. B., and Camponogara, E. Multi-agent model predictive control of signaling split in urban traffic networks. *Transportation Research Part C : Emerging Technologies* 18, 1 (2010), 120 – 139. [Page 37]
- [44] del Castillo, J. M. Three new models for the flow–density relationship : derivation and testing for freeway and urban data. *Transportmetrica* 8, 6 (2012), 443–465. [Page 23]
- [45] Demazeau, Y. From interactions to collective behaviour in agent-based systems. In *In : Proceedings of the 1st. European Conference on Cognitive Science. Saint-Malo* (1995), Citeseer. [Pages 32, 34]
- [46] Deneubourg, J.-L., Goss, S., Franks, N., Sendova-Franks, A., Detrain, C., and Chrétien, L. The dynamics of collective sorting robot-like ants and ant-like robots. In *Proceedings of the first international conference on simulation of adaptive behavior on From animals to animats* (1991), pp. 356–363. [Page 31]
- [47] Desai, P., Loke, S., Desai, A., and Singh, J. Multi-agent based vehicular congestion management. In *Intelligent Vehicles Symposium (IV), 2011 IEEE* (2011), pp. 1031–1036. [Page 37]
- [48] Dessalles, J. L., Ferber, J., and Phan, D. *Intelligent complex adaptive systems*. Y. Shyan and A. Yang, 2007, ch. Emergence in Agent based Computational Social Science : conceptual, formal and diagrammatic analysis., pp. 255–299. [Page 35]

- [49] Di Marzo Serugendo, G., Gleizes, M.-P., and Karageorgos, A. Self-organization in multi-agent systems. *The Knowledge Engineering Review* 20, 02 (2005), 165–189. [Page 35]
- [50] Di Marzo Serugendo, G., Gleizes, M.-P., and Karageorgos, A. *Self-organising software : From natural to artificial adaptation*. Springer Science & Business Media, 2011. [Page 35]
- [51] Doniec, A., Espié, S., Mandiau, R., and Piechowiak, S. Multi-agent coordination and anticipation model to design a road traffic simulation tool. In *EUMAS* (2006). [Page 31]
- [52] Doniec, A., Mandiau, R., Espié, S., and Piechowiak, S. Comportements anticipatifs dans les systèmes multi-agents : application à la simulation de trafic routier. *Revue d'Intelligence Artificielle (RIA)* 21, 7 (2007), 183–221. [Page 36]
- [53] Doniec, A., Mandiau, R., Piechowiak, S., and Espié, S. A behavioral multi-agent model for road traffic simulation. *Engineering Applications of Artificial Intelligence* 21, 8 (2008), 1443 – 1454. [Page 36]
- [54] Drescher, G. L. *Made-up minds : a constructivist approach to artificial intelligence*. MIT press, 1991. [Page 92]
- [55] Dresner, K., and Stone, P. Multiagent traffic management : A reservation-based intersection control mechanism. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 2* (2004), IEEE Computer Society, pp. 530–537. [Page 89]
- [56] Dupuis, M., and Grezlikowski, H. Opendrive®-an open standard for the description of roads in driving simulations. In *Proceedings of the Driving Simulation Conference* (2006), pp. 25–36. [Pages 55, 65]
- [57] Edie, L. C. Discussion on traffic stream measurements and definitions. In *The 2nd International Symposium on the Theory of Traffic Flow* (1963), pp. 139–154. [Page 22]
- [58] Esfandiari, B., and Chandrasekharan, S. On how agents make friends : Mechanisms for trust acquisition. In *Proceedings of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies* (2001), pp. 27–34. [Page 41]
- [59] Espié, S., Auberlet, J. M., et al. Archisim : A behavioral multi-actors traffic simulation model for the study of a traffic system including its aspects. *International Journal of ITS Research*, n1 (2007). [Page 52]
- [60] Even-Dar, E., Mannor, S., and Mansour, Y. Pac bounds for multi-armed bandit and markov decision processes. In *International Conference on Computational Learning Theory* (2002), Springer, pp. 255–270. [Page 110]
- [61] Even-Dar, E., Mannor, S., and Mansour, Y. Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. *Journal of machine learning research* 7, Jun (2006), 1079–1105. [Page 110]
- [62] Fellendorf, M. Vissim : A microscopic simulation tool to evaluate actuated signal control including bus priority. In *64th Institute of Transportation Engineers Annual Meeting* (1994), pp. 1–9. [Page 52]
- [63] Ferber, J. *Les Systèmes Multi-Agents : Vers une intelligence collective*. InterEdition, 1995. [Pages 31, 32, 38]
- [64] Ferber, J. *Multi-agent systems : an introduction to distributed artificial intelligence*, vol. 1. Addison-Wesley Reading, 1999. [Pages 32, 33]

- [65] Figueiredo, L., Jesus, I., Machado, J., Ferreira, J., and de Carvalho, J. Towards the development of intelligent transportation systems. In *Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE* (2001), pp. 1206–1211. [Page 36]
- [66] Fogue, M., Garrido, P., Martinez, F. J., Cano, J.-C., Calafate, C. T., Manzoni, P., and Sanchez, M. Prototyping an automatic notification scheme for traffic accidents in vehicular networks. In *IFIP Wireless Days (WD)*. 2011. [Page 7]
- [67] Fritzke, B. A growing neural gas network learns topologies. In *Advances in Neural Information Processing Systems 7*, G. Tesauro, D. Touretzky, and T. Leen, Eds. MIT Press, 1995, pp. 625–632. [Pages 101, 106]
- [68] Fritzke, B. A self-organizing network that can follow non-stationary distributions. In *Artificial Neural Networks – ICANN'97*, W. Gerstner, A. Germond, M. Hasler, and J.-D. Nicoud, Eds., vol. 1327 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 1997, pp. 613–618. [Pages 101, 127]
- [69] Fulda, N., and Ventura, D. Predicting and preventing coordination problems in cooperative q-learning systems. In *IJCAI* (2007), vol. 2007, pp. 780–785. [Page 99]
- [70] Gabbai, J., Yin, H., Wright, W. A., and Allinson, N. Self-organization, emergence and multi-agent systems. In *Neural Networks and Brain, 2005. ICNN B '05. International Conference on* (2005), vol. 3, pp. 1824–1863. [Page 35]
- [71] Ge, H., Dai, S., and Dong, L. An extended car-following model based on intelligent transportation system application. *Physica A : Statistical Mechanics and its Applications* 365, 2 (2006), 543 – 548. [Page 31]
- [72] Geroliminis, N., Srivastava, A., and Michalopoulos, P. G. Experimental observations of capacity drop phenomena in freeway merges with ramp metering control and integration in a first-order model. In *Transportation Research Board 90th Annual Meeting* (2011), no. 11-0748. [Page 129]
- [73] Geroliminis, N., and Sun, J. Properties of a well-defined macroscopic fundamental diagram for urban traffic. *Transportation Research Part B : Methodological* 45, 3 (2011), 605 – 617. [Pages 23, 24]
- [74] Gershenson, C., and Helbing, D. When slower is faster. *Complexity* 21, 2 (2015), 9–15. [Page 43]
- [75] Gettman, D., and Head, L. Surrogate safety measures from traffic simulation models. *Transportation Research Record : Journal of the Transportation Research Board*, 1840 (2003), 104–115. [Page 70]
- [76] Gipps, P. A behavioural car-following model for computer simulation. *Transportation Research Part B : Methodological* 15, 2 (1981), 105 – 111. [Pages 25, 55]
- [77] Goodfellow, I. J., Mirza, M., Xiao, D., Courville, A., and Bengio, Y. An empirical investigation of catastrophic forgetting in gradient-based neural networks. In *Proceedings of the International Conference on Learning Representations (ICLR) 2014* (2014). [Page 102]
- [78] Graczyk, M., Lasota, T., Trawiński, B., and Trawiński, K. Comparison of bagging, boosting and stacking ensembles applied to real estate appraisal. In *Asian Conference on Intelligent Information and Database Systems* (2010), Springer, pp. 340–350. [Pages 98, 111]
- [79] Greenshields, B. D. The photographic method of studying traffic behavior. In *Highway Research Board Proceedings* (1934), vol. 13. [Pages 20, 22]

- [80] Guériau, M., Billot, R., Armetta, F., Hassas, S., and El Faouzi, N.-E. Un simulateur multiagent de trafic coopératif. In *23es Journées Francophones sur les Systèmes Multi-Agents (JFSMA'15)* (2015), Cépaduès, pp. 165–174. [Pages 73, 81]
- [81] Guériau, M., Billot, R., El Faouzi, N.-E., Hassas, S., and Armetta, F. Multi-agent dynamic coupling for cooperative vehicles modeling. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (DEMO Track)* (2015), AAAI Press, pp. 4276–4277. [Pages 55, 73, 81]
- [82] Guériau, M., Billot, R., El Faouzi, N.-E., Monteil, J., Armetta, F., and Hassas, S. How to assess the benefits of connected vehicles ? a simulation framework for the design of cooperative traffic management strategies. *Transportation Research Part C : Emerging Technologies 67* (2016), 266 – 279. [Page 42]
- [83] Guériau, M., Billot, R., Hassas, S., Armetta, F., and El Faouzi, N.-E. An extension of movsim for multi-agent cooperative vehicles modeling. In *2014 International Conference on Connected Vehicles and Expo (ICCVE)* (2014), IEEE, pp. 859–860. [Page 55]
- [84] Guerin, F. Learning like a baby : a survey of artificial intelligence approaches. *The Knowledge Engineering Review 26*, 02 (2011), 209–236. [Page 92]
- [85] Hannoun, M., Sichman, J. S., Boissier, O., and Sayettat, C. Dependence relations between roles in a multi-agent system. In *International Workshop on Multi-Agent Systems and Agent-Based Simulation* (1998), Springer, pp. 169–182. [Page 35]
- [86] Hassas, S., Di Marzo-Serugendo, G., Karageorgos, A., and Castelfranchi, C. Self-organising mechanisms from social and business/economics approaches. *Informatica 30*, 1 (2006). [Page 35]
- [87] Hayward, J. C. Near-miss determination through use of a scale of danger. *Highway Research Record*, 384 (1972). [Page 70]
- [88] Helbing, D. Traffic and related self-driven many-particle systems. *Reviews of modern physics 73*, 4 (2001), 1067. [Pages 20, 24]
- [89] Helbing, D. Derivation of a fundamental diagram for urban traffic flow. *The European Physical Journal B 70*, 2 (2009), 229–241. [Page 23]
- [90] Hernandez-Leal, P., Taylor, M. E., Rosman, B., Sucar, L. E., and de Cote, E. M. Identifying and tracking switching, non-stationary opponents : a bayesian approach. In *Workshops at the Thirtieth AAAI Conference on Artificial Intelligence* (2016). [Page 102]
- [91] Herzig, A., Lorini, E., Hübner, J. F., Ben-Naim, J., Castelfranchi, C., Demolombe, R., Longin, D., Vercouter, L., and Boissier, O. Prolegomena for a logic of trust and reputation. *NorMAS 8* (2008), 143–157. [Page 41]
- [92] Hoogendoorn, S. P., and Bovy, P. H. State-of-the-art of vehicular traffic flow modelling. *Proceedings of the Institution of Mechanical Engineers, Part I : Journal of Systems and Control Engineering 215*, 4 (2001), 283–303. [Pages 22, 25]
- [93] Hoogendoorn, S. P., and Minderhoud, M. M. Motorway flow quality impacts of advanced driver assistance systems. *Transportation Research Record : Journal of the Transportation Research Board 1800* (2002), 69–77. [Page 5]
- [94] Huynh, T. D., Jennings, N. R., and Shadbolt, N. R. An integrated trust and reputation model for open multi-agent systems. *Autonomous Agents and Multi-Agent Systems 13*, 2 (2006), 119–154. [Pages 40, 41, 78]

- [95] Illner, R., Kirchner, C., and Pinneau, R. A derivation of the aw-rasclle traffic models from fokker-planck type kinetic models. *Quant. Appl. Math* 27 (2009), 13–45. [Page 24]
- [96] Jakobi, N., Husbands, P., and Harvey, I. Noise and the reality gap : The use of simulation in evolutionary robotics. In *European Conference on Artificial Life* (1995), Springer, pp. 704–720. [Page 117]
- [97] Jennings, N. R. On agent-based software engineering. *Artificial intelligence* 117, 2 (2000), 277–296. [Page 31]
- [98] Jiang, R., Wu, Q., and Zhu, Z. Full velocity difference model for a car-following theory. *Phys. Rev. E* 64 (Jun 2001), 017101. [Page 26]
- [99] Jøsang, A., Ismail, R., and Boyd, C. A survey of trust and reputation systems for online service provision. *Decision support systems* 43, 2 (2007), 618–644. [Page 41]
- [100] Kaelbling, L. P., Littman, M. L., and Moore, A. W. Reinforcement learning : A survey. *Journal of artificial intelligence research* 4 (1996), 237–285. [Page 96]
- [101] Kakkasageri, M., and Manvi, S. Information management in vehicular ad hoc networks : A review. *Journal of Network and Computer Applications* 39 (2014), 334–350. [Page 8]
- [102] Kaufmann, E., Cappé, O., and Garivier, A. On bayesian upper confidence bounds for bandit problems. In *AISTATS* (2012), pp. 592–600. [Page 110]
- [103] Kesting, A., and Treiber, M. Calibrating car-following models by using trajectory data : Methodological study. *Transportation Research Record : Journal of the Transportation Research Board* 2088 (2008), 148–156. [Page 26]
- [104] Kesting, A., Treiber, M., and Helbing, D. General lane-changing model mobil for car-following models. *Transportation Research Record : Journal of the Transportation Research Board* 1999 / 2007 Traffic Flow Theory 2007 (2007), 86–94. [Pages 27, 44, 45, 46, 55, 62, 81, 131]
- [105] Kesting, A., Treiber, M., and Helbing, D. *Multi-Agent Systems : Simulation and Applications*. 2009, ch. Agents for Traffic Simulation, pp. 325–356. [Pages 52, 53]
- [106] Kesting, A., Treiber, M., and Helbing, D. Connectivity statistics of store-and-forward intervehicle communication. *Intelligent Transportation Systems, IEEE Transactions on* 11, 1 (2010), 172–181. [Page 53]
- [107] Kesting, A., Treiber, M., Schönhof, M., and Helbing, D. Adaptive cruise control design for active congestion avoidance. *Transportation Research Part C : Emerging Technologies* 16, 6 (2008), 668–683. [Page 89]
- [108] Kesting, A., Trieber, M., and Helbing, D. Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity. *Philosophical Transactions of the Royal Society A : Mathematical Physical & Engineering Sciences* 368, 1928 (October 2010), 4585–4605. [Pages 26, 28, 45, 55]
- [109] Khondaker, B., and Kattan, L. Variable speed limit : an overview. *Transportation Letters* 7, 5 (2015), 264–278. [Page 89]
- [110] Knospe, W., Santen, L., Schadschneider, A., and Schreckenberg, M. Empirical test for cellular automaton models of traffic flow. *Physical Review E* 70, 1 (2004), 016115. [Page 25]
- [111] Kohonen, T. *Self-Organizing Maps*, vol. 30 of *Springer Series in Information Sciences*. 1995. [Pages 101, 106]
- [112] Koos, S., Mouret, J.-B., and Doncieux, S. The transferability approach : Crossing the reality gap in evolutionary robotics. *IEEE Transactions on Evolutionary Computation* 17, 1 (2013), 122–145. [Page 54]

- [113] Kosonen, I. Multi-agent fuzzy signal control based on real-time simulation. *Transportation Research Part C : Emerging Technologies* 11, 5 (2003), 389 – 403. World Congress on Intelligent Transport Systems. [Page 36]
- [114] Krajzewicz, D., Erdmann, J., Behrisch, M., and Bieker, L. Recent development and applications of SUMO - Simulation of Urban MObility. *International Journal On Advances in Systems and Measurements* 5, 3&4 (December 2012), 128–138. [Pages 52, 53]
- [115] Kuipers, B. J., Beeson, P., Modayil, J., and Provost, J. Bootstrap learning of foundational representations. *Connection Science* 18, 2 (2006), 145–158. [Page 101]
- [116] Laval, J. A., and Daganzo, C. F. Lane-changing in traffic streams. *Transportation Research Part B : Methodological* 40, 3 (2006), 251 – 264. [Page 26]
- [117] Leclercq, L., Knoop, V. L., Marczak, F., and Hoogendoorn, S. P. Capacity drops at merges : New analytical investigations. *Transportation Research Part C : Emerging Technologies* 62 (2016), 171–181. [Page 129]
- [118] Lenz, H., Wagner, C., and Sollacher, R. Multi-anticipative car-following model. *The European Physical Journal B - Condensed Matter and Complex Systems* 7, 2 (1999), 331–335. [Pages 25, 30]
- [119] Li, K., and Ioannou, P. Modeling of traffic flow of automated vehicles. *Intelligent Transportation Systems, IEEE Transactions on* 5, 2 (2004), 99–113. [Page 24]
- [120] Li, Z., and Zhang, R. An extended non-lane-based optimal velocity model with dynamic collaboration. *Mathematical Problems in Engineering* 2013 (2013), 8. [Page 26]
- [121] Lighthill, M. J., and Whitham, G. B. On kinematic waves. I. flood movement in long rivers. In *Proceedings of the Royal Society of London A : Mathematical, Physical and Engineering Sciences* (1955), vol. 229, The Royal Society, pp. 281–316. [Page 23]
- [122] Lighthill, M. J., and Whitham, G. B. On kinematic waves. II. a theory of traffic flow on long crowded roads. In *Proceedings of the Royal Society of London A : Mathematical, Physical and Engineering Sciences* (1955), vol. 229, The Royal Society, pp. 317–345. [Page 23]
- [123] Logi, F., and Ritchie, S. G. A multi-agent architecture for cooperative inter-jurisdictional traffic congestion management. *Transportation Research Part C : Emerging Technologies* 10 (2002), 507 – 527. [Page 37]
- [124] Lungarella, M., Metta, G., Pfeifer, R., and Sandini, G. Developmental robotics : a survey. *Connection Science* 15, 4 (2003), 151–190. [Page 92]
- [125] Mabrouk, A., Senhadji, M., Kobbane, A., Walid, A., Sabir, E., and EL Koutbi, M. A congestion game-based routing algorithm for communicating vanets. In *Wireless Communications and Mobile Computing Conference (IWCMC), 2014 International* (2014), IEEE, pp. 672–677. [Page 89]
- [126] Mahmassani, H. S. Dynamic network traffic assignment and simulation methodology for advanced system management applications. *Networks and spatial economics* 1, 3-4 (2001), 267–292. [Page 52]
- [127] Mano, J.-P., Bourjot, C., Lopardo, G., and Glize, P. Bio-inspired mechanisms for artificial self-organised systems. *Informatica* 30, 1 (2006), 55–62. [Page 35]
- [128] Marsden, G., McDonald, M., and Brackstone, M. Towards an understanding of adaptive cruise control. *Transportation Research Part C : Emerging Technologies* 9, 1 (2001), 33–51. [Page 29]

- [129] Marsh, S. P. *Formalising Trust as a Computational Concept*. PhD thesis, 1994. [Page 41]
- [130] Martinez, F. J., Toh, C. K., Cano, J.-C., Calafate, C. T., and Manzoni, P. A survey and comparative study of simulators for vehicular ad hoc networks (vanets). *Wireless Communications and Mobile Computing* 11, 7 (2011), 813–828. [Page 52]
- [131] Mazac, S., Armetta, F., and Hassas, S. On bootstrapping sensori-motor patterns for a constructivist learning system in continuous environments. In *Alife 14 : Fourteenth International Conference on the Synthesis and Simulation of Living Systems* (2014). [Pages 92, 97, 101]
- [132] Meeden, L. A., and Blank, D. S. Introduction to developmental robotics. *Connection Science* 18, 2 (2006), 93–96. [Page 92]
- [133] Milanés, V., and Shladover, S. E. Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data. *Transportation Research Part C : Emerging Technologies* 48 (2014), 285–300. [Page 75]
- [134] Milanés, V., Shladover, S. E., Spring, J., Nowakowski, C., Kawazoe, H., and Nakamura, M. Cooperative adaptive cruise control in real traffic situations. *IEEE Transactions on Intelligent Transportation Systems* 15, 1 (2014), 296–305. [Page 30]
- [135] Minderhoud, M. M., and Bovy, P. H. Extended time-to-collision measures for road traffic safety assessment. *Accident Analysis & Prevention* 33, 1 (2001), 89–97. [Pages 70, 78]
- [136] Monteil, J., Billot, R., Sau, J., Armetta, F., Hassas, S., and El Faouzi, N.-E. Cooperative highway traffic : multiagent modeling and robustness assessment of local perturbations. *Transportation Research Record : Journal of the Transportation Research Board*, 2391 (2013), 1–10. [Pages 42, 48]
- [137] Monteil, J., Billot, R., Sau, J., Buisson, C., and El Faouzi, N.-E. Calibration, estimation, and sampling issues of car-following parameters. *Transportation Research Record : Journal of the Transportation Research Board*, 2422 (2014), 131–140. [Pages 26, 72]
- [138] Monteil, J., Billot, R., Sau, J., and El Faouzi, N.-E. Linear and weakly nonlinear stability analyses of cooperative car-following models. *IEEE Transactions on Intelligent Transportation Systems* 15, 5 (2014), 2001–2013. [Pages 44, 46]
- [139] Mugan, J., and Kuipers, B. Learning distinctions and rules in a continuous world through active exploration. In *Proceedings of the Seventh International Conference on Epigenetic Robotics (EpiRob-07)* (2007), pp. 101–108. [Page 92]
- [140] Mugan, J., and Kuipers, B. Learning to predict the effects of actions : Synergy between rules and landmarks. In *Proceedings of the Sixth IEEE International Conference on Development and Learning (ICDL-07)* (2007), pp. 253–258. [Page 101]
- [141] Mugan, J., and Kuipers, B. Autonomous representation learning in a developing agent. In *Computational and Robotic Models of the Hierarchical Organization of Behavior*. Springer, 2013, pp. 63–80. [Page 100]
- [142] Mui, L., Mohtashemi, M., and Halberstadt, A. A computational model of trust and reputation. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences* (2002), pp. 2431–2439. [Page 41]
- [143] Muller, G., and Vercouter, L. Decentralized monitoring of agent communications with a reputation model. In *Trusting Agents for Trusting Electronic Societies*. Springer, 2005, pp. 144–161. [Page 41]
- [144] Nagel, K., and Schreckenberg, M. A cellular automaton model for freeway traffic. *Journal de physique I* 2, 12 (1992), 2221–2229. [Page 25]

- [145] Najjar, A., and Reignier, P. Constructivist ambient intelligent agent for smart environments. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2013 IEEE International Conference on* (2013), IEEE, pp. 356–359. [Page 92]
- [146] Newell, G. F. A simplified car-following theory : a lower order model. *Transportation Research Part B : Methodological* 36, 3 (2002), 195 – 205. [Pages 26, 55]
- [147] Ngoduy, D., Hoogendoorn, S., and Liu, R. Continuum modeling of cooperative traffic flow dynamics. *Physica A : Statistical Mechanics and its Applications* 388, 13 (2009), 2705–2716. [Page 28]
- [148] Ngoduy, D., and Wilson, R. Multianticipative nonlocal macroscopic traffic model. *Computer-Aided Civil and Infrastructure Engineering* 29, 4 (2014), 248–263. [Page 28]
- [149] Nguyen, V. Q.-A., Canal, R., Gaudou, B., Hassas, S., and Armetta, F. Trustsets : using trust to detect deceitful agents in a distributed information collecting system. *Journal of Ambient Intelligence and Humanized Computing* 3, 4 (2012), 251–263. [Pages 40, 41, 48]
- [150] Odell, J. J., Parunak, H. V. D., Fleischer, M., and Brueckner, S. Modeling agents and their environment. In *Agent-oriented software engineering III*. Springer, 2002, pp. 16–31. [Page 33]
- [151] Olfati-Saber, R., Fax, J. A., and Murray, R. M. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE* 95, 1 (2007), 215–233. [Page 38]
- [152] Öncü, S., van de Wouw, N., Heemels, W. M. H., and Nijmeijer, H. String stability of interconnected vehicles under communication constraints. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)* (2012), IEEE, pp. 2459–2464. [Page 29]
- [153] Onishi, H., and Mlinarsky, F. Wireless technology assessment for automotive applications. In *Proceedings of the ITS World Congress* (2012). [Pages 8, 59]
- [154] Ossen, S., and Hoogendoo, S. Multi-anticipation and heterogeneity in car-following empirics and a first exploration of their implications. In *2006 IEEE Intelligent Transportation Systems Conference* (2006), IEEE, pp. 1615–1620. [Page 30]
- [155] Padovan, B., Sackmann, S., Eymann, T., and Pippow, I. A prototype for an agent-based secure electronic marketplace including reputation tracking mechanisms. In *Proceedings of the 34th Annual Hawaii International Conference on System Sciences* (2001), p. 10. [Page 41]
- [156] Panait, L., and Luke, S. Cooperative multi-agent learning : The state of the art. *Autonomous agents and multi-agent systems* 11, 3 (2005), 387–434. [Pages 95, 97]
- [157] Papadimitratos, P., De La Fortelle, A., Evensen, K., Brignolo, R., and Cosenza, S. Vehicular communication systems : Enabling technologies, applications, and future outlook on intelligent transportation. *IEEE Communications Magazine* 47, 11 (2009), 84–95. [Page 7]
- [158] Papageorgiou, M., Diakaki, C., Dinopoulou, V., Kotsialos, A., and Wang, Y. Review of road traffic control strategies. *Proceedings of the IEEE* 91, 12 (2003), 2043–2067. [Page 88]
- [159] Papageorgiou, M., Hadj-Salem, H., and Blosseville, J.-M. Alinea : A local feedback control law for on-ramp metering. *Transportation Research Record*, 1320 (1991). [Page 5]
- [160] Paveri-Fontana, S. On boltzmann-like treatments for traffic flow : A critical review of the basic model and an alternative proposal for dilute traffic analysis. *Transportation Research* 9, 4 (1975), 225 – 235. [Page 24]
- [161] Piaget, J. The construction of reality in the child. *Journal of Consulting Psychology* 19, 1 (1955), 77. [Page 91]

- [162] Pinyol, I., and Sabater-Mir, J. Pragmatic-strategic reputation-based decisions in bdi agents. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems* (2009), vol. 2, International Foundation for Autonomous Agents and Multiagent Systems, pp. 1001–1008. [Page 41]
- [163] Pinyol, I., and Sabater-Mir, J. Computational trust and reputation models for open multi-agent systems : a review. *Artificial Intelligence Review* 40, 1 (2013), 1–25. [Page 39]
- [164] Pinyol, I., Sabater-Mir, J., Dellunde, P., and Paolucci, M. Reputation-based decisions for logic-based cognitive agents. *Autonomous Agents and Multi-Agent Systems* 24, 1 (2012), 175–216. [Page 41]
- [165] Ploeg, J., Scheepers, B. T., van Nunen, E., van de Wouw, N., and Nijmeijer, H. Design and experimental evaluation of cooperative adaptive cruise control. In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)* (2011), IEEE, pp. 260–265. [Page 29]
- [166] Prigogine, I., and Andrews, F. C. A boltzmann-like approach for traffic flow. *Operations Research* 8 (1960), 789–797. [Page 24]
- [167] Prigogine, I., Herman, R., and Schechter, R. Kinetic theory of vehicular traffic. *Systems, Man and Cybernetics, IEEE Transactions on* 2, 2 (1972). [Page 24]
- [168] Provost, J., Kuipers, B. J., and Miikkulainen, R. Self-organizing distinctive state abstraction using options. In *Proc. of the 7th International Conference on Epigenetic Robotics* (2007). [Page 101]
- [169] Rasmusson, L., Rasmusson, A., and Janson, S. Using agents to secure the internet marketplace reactive security and social control. In *Proceedings of PAAM* (1997), vol. 97. [Page 41]
- [170] Regan, K., and Cohen, R. A model of indirect reputation assessment for adaptive buying agents in electronic markets. *Proceedings of the Business Agents and Semantic Web (BASeWEB05)* (2005). [Page 41]
- [171] Rennard, J.-P., and Mange, D. *Vie artificielle : où la biologie rencontre l'informatique : illustré avec Java*. Vuibert, 2002. [Page 31]
- [172] Reynolds, C. W. Flocks, herds and schools : A distributed behavioral model. In *ACM SIGGRAPH computer graphics* (1987), vol. 21, ACM, pp. 25–34. [Page 31]
- [173] Richards, P. I. Shock waves on the highway. *Operations research* 4, 1 (1956), 42–51. [Page 23]
- [174] Russell, S., Norvig, P., and Intelligence, A. *Artificial Intelligence : A modern approach*, vol. 25. Citeseer, 1995. [Pages 31, 33]
- [175] Sabater, J., Paolucci, M., and Conte, R. Repage : Reputation and image among limited autonomous partners. *Journal of artificial societies and social simulation* 9, 2 (2006). [Page 41]
- [176] Sabater, J., and Sierra, C. Regret : A reputation model for gregarious societies. In *Fourth workshop on deception fraud and trust in agent societies* (2001), vol. 70. [Page 41]
- [177] Sabater, J., and Sierra, C. Review on computational trust and reputation models. *Artificial intelligence review* 24, 1 (2005), 33–60. [Pages 38, 39]
- [178] Schakel, W. J., van Arem, B., and Netten, B. D. Effects of cooperative adaptive cruise control on traffic flow stability. In *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on* (2010), IEEE, pp. 759–764. [Page 30]
- [179] Schillo, M., Funk, P., and Rovatsos, M. Using trust for detecting deceitful agents in artificial societies. *Applied Artificial Intelligence* 14, 8 (2000), 825–848. [Pages 41, 48, 78]

- [180] Schönhof, M., Treiber, M., Kesting, A., and Helbing, D. Autonomous detection and anticipation of jam fronts from messages propagated by intervehicle communication. *Transportation Research Record : Journal of the Transportation Research Board* 1999, 1 (2007), 3–12. [Pages 53, 89]
- [181] Sen, S., and Sajja, N. Robustness of reputation-based trust : Boolean case. In *Proceedings of the 1st International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS '02)* (2002), ACM Press, pp. 288–293. [Page 41]
- [182] Servat, D., Perrier, E., Treuil, J.-P., and Drogoul, A. When agents emerge from agents : Introducing multi-scale viewpoints in multi-agent simulations. In *Multi-Agent Systems and Agent-Based Simulation*, J. Sichman, R. Conte, and N. Gilbert, Eds., vol. 1534 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 1998, pp. 183–198. [Page 35]
- [183] Sierra, C., and Debenham, J. An information-based model for trust. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems* (2005), pp. 497–504. [Page 41]
- [184] Sommer, C., German, R., and Dressler, F. Bidirectionally Coupled Network and Road Traffic Simulation for Improved IVC Analysis. *IEEE Transactions on Mobile Computing* 10, 1 (January 2011), 3–15. [Page 53]
- [185] Srivastava, A., and Geroliminis, N. Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model. *Transportation Research Part C : Emerging Technologies* 30 (2013), 161–177. [Page 43]
- [186] Talebpour, A., and Mahmassani, H. Influence of connected and autonomous vehicles on traffic flow stability and throughput. *Transportation Research Part C : Emerging Technologies* 71 (2016), 143–163. [Page 74]
- [187] Talebpour, A., Mahmassani, H., and Hamdar, S. Speed harmonization : evaluation of effectiveness under congested conditions. *Transportation Research Record : Journal of the Transportation Research Board*, 2391 (2013), 69–79. [Pages 44, 89]
- [188] Talebpour, A., Mahmassani, H. S., and Hamdar, S. H. Modeling lane-changing behavior in a connected environment : A game theory approach. *Transportation Research Part C : Emerging Technologies* 59 (2015), 216–232. [Page 27]
- [189] Thiemann, C., Treiber, M., and Kesting, A. Longitudinal hopping in intervehicle communication : Theory and simulations on modeled and empirical trajectory data. *Phys. Rev. E* 78 (Sep 2008), 036102. [Page 53]
- [190] Toledo, T., Koutsopoulos, H. N., and Ben-Akiva, M. E. Modeling integrated lane-changing behavior. *Transportation Research Record : Journal of the Transportation Research Board* 1857 (2003), 30–38. [Page 27]
- [191] TRB. 75 years of the fundamental diagram for traffic flow theory : Greenshields symposium. transportation research circular, number e-c149. Tech. rep., Transportation Research Board of the National Academies, 2011. [Page 20]
- [192] Treiber, M., Hennecke, A., and Helbing, D. Congested traffic states in empirical observations and microscopic simulations. *Phys. Rev. E* 62 (2000), 1805–1824. [Pages 26, 27, 44, 45, 55, 64, 131]
- [193] Treiber, M., and Kesting, A. An open-source microscopic traffic simulator. *IEEE Intelligent Transportation Systems Magazine* 2, 3 (2010), 6–13. [Pages 52, 53, 54]

- [194] Treiber, M., and Kesting, A. *Traffic Flow Dynamic : Data, Models and Simulation*. 2013. [Pages 25, 28, 52, 55, 75]
- [195] Treiber, M., Kesting, A., and Helbing, D. Three-phase traffic theory and two-phase models with a fundamental diagram in the light of empirical stylized facts. *Transportation Research Part B : Methodological* 44, 8 (2010), 983–1000. [Page 22]
- [196] Treiber, M., Kesting, A., and Thiemann, C. How much does traffic congestion increase fuel consumption and emissions ? applying a fuel consumption model to the ngsim trajectory data. In *87th Annual Meeting of the Transportation Research Board, Washington, DC* (2008). [Page 71]
- [197] Uny Cao, Y., Fukunaga, A. S., and Kahng, A. B. Cooperative mobile robotics : Antecedents and directions. *Autonomous Robots* 4, 1 (1997), 7–27. [Pages 31, 36]
- [198] Vahidi, A., and Eskandarian, A. Research advances in intelligent collision avoidance and adaptive cruise control. *IEEE transactions on intelligent transportation systems* 4, 3 (2003), 143–153. [Page 5]
- [199] Van Arem, B., van Driel, C. J., and Visser, R. The impact of cooperative adaptive cruise control on traffic-flow characteristics. *Intelligent Transportation Systems, IEEE Transactions on* 7, 4 (2006), 429–436. [Pages 10, 30, 75]
- [200] Van Katwijk, R., and Van Koningsbruggen, P. Coordination of traffic management instruments using agent technology. *Transportation Research Part C : Emerging Technologies* 10 (2002), 455 – 471. [Page 37]
- [201] van Wageningen-Kessels, F. *Multi-class continuum traffic flow models : Analysis and simulation methods*. PhD thesis, TU Delft, Delft University of Technology, 2013. [Pages 21, 159]
- [202] Van Willigen, W., Haasdijk, E., and Kester, L. A multi-objective approach to evolving platooning strategies in intelligent transportation systems. In *Proceeding of the fifteenth annual conference on Genetic and evolutionary computation conference* (2013), ACM, pp. 1397–1404. [Page 53]
- [203] Vasirani, M., and Ossowski, S. A market-inspired approach to reservation-based urban road traffic management. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1* (2009), International Foundation for Autonomous Agents and Multiagent Systems, pp. 617–624. [Page 89]
- [204] Vieira, D. C., Adeodato, P. J., and Júnior, P. M. G. A temporal difference gng-based algorithm that can learn to control in reinforcement learning environments. In *Machine Learning and Applications (ICMLA), 2013 12th International Conference on* (2013), vol. 1, IEEE, pp. 329–332. [Pages 101, 127]
- [205] Wang, M., Daamen, W., Hoogendoorn, S. P., and van Arem, B. Rolling horizon control framework for driver assistance systems. part i : Mathematical formulation and non-cooperative systems. *Transportation research part C : emerging technologies* 40 (2014), 271–289. [Page 31]
- [206] Wang, M., Daamen, W., Hoogendoorn, S. P., and van Arem, B. Rolling horizon control framework for driver assistance systems. part ii : Cooperative sensing and cooperative control. *Transportation research part C : emerging technologies* 40 (2014), 290–311. [Page 31]
- [207] Wei, G. Learning to coordinate actions in multi-agent systems. In *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence (IJCAI-93)* (1993), pp. 311–316. [Page 97]
- [208] Weiss, G. *Multiagent systems : a modern approach to distributed artificial intelligence*. MIT press, 1999, ch. Learning in Multiagent Systems, pp. 259–298. [Page 95]

- [209] Weyns, D., Parunak, H. V. D., Michel, F., Holvoet, T., and Ferber, J. Environments for multiagent systems state-of-the-art and research challenges. In *Environments for multi-agent systems*. Springer, 2004, pp. 1–47. [Page 33]
- [210] Wilmink, I., Klunder, G., and Van Arem, B. Traffic flow effects of integrated full-range speed assistance (irsa). In *2007 IEEE Intelligent Vehicles Symposium (2007)*, IEEE, pp. 1204–1210. [Page 30]
- [211] Winner, H., Witte, S., Uhler, W., and Lichtenberg, B. Adaptive cruise control system aspects and development trends. Tech. rep., SAE Technical Paper, 1996. [Page 28]
- [212] Wolf, T. D., and Holvoet, T. Emergence versus self-organisation : Different concepts but promising when combined. Springer-Verlag, pp. 1–15. [Page 35]
- [213] Wooldridge, M., Jennings, N. R., et al. Intelligent agents : Theory and practice. *Knowledge engineering review* 10, 2 (1995), 115–152. [Page 32]
- [214] Xu, Z., Jin, P. J., Jiang, X., and Ran, B. Macroscopic fundamental diagram (mfd) based on vehicle-to-infrastructure (v2i) connected vehicle data. In *Transportation Research Board 94th Annual Meeting (2015)*, no. 15-0749. [Page 24]
- [215] Yan, X., Gu, F., Hu, X., and Engstrom, C. Dynamic data driven event reconstruction for traffic simulation using sequential monte carlo methods. In *Simulation Conference (WSC), 2013 Winter (Dec 2013)*, pp. 2042–2053. [Page 53]
- [216] Yu, B., and Singh, M. P. Detecting deception in reputation management. In *AAMAS '03 : proceedings of the second international joint conference on autonomous agents and multiagent systems (2003)*, pp. 73–80. [Page 41]
- [217] Zacharia, G. Collaborative reputation mechanisms for online communities. Master's thesis, Massachusetts Institute of Technology, 1999. [Page 41]
- [218] Zacharia, G., and Maes, P. Trust management through reputation mechanisms. *Applied Artificial Intelligence* 14, 9 (2000), 881–907. [Page 41]
- [219] Zheng, Z., Ahn, S., Chen, D., and Laval, J. The effects of lane-changing on the immediate follower : Anticipation, relaxation, and change in driver characteristics. *Transportation research part C : emerging technologies* 26 (2013), 367–379. [Page 130]
- [220] Zhu, F., and Ukkusuri, S. V. Accounting for dynamic speed limit control in a stochastic traffic environment : A reinforcement learning approach. *Transportation Research Part C : Emerging Technologies* 41 (2014), 30 – 47. [Page 89]
- [221] Zlatev, J., and Balkenius, C. Introduction : Why "epigenetic robotics" ? In *Proceedings of the First Conference on Epigenetic Robotics*. In : Balkenius, C., Zlatev, J., Kozima, H., Dautenhahn, K., Breazeal, C.(eds.) : *Lund University Cognitive Science Series* (2001), no. 85. [Page 91]

Articles de l'auteur et valorisation

Revue Internationale

Guériau M., Billot R., El Faouzi N.-E., Monteil J., Armetta F., Hassas S.
How to assess the benefits of connected vehicles? A simulation framework for the design of cooperative traffic management strategies.
Transportation Research Part C: Emerging Technologies 67 (2016), pages 266 – 279.

Conférences Internationales

Guériau M., Hassas S., Armetta F., Billot R., El Faouzi N.-E.
A constructivist approach for a self-adaptive decision-making system: application to road traffic control.
28th IEEE International Conference on Tools with Artificial Intelligence (ICTAI) (2016), IEEE, pages 670 – 677.

Billot R., El Faouzi N.-E., Guériau M., Monteil J.
Can C-ITS lead to the emergence of traffic management 2.0?
17th International IEEE Conference on Intelligent Transportation Systems (ITSC) (2014), IEEE, pages 483 – 488.

Conférences Nationales

Guériau M., Billot R., Armetta F., Hassas S., El Faouzi N.-E.
Un simulateur multiagent de trafic coopératif.
23èmes Journées Francophones sur les Systèmes Multi-Agents (JFSMA'15) (2015), Cépaduès, pages 165 – 174.

Présentations et démonstrations

Articles courts

Guériau M., Billot R., El Faouzi N.-E., Hassas S., Armetta F.
Multi-agent dynamic coupling for cooperative vehicles modeling.
Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (DEMO Track) (2015),
AAAI Press, pages 4276 – 4277.

Guériau M., Billot R., Hassas S., Armetta F., El Faouzi N.-E.
An extension of movsim for multi-agent cooperative vehicles modeling.
2014 International Conference on Connected Vehicles and Expo (ICCVE) (2014), IEEE,
pages 859 – 860.

Présentations en conférence

Guériau M., Billot R., El Faouzi N.-E., Hassas S., Armetta F.
X2V-based information dissemination for highway congestion reduction.
18th Euro Working Group on Transportation (EWGT) (2015).

Guériau M., Billot R., Monteil J., El Faouzi N.-E., Armetta F., Hassas S.
Agent-Based Cooperative Traffic Modeling and Simulation.
94th Transportation Research Board Annual Meeting (2015).

Guériau M., Billot R., El Faouzi N.-E.
Flocking behavior based microscopic cooperative traffic modeling: Simulation studies.
4th International Symposium of Transport Simulation (ISTS) (2014)

Table des figures

1.1	Les interactions d'un véhicule intelligent et son environnement.	6
1.2	Représentation simplifiée des véhicules différents types de véhicules dans le contexte des Systèmes Coopératifs.	7
1.3	Échanges d'information dans les systèmes coopératifs	9
1.4	Illustration de certains cas d'usages envisagés dans les C-ITS.	10
1.5	Le <i>C-ITS Corridor</i> européen	11
1.6	Logo du projet SCOOP@F	11
1.7	Carte des sites pilotes du projet SCOOP@F	12
1.8	Illustration de l'organisation du manuscrit, permettant de visualiser les contributions en rapport avec les différentes phases de déploiement des Systèmes Coopératifs.	17
2.1	"Arbre généalogique" des modèles de trafic, tiré des récents travaux de thèse de van Wageningen-Kessels [201].	21
2.2	Exemple d'observations spatio-temporelles de trajectoires individuelles ou "diagramme espace-temps".	22
2.3	Schéma d'un diagramme fondamental (débit/concentration)	23
2.4	Numérotation des véhicules dans les modèles microscopiques.	24
2.5	Exemple de discrétisation d'un réseau pour les automates cellulaires.	25
2.6	Interprétation graphique des paramètres utilisés dans les lois de poursuites.	26
2.7	Représentation multi-voie d'une section et modélisation du changement de voie.	27
2.8	Représentation de la perception d'un véhicule équipé du système ACC.	28
2.9	Représentation de la perception d'un véhicule équipé du système C-ACC.	29
2.10	Représentation de la perception d'un véhicule équipé d'un système de contrôle basé sur le principe de la multi-anticipation.	30
2.11	Schéma simplifié des interactions d'un agent dans son environnement et conceptualisation de la notion de représentation	32
2.12	Couplage de dynamiques pour les C-ITS	42
2.13	Illustration de la stratégie de changement de voie MOBIL	45
2.14	Illustration de l'application de la règle de proximité.	47
3.1	Capture de l'interface du simulateur MovSim	54
3.2	Diagramme de classes simplifié de MASCAT, extension de Movsim	56
3.3	Schéma d'opérations de la simulation dans MASCAT	56
3.4	Représentation de la perception directe d'un véhicule équipé d'un capteur (télémètre laser).	57

3.5	Représentation des boucles électromagnétiques (rectangles gris) et d'une UBR (cercle blanc) dans l'interface de MASCAT.	58
3.6	Schéma des communications (V2V et I2V/V2I) possibles dans MASCAT.	59
3.7	Principe de rétro-propagation (<i>re-broadcasting</i>) des messages en V2V.	60
3.8	Passage du référentiel global au référentiel local du véhicule.	62
3.9	Illustration des paramètres des scénarios dans MovSim et son extension MASCAT. . .	65
3.10	Capture d'écran de la vue trafic adaptée pour le trafic coopératif dans MASCAT. . . .	66
3.11	Exemple d'un tracé des trajectoires sous la forme d'un diagramme espace-temps- vitesses dans MASCAT.	67
3.12	Capture d'écran du simulateur avec le scénario de l'autoroute péri-urbaine à 3 voies .	72
3.13	Résultats du modèle bilatéral multi-anticipatif sur les trajectoires des véhicules pour différents taux de pénétration.	74
3.14	Comparaison de la valeur d'hétérogénéité du groupe entre les modèles BMA et ACC, pour différents taux de pénétration.	75
3.15	Résultats du modèle de confiance sur les trajectoires d'un flux de trafic en présence de véhicules connectés équipés de capteurs défaillants	76
3.16	Comparaison des valeurs de G_d pour un flux de trafic en présence de véhicules connectés équipés de capteurs défaillants, avec et sans le modèle de confiance . .	77
3.17	Distributions des PET pour trois variantes de modèles de confiance.	79
3.18	Capture d'écran du simulateur pour le scénario d'insertion.	82
3.19	Mesures des débits et concentrations dans le scénario d'insertion.	82
3.20	Capture d'écran du simulateur pour le scénario d'insertion.	83
3.21	Diagrammes débits-concentration du scénario de fermeture de voie.	84
4.1	Illustration d'un scénario de contrôle décentralisé des systèmes coopératifs	90
4.2	Schéma des processus internes d'un agent en interaction avec son environnement capable de construire sa représentation en suivant un apprentissage constructiviste. .	94
4.3	Schéma simplifié des interactions d'un agent faisant évoluer sa représentation sous la forme d'un espace d'états dans un contexte d'apprentissage par renforcement.	96
4.4	Aperçu général du modèle d'élaboration de représentation par apprentissage.	105
4.5	Processus de perception visant à proposer une discrétisation de l'environnement à partir de représentations concurrentes.	108
4.6	Processus de sélection de l'action et renforcement des représentations concurrentes. .	109
4.7	Illustration simplifiée des conflits entre les états sous la forme d'une matrice de conflits. .	114
4.8	Exemple d'environnement virtuel généré dans le prototype expérimental	118
4.9	Capture d'écran du prototype expérimental montrant le résultat d'une simulation. . . .	119
4.10	Capture d'écran du prototype montrant les actions résultant de la politique du système comparées aux actions attendues.	119
4.11	Aperçu de l'environnement et de la discrétisation d'un agent discrétiseur.	120
4.12	Comparaison de la performance de 7 implémentations du modèle utilisant jusqu'à 3 agents discrétiseurs	121
4.13	Comparaison des temps de convergence de l'apprentissage pour deux implémentations du modèle.	123
4.14	Illustration des environnements générés (deux contextes). Chaque couleur représente une action attendue différente.	123
4.15	Comparaison de la performance de trois implémentations face à des changements de contexte.	124

4.16	Performance moyenne des trois implémentations face à des interpolations entre les contextes A et B.	126
4.17	Pourcentage de participation des agents spécialisés dans les contextes A et B face à des contextes interpolés.	126
4.18	Comparaison de trois implémentations du modèle impliquant ou non un agent associateur.	128
4.19	Capture d'écran du scénario de contrôle modélisé dans le simulateur.	130
4.20	Discretisations générées par les deux agents discrétiseurs spécialisés dans les capteurs de l'UBR	133
4.21	Résultats de trois simulations de contrôle I2V d'un flux de véhicules partiellement connecté.	135
4.22	Actions des agents internes au système dans l'implémentation du modèle	136

Liste des tableaux

2.1	Comparaison des modèles basés sur la confiance et la réputation parmi les plus utilisés de la littérature.	41
2.2	Résumé des caractéristiques des véhicules coopératifs et non coopératifs.	44

Liste des sigles

Sigle	Signification	Traduction (le cas échéant)
ABS	Antiblockiersystem	Système anti-blocage des roues
ACC	Adaptive Cruise Control	Régulateur de vitesse à contrôle de distance
AICC	Autonomous Intelligent Cruise Control	
ADAS	Advanced Driver Assistance Systems	Système avancé d'assistance au conducteur
BMA	Bilateral Multi-Anticipative model	Modèle Multi-Anticipatif Bilatéral
C-ACC	Cooperative Adaptive Cruise Control	
C-ITS	Cooperative Intelligent Transportation Systems	Systèmes de Transports Intelligents Coopératifs
CAM	Cooperative Awareness Message	
CSV	Comma-Separated Values	
DENM	Decentralized Environmental Notification Message	
DIR	Direction Inter-départemental des Routes	
ESP	Electronic Stability Program	Correcteur électronique de trajectoire
FCD	Floating Car Data	Données de véhicule traceur
GNSS	Global Navigation Satellite System	
GPS	Global Positioning System	
GDM	Global System for Mobile communications	
GNG	Growing Neural Gas	
I2V	Infrastructure to Vehicle	
IA	Intelligence Artificielle	Artificial Intelligence
IAD	Intelligence Artificielle Distribuée	Distributed Artificial Intelligence
IDM	Intelligent Driver Model	
ITS	Intelligent Transportation Systems	Systèmes de Transports Intelligents
IVC	Inter-Vehicular Communication	Communication Inter-Véhiculaire
LDWS	Lane Departure Warning System	Système anti-franchissement de ligne continue
LIDAR	Light Detection And Ranging	Télémètre Laser
LKS	Lane Keeping System	Système de maintien de trajectoire
MPC	Model Predictive Control	
NHTSA	National Highway Transportation Safety Administration	
PET	Post-Encroachment Time	
PL	Poids Lourd	
PMV	Panneau à Message Variable	
SMA	Système Multi-Agent	Multi-Agent System
SOM	Self-Organizing Maps	Cartes auto-organisées
TIC	Technologies de l'Information et de la Communication	
TMC	Traffic Management Center	Centre de gestion du trafic
TTC	Time To Collision	
TTS	Total Time Spent	
UA	Unité Application	
UBR	Unité de Bord de Route	Road Side Unit
UCB	Upper Confidence Bound	
UEV	Unité Embarquée Véhicule	

Sigle	Signification	Traduction (le cas échéant)
V2I	Vehicle to Infrastructure	
V2V	Vehicle to Vehicle	
VANET	Vehicular Ad-hoc NETwork	Réseaux Ad-Hoc Véhiculaires
VCA	Véhicule connecté et autonomes	
VP	Véhicule Particulier	
VSL	Variable Speed Limit	Limite dynamique de vitesse
WAVE	Wireless Access in Vehicular Environments	
XML	Extensible Markup Language	Langage à balises