

Une approche orientée risques pour la prise de décision dans les trains autonomes : Cas de la fonction anti-collision

CHELOUATI Mohammed
Univ Gustave Eiffel
Villeneuve d'Ascq
mohammed.chelouati2@univ-
eiffel.com

BOUSSIF Abderraouf
Univ Gustave Eiffel
Villeneuve d'Ascq
abderraouf.boussif@univ-
eiffel.com

BEUGIN Julie
Univ Gustave Eiffel
Villeneuve d'Ascq
julie.beugin@univ-eiffel.com

EL KOURSI El-Miloudi
Univ Gustave Eiffel
Villeneuve d'Ascq
el-miloudi.el-kourisi@univ-
eiffel.fr

Résumé — En vue de l'intégration de systèmes autonomes dans le secteur ferroviaire, garantir un niveau de sécurité équivalent à celui des systèmes classiques pilotés par des opérateurs humains constitue un enjeu critique. L'exploitation de trains entièrement autonomes implique le déploiement de solutions avancées capables, sans perception humaine du contexte opérationnel, d'analyser et évaluer les risques en temps réel et de prendre des décisions complexes. Cet article présente une approche innovante pour la prise de décision orientée risques dans les trains autonomes, en se basant sur la mise en œuvre de Processus Décisionnels de Markov Partiellement Observables (POMDPs). La méthodologie développée permet une évaluation et un suivi continus des risques de collision avec des obstacles de l'environnement, en veillant à ce que le système de conduite autonome des trains maintienne systématiquement un niveau de risque acceptable, caractérisé par une probabilité d'occurrence et un degré de sévérité spécifiques. Cette approche renforce les capacités décisionnelles du système autonome en favorisant l'adoption de décisions sécuritaires face aux incertitudes liées à l'état opérationnel du train et aux variables environnementales. La validité et l'efficacité de cette approche sont démontrées à travers son intégration dans un dispositif d'anticollision destiné aux trains autonomes.

Mots-clés — *Train autonome, appréciation dynamique des risques, processus de décision Markovien, assurance sécurité*

Abstract — In view of the integration of autonomous systems into the railway sector, ensuring a level of safety equivalent to that of conventional systems operated by human operators is a critical challenge. The operation of fully autonomous trains involves the deployment of advanced solutions capable of analyzing and evaluating real-time risk and making complex decisions. This paper presents an innovative approach to risk-oriented decision-making in autonomous trains, based on the implementation of Partially Observable Markov Decision Processes (POMDPs). The developed methodology enables continuous assessment and monitoring of collision risks with the environmental obstacles, ensuring that the autonomous train driving system consistently maintains an acceptable level of risk, characterized by specific probabilities of occurrence and degrees of severity. This approach enhances the decision-making capabilities of the autonomous system by facilitating the adoption of safe decisions in the face of uncertainties related to the operational state of the train and environmental variables. The validity and effectiveness of this approach are demonstrated through its integration into an anti-collision device designed for autonomous trains.

Key words — *Autonomous train, dynamic risk assessment, Markov decision process, safety assurance*

I. INTRODUCTION

L'émergence de trains autonomes promet des transformations substantielles au sein de l'industrie du transport ferroviaire (Trentesaux et al. 2018). En effet, leur potentiel disruptif quant à l'optimisation de la capacité des lignes et d'amélioration des services ferroviaires est prometteur. Cette évolution, loin d'être hypothétique, est actuellement au cœur des efforts de recherche et développement à l'échelle internationale. Par exemple, en 2016 en France, la SNCF a inauguré Tech4Rail, un programme technologique visant à poser les jalons des futurs systèmes ferroviaires, incluant l'introduction de trains autonomes et semi-autonomes. En collaboration avec le programme Train Autonome de l'Institut de Recherche Technologique Railenium, ainsi que divers partenaires industriels et académiques, cette initiative a conduit à la formation de deux consortiums dédiés à la réalisation de trois projets d'envergure (Lagay et Adell 2018; Trentesaux et al. 2018) : (1) Train Autonome-Fret, (2) Train Autonome-Service Voyageurs, et (3) TC-Rail pour la mise en œuvre d'un train télécommandé. Ces projets, au cœur de nombreux défis d'ingénierie et de recherche, se concentrent principalement sur l'intégration de l'intelligence artificielle pour améliorer les fonctions de perception (Boudardara et al. 2023), de contrôle, et de prise de décision (Chelouati et al. 2023b). Ils s'intéressent également à des aspects tels que la coopération homme-machine (Alsaba et al. 2020), l'acceptation sociale des technologies autonomes, et les stratégies d'appréciation des risques et d'assurance sécurité.

Pour qu'un nouveau système, en particulier le train autonome, puisse devenir un élément constitutif du système ferroviaire global de l'Union Européenne, il est essentiel que le niveau de sécurité du système global soit préservé, voire amélioré, conformément à la réglementation nationale et européenne en matière de sécurité ferroviaire. Dans le cadre des trains conventionnels, les processus d'analyse, d'évaluation et de maîtrise des risques sont intégrés dès les phases initiales de conception et de développement (Parhizkar 2021). De plus, ces processus reposent sur la présence d'un conducteur humain pour réaliser une analyse et une appréciation dynamique des risques lors de l'exploitation, tâches essentielles au maintien de la sécurité du train et de ses passagers face à d'éventuels obstacles (Gebauer, Pree, et Stadlmann 2012). Pour les trains autonomes, le défi réside dans la capacité du système de conduite autonome (ADS, *Autonomous Driving System*) à effectuer ces tâches d'appréciation des risques. Cela nécessite une analyse en temps réel pour assurer un fonctionnement en toute sécurité dans diverses conditions opérationnelles. L'appréciation des risques en temps réel (DRA, *Dynamic Risk Assessment*) implique l'intégration de fonctions avancées telles que la détection d'obstacles et la fonction d'anticollision, permettant ainsi des opérations autonomes sûres et efficaces (Rosique et al. 2019).

Ce travail est organisé comme suit : la section II présente les travaux issus de la littérature scientifique concernant la prise en compte des incertitudes dans les processus de prise de décision des systèmes autonomes. Elle discute également des avantages de l'intégration des POMDPs dans ces processus pour la maîtrise des risques. La section III détaille la problématique liée à la fonction d'anticollision pour le train autonome et la manière de structurer les données de risque associées, nécessaires pour pouvoir réaliser la tâche de DRA. La méthodologie proposée pour mettre en œuvre la fonction d'anticollision est décrite dans la section IV, incluant la définition du modèle POMDP et la proposition d'un modèle de risque associé. Les résultats de simulation sont présentés dans la section V. Enfin, la section VI conclut et esquisse des perspectives pour les recherches futures.

II. VERS L'UTILISATION DES POMDPs DANS LES SYSTEMES DE CONDUITE AUTONOME (ADS)

Cette section débute par une présentation des objectifs et des fonctions de l'ADS, en mettant spécifiquement l'accent sur ses aspects décisionnels. Elle passe ensuite en revue les travaux de recherche traitant des incertitudes dans les processus de prise de décision de différents systèmes de transport terrestre, mettant en lumière diverses techniques employées pour atténuer ces incertitudes. Une attention particulière est accordée aux Processus Décisionnels de Markov Partiellement Observables (POMDP) en raison de leurs caractéristiques appropriées à l'appréciation dynamique des risques et leur adaptabilité à la variabilité des contextes opérationnels ferroviaires rencontrés. Les principes de mise en œuvre de ses processus seront vus plus en détails dans la Section IV.

A. Objectifs et fonctions de l'ADS dans les trains autonomes

L'ADS d'un train autonome exécute plusieurs fonctions afin de garantir une exploitation sûre dans un environnement ouvert. Ces fonctions englobent la gestion de l'état du train, la détection d'obstacles pour éviter les collisions, ainsi que l'exécution d'autres tâches critiques essentielles regroupées sous le concept de « Tâche de Conduite Dynamique » (DDT, *Dynamic Driving Task*). Ce concept, tel que défini dans (Mendiboure et al. 2023), inclut les tâches traditionnelles de conduite de train telles que la perception de l'environnement, la conscience situationnelle, l'appréciation dynamique des risques, la prise de décision et les tâches de contrôle. Cette section met particulièrement l'accent sur la tâche de prise de décision réalisée par l'ADS, sous-système au sein des trains autonomes.

La Figure 1 illustre les composants critiques de l'ADS, où l'unité de prise de décision traite toutes les informations nécessaires provenant de l'unité de perception, calcule les indicateurs clés opérationnels et de sécurité, et formule des actions appropriées. Un pilier du processus de prise de décision est l'appréciation dynamique des risques (DRA), qui constitue le fondement du contrôle de la sécurité à travers un modèle de risque. Selon le niveau de risque évalué, l'ADS élabore un plan d'action, qui peut inclure des actions telles que l'accélération, le maintien de la vitesse, la décélération en cas de détection d'un obstacle potentiel jusqu'à garantir une distance minimum, ou l'initiation d'une procédure de freinage d'urgence en cas de risque de collision immédiat.

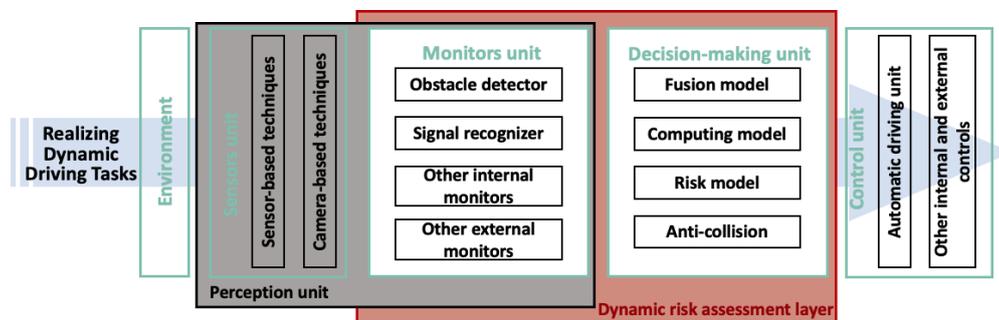


Figure 1 : architecture simplifiée de l'ADS

B. Gestion des incertitudes dans les processus de prise de décision

La gestion des incertitudes dans la prise de décision des systèmes autonomes est devenue un point central de recherche, puisqu'elles amènent des difficultés dans les processus de fusion de données issues de capteurs, de perception sous des conditions environnementales variables et d'estimation de l'état du système. La littérature révèle que les conditions environnementales influencent considérablement la précision de la perception, où des facteurs tels que l'éclairage, la météo et les obstructions peuvent

introduire des incertitudes dans la détection et la classification des objets. De plus, le maintien d'un état précis du système, malgré la variabilité des conditions opérationnelles et la nécessité de réactions en temps réel, est impératif car il constitue la base de toutes les décisions subséquentes. Cela souligne le besoin de cadres et méthodologies robustes malgré les incertitudes.

En fait, plusieurs travaux de recherche, tels que ceux fondés sur l'utilisation de Réseaux Bayésiens (BN) (Heckerman 2008) et de Réseaux Bayésiens Dynamiques (DBN) (Ghahramani 1998) offrent des modèles pour comprendre les relations probabilistes et gérer les dépendances temporelles entre les variables. Les arbres de décision fournissent une approche simple et intuitive pour modéliser les décisions et leurs conséquences potentielles, y compris les résultats, les coûts des ressources et l'utilité. Enfin, l'Apprentissage par Renforcement (RL) (Kaelbling, Littman, et Moore 1996) propose une approche interactive pour apprendre une politique optimale par interaction directe avec un environnement dynamique. Parmi ces méthodologies, les POMDPs se distinguent par leur applicabilité aux systèmes autonomes, comme détaillé dans les sections suivantes (Monahan 1982).

C. Avantages des POMDPs dans les processus de prise de décision

Les POMDPs offrent plusieurs avantages dans le cadre de la prise de décision. Premièrement, les POMDP prennent explicitement en compte l'incertitude à la fois dans l'état du système et dans les observations (Tran et Bae 2021). Cette caractéristique est essentielle dans les systèmes autonomes où les lectures des capteurs peuvent ne pas toujours être fiables ou complètes, et où l'état réel de l'environnement est difficile à spécifier et à prédire. Deuxièmement, contrairement à des méthodologies telles que les arbres de décision qui opèrent sur des modèles discrets, les POMDP peuvent gérer des états, des actions et des espaces d'observation continus (Temizer et al. s. d.). Ceci est particulièrement utile dans le cas où l'environnement est souvent mieux représenté comme un espace continu, tel que les positions relatives et les vitesses des véhicules (Leurent 2018). Enfin, bien que l'apprentissage par renforcement soit également un outil puissant pour la prise de décision sous incertitude, il nécessite généralement un grand nombre d'essais pour apprendre la politique optimale de fonctionnement, ce qui peut ne pas toujours être faisable ou sûr pour les systèmes critiques comme les trains autonomes. D'autre part, les POMDP s'appuient sur un modèle qui permet un calcul efficace de cette politique. Outre leur capacité à traiter l'incertitude, les POMDP peuvent également modéliser à la fois la stochasticité dans les événements liés à l'environnement et les informations sensorielles imparfaites.

Nous mettrons maintenant l'accent sur le rôle central que joue la prise de décision pour assurer la sécurité des trains autonomes en nous concentrant sur l'une de ses principales fonctions : la fonction d'anticollision.

III. PROBLEMATIQUE DE LA FONCTION D'ANTICOLLISION

Cette section décrit la problématique liée à la fonction d'anticollision pour le train autonome et la manière de structurer les données de risque associées, nécessaires à la réalisation de l'appréciation dynamique des risques (DRA). Pour évaluer le risque, il est nécessaire au préalable de disposer d'informations sur le profil de risque. Pour ce faire, une structuration de ces informations qualifiée de *DRA framework* a été proposée par les auteurs dans (Chelouati et al. 2022). Cette structuration est synthétisée dans cette section, ainsi que son application à la fonction d'anticollision. Elle sera utilisée dans la méthodologie qui suit.

A. Objectifs de la fonction d'anticollision

La fonction d'anticollision représente la capacité du train à détecter et à réagir de manière appropriée et sûre à tout obstacle potentiel pouvant provoquer une collision. Il est essentiel qu'un train autonome soit équipé des capteurs et algorithmes capables d'identifier avec précision la nature d'un obstacle, estimer sa distance par rapport au train et sa trajectoire, afin de calculer et d'évaluer le risque associé. Pour réaliser la fonction d'anticollision, le système ADS surveille l'état opérationnel du train et son milieu environnant, recherchant constamment des obstacles potentiels.

La figure 2 illustre un scénario où un train autonome s'approche d'une zone de passage à niveau avec intersection routière et traversée possibles de personnes. Cet endroit représente une zone potentielle d'obstacles que l'ADS doit identifier. De plus, l'ADS doit être capable de réagir à tout obstacle potentiel et de prendre des décisions garantissant un niveau de sécurité acceptable.

Par ailleurs, la présence d'arbres et d'animaux le long des voies ne constituent pas simplement des éléments environnementaux dans l'image, il s'agit de risques potentiels tels que la chute de branches cassées ou l'intrusion d'animaux sur la voie. Dans de tels scénarios complexes et imprévisibles, la fonction d'anticollision doit traiter efficacement les risques des différents scénarios et leurs évolutions. L'objectif de la fonction d'anticollision n'est pas seulement d'identifier ces menaces, mais aussi d'évaluer le niveau de risque associé à chacune d'elles afin que la décision puisse être prise sur la base des informations de risque les plus récentes et les plus précises.

B. La tâche d'appréciation dynamique des risques inhérente à la fonction d'anticollision

Compte tenu des incertitudes liées aux environnements réels et aux informations fournies par les capteurs, les observations contribuent à former une estimation des incertitudes. Cette estimation représente de manière probabiliste la situation actuelle du train, résumant les états possibles du train et de son environnement. Une fois l'estimation de l'incertitude établie, l'analyse et l'évaluation du risque inhérente à la fonction d'anticollision doivent être réalisées. Cela renvoie à la tâche DRA, qui doit être effectuée par l'ADS pour évaluer et mettre à jour les risques associés à l'état actuel du train, à l'environnement et aux actions disponibles que le train pourrait entreprendre. Cette évaluation se base sur l'estimation de chaque incertitude, en considérant à la

fois la probabilité et les conséquences potentielles d'une collision. De plus, l'estimation des incertitudes joue un rôle important dans l'établissement du profil de risque, car elle fournit la base probabiliste à partir de laquelle les scénarios potentiellement dangereux et leurs risques associés sont évalués et classifiés comme étant acceptable ou non au sein du profil de risque.

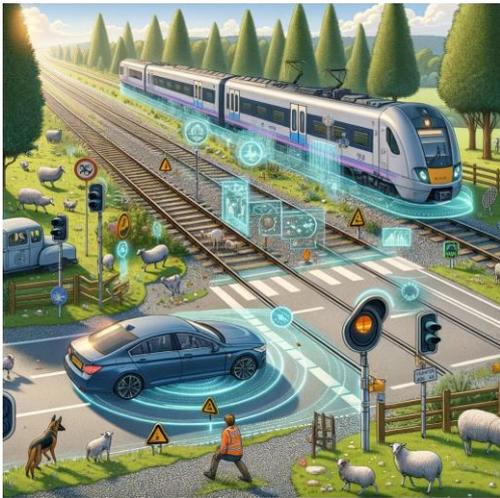


Figure 2 : Contexte de la fonction d'anticollision

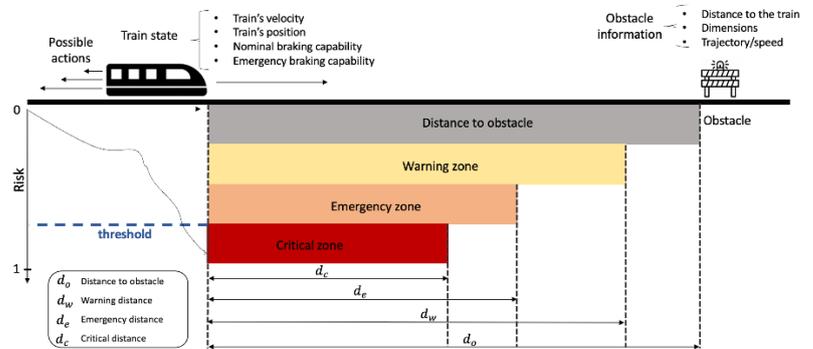


Figure 3 : Illustration de la fonction d'anticollision

Différentes actions de contrôle sont disponibles pour le train, en réponse à l'environnement et en fonction de la criticité des risques évalués, à savoir l'accélération, le maintien de la vitesse actuelle et divers types de freinage. Dans la figure 3, un obstacle se trouve à une certaine distance sur la voie du train. Par rapport à la distance du train, trois zones sont considérées : les zones d'avertissement, d'urgence et critique. La zone d'avertissement (en couleur jaune) indique une distance à partir de laquelle aucune action immédiate n'est nécessaire (c'est-à-dire que l'obstacle est encore loin ou non détecté). La zone d'urgence (en couleur orange) signifie une distance de prudence à partir de laquelle le train peut devoir ajuster sa vitesse ou freiner pour éviter une collision. Enfin, la zone critique (en couleur rouge) signifie que la présence d'un obstacle peut entraîner une collision (c'est-à-dire que dans cette zone, l'obstacle est considéré comme proche du train, et même avec un freinage d'urgence, le risque de collision est élevé). Le niveau de risque associé, représenté sur l'axe vertical avec une échelle de 0 à 1, est estimé en fonction de la distance jusqu'à l'obstacle. Évidemment, plus l'obstacle est proche du train, plus le niveau de risque est élevé. Le seuil de risque inacceptable (visualisé dans la figure par l'intersection entre la ligne pointillée bleue et l'axe vertical) est franchi lorsque le train entre dans la zone critique.

Il est à noter que la fonction d'anticollision, réalisée par la DRA, agit comme une couche d'apprentissage dynamique et de protection, garantissant un niveau de sécurité plus élevé dans l'exploitation des trains autonomes.

IV. METHODOLOGIE

Dans cette section, nous définissons les notions liées au POMDP, puis nous décrivons les différents éléments du modèle POMDP entrant dans le processus de prise de décision développé.

A. Définition des notions liées au POMDP

Un POMDP est une méthode probabiliste qui modélise le processus séquentiel d'un système fonctionnant sous incertitude. Il s'agit d'une généralisation du Processus Décisionnel de Markov pour les situations où l'état du système est partiellement inconnu. Formellement, un POMDP est un tuple $\langle S, A, O, T, Z, R, \gamma \rangle$, où S et A sont les ensembles d'états et d'actions, T est la fonction de transition qui définit la probabilité conditionnelle P de passer d'un état $s \in S$ à un autre état $s' \in S$ à la suite de l'exécution d'une action $a \in A$, c'est-à-dire, $T(s, a, s') = P(s' | a, s)$. O est l'espace d'observation qui définit les informations reçues (des capteurs) après l'exécution d'une action. Z est la fonction d'observation correspondante qui définit la probabilité conditionnelle d'observer un résultat particulier $o \in O$ après avoir exécuté une action $a \in A$ pour atteindre l'état $s' \in S$, c'est-à-dire, $Z(o, a, s') = P(o | s', a)$. R est la fonction de récompense $R(s, a)$ qui définit la récompense immédiate reçue pour être dans un état particulier $s \in S$ et prendre une action particulière $a \in A$. Enfin, $\gamma \in [0, 1]$ est le facteur de remise qui détermine la pertinence (ou non) des récompenses futures.

Dans un POMDP, seule une connaissance partielle et bruitée du système et de son environnement est considérée ; ainsi, une croyance sur les états du modèle, connue sous le nom d'état de croyance $b(s)$, est continuellement inférée. L'état de croyance est une distribution de probabilité sur l'espace d'état qui reflète le degré de certitude maintenu par le modèle POMDP sur l'état actuel du système. En conséquence, une politique $\pi : B \rightarrow A$ est utilisée comme une fonction de l'ensemble des états de croyance possibles vers l'ensemble des actions, afin de déterminer l'action adéquate à prendre.

Résoudre un POMDP implique de trouver la politique optimale π^* en termes d'action courante ou de séquence finie d'actions à exécuter afin de maximiser (ou d'optimiser) la récompense cumulative attendue au fil du temps, en tenant compte de l'état de croyance. Formellement :

$$\pi^*(b) = \underset{a \in A}{\operatorname{argmax}} \{ \sum_{s'} P(s' | b, a) [R(b, a, s') + \gamma \cdot E[V^*(b')]] \} \quad (1)$$

Pour évaluer la récompense potentielle de prendre une action a et de passer dans l'état s' , l'équation 1 considère la probabilité $P(s' | b, a)$ de passer dans l'état s' étant donné l'état de croyance b actuel et l'action a prise. Elle prend également en compte la récompense immédiate $R(b, a, s')$ obtenue de l'action a à partir de l'état de croyance b vers l'état s' . De plus, l'équation considère la valeur attendue (c'est-à-dire, la récompense attendue) de la fonction de valeur optimale, $E[V^*(b')]$ pour le prochain état de croyance b' résultant de la transition vers l'état s' . Cette composante tient compte des récompenses et résultats futurs potentiels de prendre l'action a . La politique optimale dans un POMDP peut être calculée en utilisant deux catégories principales de solveurs : les solveurs en ligne et hors ligne. Ces solveurs diffèrent dans la manière dont ils trouvent la politique optimale et les ressources informatiques qu'ils requièrent. Les solveurs en ligne sont conçus pour fonctionner en temps réel et prendre des décisions basées sur l'état actuel du système, tandis que les solveurs hors ligne sont conçus pour fonctionner hors ligne et prendre des décisions basées sur des données historiques. Le choix du solveur dépend du cas d'utilisation spécifique et des ressources informatiques disponibles.

B. Modélisation de la fonction d'anticollision du train en utilisant un modèle POMDP

Le système d'anticollision prend comme entrées des informations internes concernant l'état du train et des informations externes sur l'environnement, comme expliqué dans la sous-section III-C présentant le *DRA framework*. La figure 4 présente une vue générale de la structure entrée-sortie du POMDP utilisée pour mettre en œuvre la fonction d'anticollision.

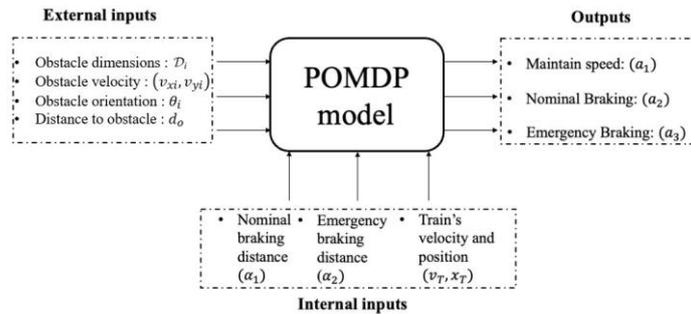


Figure 3 : Illustration du modèle POMDP

L'espace d'état continu du modèle POMDP comprend l'état du train et les états des obstacles (éventuellement) environnants. L'état du train s_T contient sa position (x_T, y_T) , sa vitesse v_T et son orientation θ_T . De même, l'état de chaque obstacle si est composé de sa position (x_i, y_i) , sa dimension D_i , sa vitesse instantanée (v_{xi}, v_{yi}) et son orientation θ_i . Il convient de noter qu'une telle formulation de l'espace d'état est réalisée sur un système de coordonnées global (ou terrestre).

Bien que la formulation continue de l'espace d'état soit une représentation fidèle du système réel, il demeure un espace continu de très haute dimension, nécessitant un temps et un espace de calcul significatifs pour résoudre le modèle et trouver la politique adéquate. De plus, les algorithmes existants pour résoudre les POMDP continus ne passent pas à l'échelle de manière efficace lorsqu'il s'agit de modèles continus de haute dimension. Afin de remédier à ce problème, nous considérons dans cet article un POMDP discret avec une représentation discrète de l'espace d'état, de l'espace d'action, et de l'espace d'observation.

La discrétisation de l'espace d'état est réalisée à l'aide d'une grille adaptative bidimensionnelle fixée à l'avant du train. Ainsi, un système de coordonnées locales (égocentriques), avec la tête du train comme origine du système, est considéré. Cela signifie qu'au lieu de représenter explicitement les positions des obstacles comme des variables continues dans les états du modèle, ils sont représentés implicitement par plusieurs variables indiquant l'occupation ou non des cellules de la grille.

L'axe positif des x est dans la direction de la conduite du train et l'axe positif des y est dirigé vers la gauche de la tête du train. Notez que la taille des cellules de la grille adaptative dépend des capacités de freinage du train, de la présence d'obstacles sur (ou à côté de) la voie, et du gabarit du train.

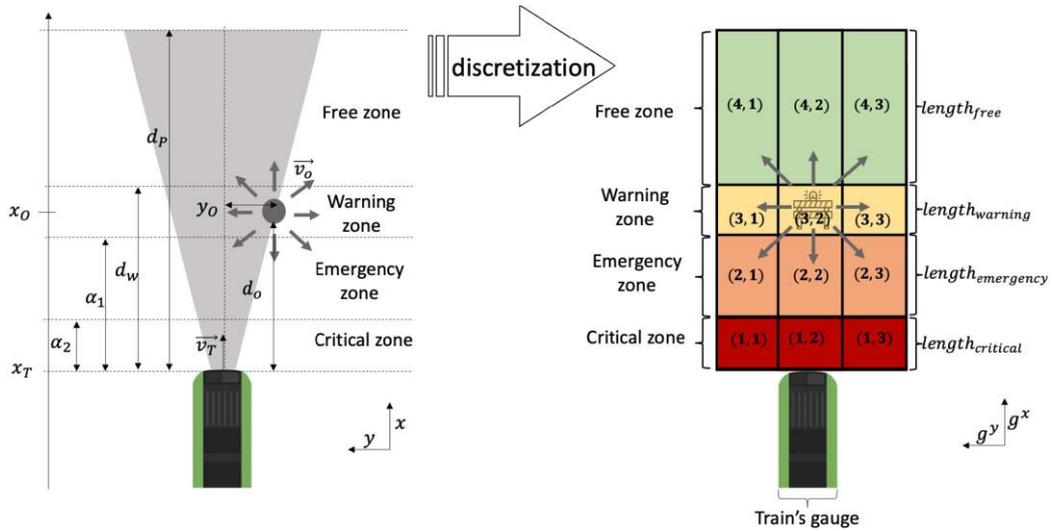


Figure 4 : la discrétisation du plan Cartésien à une grille adaptative for la navigation du train autonome

La figure 4 présente une illustration en deux parties de la grille adaptative. La première partie (à gauche de la figure) montre un train se déplaçant le long de la voie avec un obstacle apparaissant sur son chemin, visualisé à l'aide d'un système de coordonnées global. La deuxième partie de l'illustration (à droite de la figure) représente la grille adaptative résultant de ce processus de discrétisation. Ce processus est représenté par la superposition de grille sur la voie, avec les cellules numérotées entre parenthèses. Les cellules sont codées en couleur de manière cohérente avec les zones décrites à gauche : vert pour la zone libre, jaune pour la zone d'avertissement, orange pour la zone d'urgence et rouge pour la zone critique. Cette grille représente une méthode de discrétisation de l'espace continu autour du train pour que le système anticollision évalue le risque plus efficacement. Les longueurs de chaque zone dans cette grille adaptative sont indiquées sur le côté droit de la grille comme $length_{free}$, $length_{warning}$, $length_{emergency}$ et $length_{critical}$. Le gabarit du train est également noté en bas de la grille. En fait, la figure 4 illustre notre approche de quantification du risque, qui, à première vue, met l'accent sur la proximité et les distances de freinage. Cependant, l'architecture du modèle prend intrinsèquement en compte des paramètres critiques supplémentaires.

La position latérale est intégrée dans la grille discrétisée, où chaque cellule correspond à une zone latérale et longitudinale spécifique par rapport au train, nous permettant de prendre en compte le positionnement latéral des obstacles. De plus, la vitesse de l'obstacle est intégrée dans l'évaluation du risque grâce à des mises à jour dynamiques des cellules qui reflètent les positions changeantes des obstacles dans le temps. Cela permet au système d'anticiper et de réagir aux obstacles mobiles, avec une attribution de risque plus élevée pour ceux ayant une vitesse relative significative vers le train. La grille adaptative est structurée comme une grille de 12 cellules, où chaque cellule est définie en fonction de la position relative de l'obstacle (g_x, g_y) et de son orientation discrète relative θ_d . Notez que l'orientation de l'obstacle est déterminée en fonction de ses projections de vitesse (v_x, v_y) (ou de sa vitesse angulaire ω_o) et représente les transitions possibles vers les huit cellules environnantes.

Ainsi, l'ensemble d'états S peut s'exprimer comme suit : l'obstacle g_x représente la discrétisation de la position de l'obstacle sur l'axe des x et peut prendre quatre valeurs $\{1, 2, 3, 4\}$, correspondant au nombre de lignes dans la grille. L'obstacle g_y représente la discrétisation de la position de l'obstacle sur l'axe des y et peut prendre trois valeurs $\{1, 2, 3\}$, correspondant au nombre de colonnes dans la grille. De plus, la variable θ_d représente l'orientation de l'obstacle et est discrétisée d'un espace continu (de 0 à 2π) à huit valeurs discrètes $\{0, \frac{2\pi}{8}, \frac{4\pi}{8}, \frac{6\pi}{8}, \pi, \frac{10\pi}{8}, \frac{12\pi}{8}, \frac{14\pi}{8}\}$, représentant les transitions possibles vers les huit cellules environnantes. En fait, chaque combinaison unique de g_x, g_y et θ_d représente un état distinct dans la grille adaptative, indiquant la position et l'orientation de l'obstacle (voir la Figure 4).

Avec quatre valeurs possibles pour g_x , trois valeurs possibles pour g_y , et huit valeurs possibles pour θ_d , le nombre total d'états possibles dans la grille adaptative est de $N_S = 4 \times 3 \times 8 = 96$. Ces 96 états capturent toutes les configurations possibles d'un obstacle dans la grille adaptative, permettant au modèle POMDP de raisonner efficacement sur son mouvement et ses interactions potentielles avec le train dans des scénarios réels.

Pour établir la taille de chaque cellule dans la carte de grille adaptative, l'étape suivante du processus de discrétisation est la définition des différentes zones (zones Libre, d'Avertissement, d'Urgence et Critique). Les limites de chaque zone sont déterminées comme des fonctions des distances de freinage nominal et d'urgence α_1 et α_2 . En fait, la longueur des cellules dans les zones Critique, d'Urgence et d'Avertissement sont respectivement égales à la distance de freinage d'urgence ($length_{critical} = \alpha_2$), à la distance de freinage nominal ($length_{emergency} = \alpha_1 - \alpha_2$), à la distance jusqu'à l'obstacle ($length_{warning} = d_w - \alpha_1$). De plus, la longueur des cellules de la zone libre est déterminée par la distance maximale de perception (ou la portée de perception) d_p du train autonome ($length_{(free)} = d_p - d_w$). D'autre part, la largeur de toutes les cellules dans la grille adaptative est égale au gabarit du train. L'équation 2 montre les frontières de chaque zone :

$$\begin{cases} \text{Freezone} = \{(g_x, g_y, \theta_d) \in S \mid d_w < g_x \leq d_p \\ \text{Warningzone} = \{(g_x, g_y, \theta_d) \in S \mid \alpha_1 < g_x \leq d_w \\ \text{Emergencyzone} = \{(g_x, g_y, \theta_d) \in S \mid \alpha_2 < g_x \leq \alpha_1 \\ \text{Criticalzone} = \{(g_x, g_y, \theta_d) \in S \mid g_x \leq \alpha_2 \end{cases} \quad (2)$$

Dans la grille adaptative, chaque zone se compose de trois cellules, ce qui donne un total de 12 cellules.

Du point de vue de la sécurité, si un obstacle se trouve dans l'une des trois cellules de chaque zone, quelle que soit la vitesse de l'obstacle par rapport à la vitesse du train, et sachant que son orientation est vers une direction latérale (c'est-à-dire, $\theta_d = 0$ ou π , ce qui signifie que l'état suivant de l'obstacle restera dans la même zone), le niveau de risques associé peut être considéré comme similaire.

Si l'orientation de l'obstacle est vers l'avant ($\theta_d = \frac{2\pi}{8}, \frac{4\pi}{8}, \frac{6\pi}{8}$) ou vers l'arrière ($\theta_d = \frac{10\pi}{8}, \frac{12\pi}{8}, \frac{14\pi}{8}$), le risque diminuera (seulement si $v_o \geq v_T$) ou augmentera (seulement si $v_o > 0$) respectivement. Afin de définir des états POMDP avec des niveaux de risque compréhensibles, nous adoptons les hypothèses suivantes.

C. Hypothèses pour la définition des niveaux de risque

Il peut être observé que la plupart des 96 états de la carte de grille adaptative peuvent présenter des conditions de sécurité similaires. En particulier, les trois cellules de chaque zone peuvent être reliées à un niveau de risque similaire. En d'autres termes, plusieurs états pourraient présenter un niveau de risque analogue pour les opérations de train autonome. De telles similarités entre divers états peuvent être attribuées à des facteurs tels que la menace immédiate qu'un obstacle peut poser, le temps de réaction disponible pour le train, et les conséquences potentielles de l'inaction. Plutôt que de distinguer parmi ces nombreux états, qui pourraient offrir seulement des différences marginales dans le risque réel, il semble plus pragmatique et efficace de les regrouper en fonction de leur niveau de risque global. Cela simplifie non seulement le processus de prise de décision, mais assure également une clarté dans la définition des niveaux de risque distincts.

De plus, dans la configuration de simulation initiale décrite ici, on suppose que les obstacles détectés par l'unité de perception du train autonome sont statiques (c'est-à-dire, $v_o = 0$) dans l'environnement immédiat. Cette hypothèse simplifie l'aspect prédictif du mouvement et de la trajectoire de l'obstacle, permettant au processus de prise de décision de ne pas tenir compte de ces dynamiques. En conséquence, l'orientation (θ_d) des obstacles n'est pas prise en compte lors de la transition vers des états de sécurité discrets. L'accent est principalement mis sur l'identification des obstacles et l'évaluation de leur proximité avec le train (c'est-à-dire, la distance jusqu'à l'obstacle d_o). Au contraire, la deuxième configuration de simulation fait évoluer ce modèle en intégrant la vitesse des obstacles et leur nature (c'est-à-dire, statique ou dynamique). Cela reflète non seulement un scénario opérationnel plus réaliste, mais pousse également le système à prendre en compte la complexité supplémentaire dans son évaluation des risques et ses algorithmes de prise de décision. De plus, développer deux configurations de simulation met en évidence l'adaptabilité de l'approche, démontrant sa capacité à intégrer de multiples facteurs, qu'ils soient externes liés aux obstacles ou internes associés au train lui-même.

Sur la base des considérations décrites, nous avons identifié quatre états discrets. Cette délimitation n'est pas seulement une réduction, mais une classification méthodique et une catégorisation basée sur les niveaux de risque que plusieurs états dans la grille adaptative pourraient être associés. Cette approche structurée fournit une représentation claire des risques de collision, facilitant une réponse efficace par le système de train autonome aux situations critiques de sécurité. Les spécificités de ces quatre états sont détaillées dans l'Équation 3, comme indiqué ci-dessous.

$$S = \begin{cases} s_1 \\ s_2 \\ s_3 \\ s_4 \end{cases} = \begin{cases} \text{Safe} & = \{(g^x, g^y, \theta^d \mid g^x = 4, \forall g^y, \theta^d\} \\ \text{ObstacleDetected} & = \{(g^x, g^y, \theta^d \mid g^x = 3, \forall g^y, \theta^d\} \\ \text{AboutToCrash} & = \{(g^x, g^y, \theta^d \mid g^x = 2, \forall g^y, \theta^d\} \\ \text{Crash} & = \{(g^x, g^y, \theta^d \mid g^x = 1, \forall g^y, \theta^d\} \end{cases} \quad (3)$$

D. Modélisation de l'espace des actions

Le comportement dynamique du train est principalement contrôlé par l'action continue d'accélération (et de décélération intrinsèque et de freinage). Pour simplifier le modèle, nous considérons une discrétisation de l'espace des actions en trois valeurs discrètes $A = \{a_1, a_2, a_3\}$, qui représentent respectivement : le maintien de la vitesse, le freinage nominal, et le freinage d'urgence.

E. Modélisation de l'espace des observations

L'espace des observations, noté O , est défini comme l'ensemble des observations possibles que le train autonome peut effectuer à chaque pas de temps. Dans notre cas, l'espace d'observation comprend la position de l'obstacle dans la grille adaptative, représentée par les variables g^x et g^y . Cette représentation capture la localisation relative de l'obstacle par rapport à la position du train et permet l'évaluation des risques potentiels de collision. Ainsi, deux observations sont définies dans l'ensemble suivant :

$$O = \begin{cases} g^x, \text{ avec } g^x \in \{1,2,3,4\} \\ g^y, \text{ avec } g^y \in \{1,2,3\} \end{cases} \quad (4)$$

F. Modélisation de la fonction de transition

Sur la base de la distribution de probabilité de l'état initial (ou actuel) du modèle, à chaque pas de temps δt , une action est prise et la distribution de probabilité sur l'espace d'état est mise à jour selon le modèle de fonction de transition :

$$T(s, a, s') = P(s' | s, a) \quad (5)$$

La fonction de transition décrit le comportement dynamique du train et des obstacles après chaque intervalle de temps δt . Nous considérons v_T , x_T , et a_{cc}^T comme étant respectivement la vitesse, la position, et l'accélération du train, avec l'échantillon de temps δt . L'équation suivante montre le modèle de transition du train (c'est-à-dire, la dynamique du train) dans le système de coordonnées global (ou terrestre) :

$$\begin{pmatrix} v_T(t+\delta t) \\ x_T(t+\delta t) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \delta t & 1 \end{pmatrix} \cdot \begin{pmatrix} v_T(t) \\ x_T(t) \end{pmatrix} + \begin{pmatrix} \delta t \\ \frac{\delta t^2}{2} \end{pmatrix} \cdot a_{cc}^T \quad (6)$$

De même, le modèle de transition de l'obstacle (c'est-à-dire, la dynamique de l'obstacle) dans le système de coordonnées global est décrit comme suit :

$$\begin{pmatrix} v_O(t+\delta t) \\ x_O(t+\delta t) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \delta t & 1 \end{pmatrix} \cdot \begin{pmatrix} v_O(t) \\ x_O(t) \end{pmatrix} \quad (7)$$

Notez que dans le cas du modèle de transition de l'obstacle, l'accélération n'est pas considérée. De plus, nous supposons que les transitions sont déterministes, et que l'obstacle reste statique dans le temps. La nouvelle distance jusqu'à l'obstacle d_o après un pas de temps (c'est-à-dire, le modèle de transition de l'obstacle) dans le système de coordonnées global est représentée par l'équation suivante :

$$d_o(t + \delta t) = d_o(t) - v_T(t) \cdot \delta t - \frac{a_{cc}^T(t) \cdot \delta t^2}{2} \quad (8)$$

Cependant, la distance à l'obstacle dans le système de coordonnées local (grille adaptative) est définie comme suit :

$$d_o(t) = g_o^x(t) \quad (9)$$

G. Modélisation de la fonction d'observation

L'objectif principal de la fonction d'observation $Z(o|s, a)$, dans ce cas, est de calculer la distance parcourue par le train après un pas de temps, dans le système de coordonnées global. Cette distance permet de suivre la nouvelle distance jusqu'à l'obstacle pour chaque action sélectionnée depuis l'espace d'actions.

$$d_{traveledT}(t + \delta t) = v_T(t) \cdot \delta t + \frac{a_{cc}^T(t) \cdot \delta t^2}{2} \quad (10)$$

La nouvelle distance à l'obstacle, après un pas de temps, devient :

$$d_o(t + \delta t) = d_o(t) - (v_T(t) \cdot \delta t + \frac{a_{cc}^T(t) \cdot \delta t^2}{2}) \quad (11)$$

De manière similaire, l'orientation de l'obstacle peut être mise à jour, à chaque pas de temps δt , sur la base de l'orientation de l'obstacle au pas de temps précédent et de la vitesse angulaire de l'obstacle ω_o . Ainsi :

$$\theta_{dt+\delta t} = \theta_a(t) + \omega_o \cdot \delta t \quad (12)$$

Notez que l'équation 12 suppose que la vitesse angulaire de l'obstacle (ω_o) reste constante sur l'intervalle de temps δt . Cela correspond à l'hypothèse faite précédemment que l'obstacle reste statique dans le système de coordonnées global. Il en est de même dans le système de coordonnées local (grille adaptative).

H. Fonction de récompense

La fonction de récompense prend la forme de coûts (ou de récompenses négatives), attribués à chaque décision (action) prise par le modèle dans un état spécifié. Le rôle de la fonction de récompense est d'encourager les décisions qui favorisent les objectifs du système, tout en imposant des pénalités à celles qui ne le font pas. Bien que l'objectif principal du système d'anti-collision soit d'éviter les collisions de trains, il reste souhaitable de considérer d'autres objectifs secondaires, tels que le respect de l'horaire, le maintien d'une vitesse stable, etc.

Pour l'objectif principal, des récompenses négatives (c'est-à-dire, des pénalités) sont attribuées aux états considérés comme dangereux, tels que ceux ayant une forte probabilité de collision avec un obstacle (par exemple, l'état de *Crash*). En attribuant des récompenses négatives plus élevées aux états plus risqués, l'ADS peut être incité à prendre des actions plus sûres et à éviter les collisions. Cette adaptation de la récompense selon le risque incarne le modèle de risque mentionné dans la section III. Elle peut être mise à jour en temps réel à mesure que de nouvelles informations sur l'environnement deviennent disponibles, permettant au système de s'adapter continuellement aux conditions changeantes et de maintenir une opération sûre. De plus, la fonction de récompense attribue des valeurs numériques à chaque paire état-action pour simuler le comportement désiré. Dans notre cas, l'objectif principal du système est d'éviter autant que possible (c'est-à-dire, de minimiser le risque de) collisions de trains. Ainsi, nous définissons une pénalité importante pour le train d'être dans l'état de *Crash* (s_4), une autre pénalité pour l'état de

AboutToCrash (s_3), et une récompense pour être dans l'état Safe (s_1). La fonction de récompense est représentée par l'équation 13 :

$$R(s) = \begin{cases} 10 & \text{si } s = s_1 \\ -10 & \text{si } s = s_2 \\ -100 & \text{si } s = s_3 \\ -1000 & \text{si } s = s_4 \end{cases} \quad (13)$$

Une considération importante lors de la conception du modèle de risque pour l'ADS est le compromis entre la sécurité et l'efficacité. En particulier, pour des états tels que *ObstacleDetected* et *AboutToCrash* (c'est-à-dire, s_2 et s_3), la fonction de récompense devrait équilibrer le besoin d'éviter les collisions avec le besoin de maintenir un comportement de conduite efficace.

V. SIMULATIONS ET RESULTATS

Dans cette section, nous décrivons la configuration du modèle POMDP, la génération des incertitudes dans les expérimentations, et le processus d'initialisation des variables. Ensuite, nous présentons les résultats de ~~la~~ simulations effectuées sur deux scénarios : l'un avec des obstacles fixes, et l'autre avec des obstacles mobiles (les résultats du deuxième scénario sont présentés sur (Chelouati et al. 2023a)).

A. Création de l'état perçu

L'état perçu est crucial pour relier les environnements réels et simulés. En fait, la distance observée et les obstacles perçus dans cette simulation sont sujets à un bruit gaussien, émulant les incertitudes inhérentes à la détection dans le monde réel. La prochaine action du train est décidée en fonction de l'état perçu, dérivé de ces observations bruitées, et non de l'état réel.

B. Processus de génération d'obstacles

Les obstacles sont générés de manière stochastique de telle sorte qu'un obstacle apparaît environ 20% du temps, avec une distance par rapport au train tirée selon une distribution uniforme. Ce processus de génération d'obstacles introduit de la diversité dans la simulation et permet de tester la fiabilité de la prise de décision du train dans diverses situations. De plus, les obstacles sont générés, suivant une distribution uniforme, entre la moyenne des distances de freinage nominal et d'urgence (α_1 et α_2), et 50 mètres au-delà de cette moyenne. Cela garantit des distances générées ni trop proches ni trop lointaines afin de laisser suffisamment de temps au train autonome pour réagir tout en créant des situations de danger pouvant être détecté et maîtrisé.

C. Mise à jour de la croyance

La mise à jour de la croyance est un composant critique du modèle. Elle retient une distribution sur les états potentiels que le train autonome peut occuper, intégrant l'état actuel, l'état perçu et les actions prises. L'état de croyance est généré à chaque pas de temps, jouant un rôle essentiel dans la gestion des incertitudes dans le système et permettant une prise de décision plus robuste. L'équation de mise à jour de la croyance est donnée par :

$$b'(s') = \eta \cdot P(o | s', a) \cdot \sum_{s \in S} P(s' | s, a) \cdot b(s) \quad (14)$$

Dans l'équation 14, η est la constante de normalisation pour assurer que l'état de croyance mis à jour b' est une distribution de probabilité valide (c'est-à-dire, la somme est à 1 sur tous les états). $b(s)$ et $b'(s')$ sont, respectivement, la probabilité d'être dans l'état de croyance courant s et l'état de croyance mis à jour s' . Cette équation met à jour la croyance sur l'état courant après avoir pris une action a et observé un résultat o . La nouvelle croyance $b'(s')$ est proportionnelle à la vraisemblance de l'observation o menant à l'état s' , multipliée par la somme des probabilités d'atteindre s' à partir de tous les états possibles s sous une action a , pondérée par la croyance courante d'être dans l'état s .

D. Initialisation des variables

Avant l'exécution de la simulation, toutes les variables nécessaires associées aux états, actions et politique sont initialisées. Les paramètres initiaux pour la position du train, sa vitesse et sa distance par rapport à l'obstacle sont également établis. Au fur et à mesure que la simulation progresse, la position et la vitesse sont continuellement mises à jour selon l'action choisie et l'état courant du train (voir Tableau 1).

TABLE I. INITIALISATION DES VARIABLES

Variables	Valeur initiale	unité
Vitesse initiale du train	40	$m \cdot s^{-1}$
Position initiale du train	0	m
Distance de freinage nominal	300	m
Distance de freinage d'urgence	100	m
Temps d'échantillonnage	0,1	s
Récompenses	[10, -10, -100, -1000]	-
Actions	[0, -1, -3]	$m \cdot s^{-2}$

<i>Variables</i>	<i>Valeur initiale</i>	<i>unité</i>
Facteur	0,95	-

E. Formulation du risque

Une fois l'environnement perçu, l'étape suivante est l'estimation du risque. Ici, les scénarios possibles pouvant conduire à des conditions dangereuses/collisions sont identifiés, et leur probabilité est estimée en fonction des états actuels et prédits. Cela implique l'identification des dangers potentiels, l'évaluation de leur impact possible et le calcul du risque associé à chaque danger. À cette fin, le risque est calculé de deux manières, comme décrit dans les équations suivantes :

$$R_1 = 1 - \frac{1}{1 + \exp\left(-5 \cdot \frac{d_o}{\alpha_1}\right)} \quad (16)$$

$$R_2 = \frac{\alpha_1 - d_o}{\alpha_1 - \alpha_2} \quad (17)$$

L'équation 16 utilise une fonction logistique pour présenter le scénario où le risque est relativement faible lorsque le train est loin de l'obstacle ($d_o > \alpha_1$). L'utilisation de la fonction logistique offre une transition douce et sigmoïde d'un état à faible risque à un état à haut risque. Cette caractéristique est idéale pour représenter des scénarios où le risque est initialement faible mais augmente à mesure que le train se rapproche de l'obstacle, et finit par saturer lorsque l'obstacle est très proche. De plus, cette caractéristique prend en compte le fait que lorsque l'obstacle est suffisamment éloigné, le train a assez de temps pour réagir et le risque est faible. D'autre part, lorsque l'obstacle est très proche ($d_o < \alpha_2$) le train pourrait déjà avoir engagé son freinage d'urgence, impliquant qu'il a déjà reconnu le risque et tente de l'atténuer.

L'équation 17 augmente linéairement le risque à mesure que le train se rapproche de l'obstacle, de la distance de freinage nominale (α_1) à la distance de freinage d'urgence (α_2). Cela est logique car lorsque le train se trouve à sa distance de freinage nominale, il devrait idéalement commencer à décélérer pour éviter une collision, et le fait de ne pas le faire augmente progressivement le risque. Le risque atteint son pic lorsque le train est à sa distance de freinage d'urgence, signifiant que si le train ne s'arrête pas immédiatement, la collision est inévitable.

F. Résultats

Les figures suivantes illustrent la performance de la fonction d'anticollision de l'ADS dans un environnement ferroviaire dynamique, c'est-à-dire sa capacité à détecter et à agir face à un obstacle fixe, estimer les niveaux de risque et garantir des opérations sûres et efficaces. Le scénario avec obstacles fixes établit une base de référence pour démontrer les avantages de notre approche.

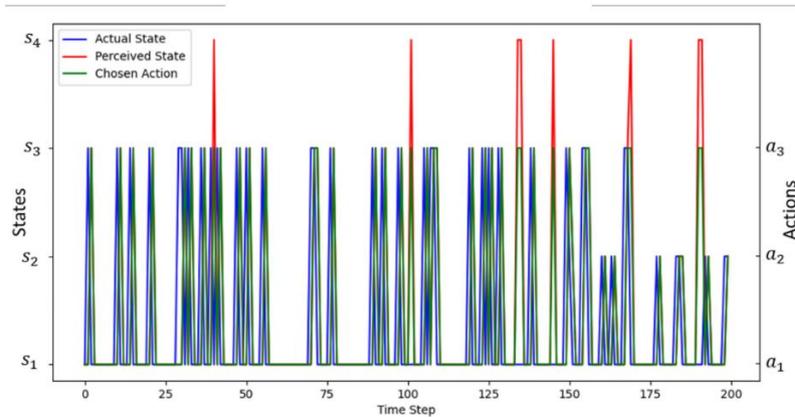


Figure 5 : les états réels, les états perçus et les actions choisies

La Figure 5 illustre l'évolution de l'état réel (en couleur bleue), de l'état perçu (en couleur rouge) et de l'action choisie (en couleur verte) dans le temps dans le cas des obstacles fixes. L'état réel représente l'état de vérité terrain du train, tandis que l'état perçu est basé sur les observations faites par les capteurs du train. L'action choisie est la décision prise par le modèle POMDP en fonction de l'état perçu. L'axe des x de la figure représente les différents pas de temps de simulation, tandis que l'axe des y (à gauche), les valeurs s_1, s_2, s_3 , et s_4 correspondent aux différents états dans lesquels le système peut se trouver. D'autre part, l'axe des y (à droite) représente également les actions disponibles que le système peut prendre en réponse à son état perçu. Ces actions sont représentées par a_1, a_2 , et a_3 .

La figure souligne l'efficacité du système à percevoir précisément son environnement étant donné que l'état perçu suit, dans la majorité des cas, la trajectoire de l'état réel. Cependant, des divergences occasionnelles entre les deux trajectoires (état perçu et état réel) sont présentes à des moments spécifiques. Ces divergences sont interprétées comme des faux positifs (percevoir un obstacle qui n'est pas présent/fausse alerte) et des faux négatifs (ne pas détecter un obstacle/détection manquée).

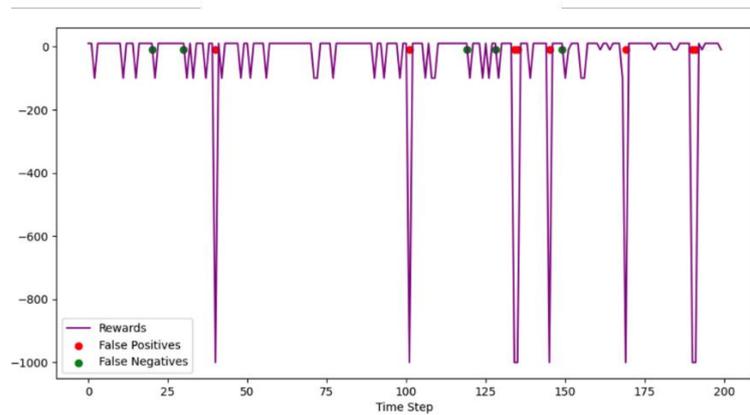


Figure 6 : les récompenses

La Figure 6 présente les récompenses (et pénalités) immédiates obtenues par le système au fil du temps. Les récompenses sont directement liées à l'état perçu et à l'action choisie. Les récompenses positives indiquent la sécurité (état *Safe*), tandis que les récompenses négatives représentent des risques potentiels (états *ObstacleDetected*, *AboutToCrash* et *Crash*). Les diagrammes de dispersion dans la figure mettent également en évidence les faux positifs (points verts) et les faux négatifs (points rouges) dans le processus de prise de décision, montrant des instances où la perception du modèle s'écarte de l'état réel. Des faux positifs notables se produisent aux moments 40, 101, 134, 135, 145, 169, 190, 191, tandis que les faux négatifs se produisent aux moments 20, 30, 119, 149.

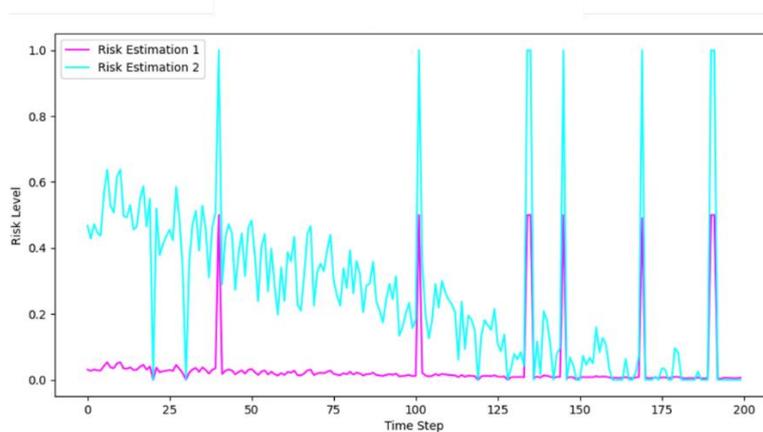


Figure 7 : l'estimation du risque

La Figure 7 illustre les deux méthodes d'estimation du risque formalisées dans le modèle par les équations 16 et 17. Ces estimations évaluent le niveau de risque associé à la distance observée jusqu'à l'obstacle. Des valeurs de risque plus élevées indiquent une probabilité plus élevée de collision. Le graphique permet une compréhension complète du processus d'évaluation du risque et de son rôle dans la détermination des actions appropriées.

L'estimation du risque 1 (représentée en couleur magenta) décrit principalement des scénarios à faible risque dans la plupart des états (c'est-à-dire, les états s_1 , s_2 et s_3), à l'exception de l'état *Crash* (c'est-à-dire, l'état s_4), où le risque est élevé. Cette approche semble optimiste, car elle maintient une évaluation du risque proche de 0. Au contraire, l'estimation du risque 2 (illustrée en couleur cyan) décrit une appréciation du risque plus dynamique. À mesure que le modèle passe de l'état *Safe* à l'état *AboutToCrash*, le risque augmente régulièrement, atteignant environ 0.5, indiquant un état de prudence accrue. Cependant, une fois que le modèle entre dans l'état *Crash*, le risque atteint sa valeur maximale de 1, soulignant les conséquences graves de cet état. Ces stratégies d'estimation du risque différentes mettent en lumière l'adaptabilité du modèle et révèlent la capacité de répondre à différents niveaux de risque.

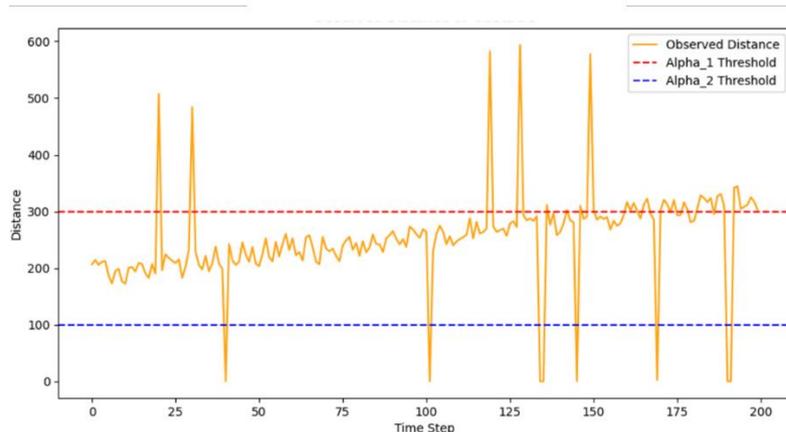


Figure 8 : la distance par rapport à l'obstacle

La Figure 8 représente la distance observée jusqu'à l'obstacle au fil du temps. Elle montre comment la distance perçue fluctue à mesure que les capteurs du train détectent et interagissent avec l'environnement. Les lignes pointillées rouge et bleue représentent les seuils pour les distances de freinage nominal et d'urgence (α_1 et α_2 , respectivement). Lorsque la distance observée franchit ces seuils, le modèle peut initier des actions de freinage en conséquence pour prévenir des collisions potentielles.

VI. CONCLUSION

Dans cet article, nous avons proposé une approche de prise de décision basée sur le risque pour les trains autonomes, exploitant les capacités des Processus Décisionnels de Markov Partiellement Observables (POMDP) pour apporter une surveillance environnementale efficace et en temps réel des trains. La contribution principale de cette étude réside dans la surveillance continue et l'estimation du risque, qui sont cruciales pour garantir l'exploitation sûre des trains autonomes. Cette approche intègre l'appréciation dynamique du risque au cœur du processus de décision, permettant au train de gérer de manière proactive les dangers potentiels de collision. Elle aborde efficacement les incertitudes tant dans l'état opérationnel du train que dans son interaction avec l'environnement. Ce faisant, l'approche améliore la capacité du train autonome à prendre des décisions pertinentes et sûres.

REMERCIEMENTS

Ce travail de recherche est financé par le programme français « investissements d'avenir » et fait partie du projet collaboratif français TASV (Train Autonome Service Voyageurs), avec Railenium, SNCF, Alstom Crespin, Thales, Bosch et Spirops.

RÉFÉRENCES

- Alsaba, Yamen, Marion Berbineau, Iyad Dayoub, Emilie Masson, Gemma Morral Adell, et Eric Robert. 2020. « 5G for Remote Driving of Trains ». P. 137-47 in *Communication Technologies for Vehicles*, édité par F. Krief, H. Aniss, L. Mendiboure, S. Chaumette, et M. Berbineau. Cham: Springer International Publishing.
- Boudardara, Fateh, Abderraouf Boussif, Pierre-Jean Meyer, et Mohamed Ghazel. 2023. « A review of abstraction methods towards verifying neural networks ». *ACM Transactions on Embedded Computing Systems*. doi: 10.1145/3617508.
- Chelouati, Mohammed, Abderraouf Boussif, Julie Beugin, et El-Miloudi El Koursi. 2022. « A framework for risk-awareness and dynamic risk assessment for autonomous trains ». P. 1-8 in *ESREL 2022, 32nd European Safety and Reliability Conference*.
- Chelouati, Mohammed, Abderraouf Boussif, Julie Beugin, et El-Miloudi El Koursi. 2023a. « A Risk-Based Decision-Making Process for Autonomous Trains Using POMDP: Case of the Anti-Collision Function ». *IEEE Access*.
- Chelouati, Mohammed, Abderraouf Boussif, Julie Beugin, et El-Miloudi El Koursi. 2023b. « Graphical safety assurance case using Goal Structuring Notation (GSN) — challenges, opportunities and a framework for autonomous trains ». *Reliability Engineering & System Safety* 230:108933. doi: 10.1016/j.ress.2022.108933.
- Gebauer, Oliver, Wolfgang Pree, et Burkhard Stadlmann. 2012. « Autonomously Driving Trains on Open Tracks — Concepts, System Architecture and Implementation Aspects ». 54(6):266-79. doi: 10.1524/itit.2012.0689.
- Ghahramani, Zoubin. 1998. « Learning Dynamic Bayesian Networks ». P. 168-97 in *Adaptive Processing of Sequences and Data Structures: International Summer School on Neural Networks "E.R. Caianiello" Vietri sul Mare, Salerno, Italy September 6–13, 1997 Tutorial Lectures*, édité par C. L. Giles et M. Gori. Berlin, Heidelberg: Springer.

- Heckerman, David. 2008. « A Tutorial on Learning with Bayesian Networks ». P. 33-82 in *Innovations in Bayesian Networks: Theory and Applications*, édité par D. E. Holmes et L. C. Jain. Berlin, Heidelberg: Springer.
- Kaelbling, L. P., M. L. Littman, et A. W. Moore. 1996. « Reinforcement Learning: A Survey ». *Journal of Artificial Intelligence Research* 4:237-85. doi: 10.1613/jair.301.
- Lagay, Rémy, et Gemma Morral Adell. 2018. « The Autonomous Train: a game changer for the railways industry ». P. 1-5 in *2018 16th International Conference on Intelligent Transportation Systems Telecommunications (ITST)*.
- Leurent, Edouard. 2018. « A survey of state-action representations for autonomous driving ».
- Mendiboure, Leo, Mohamed Lamine Benzagouta, Dominique Gruyer, Tidiane Sylla, Morayo Adedjouma, et Abdelmenname Hedhli. 2023. « Operational Design Domain for Automated Driving Systems: Taxonomy Definition and Application ». P. 1-6 in *2023 IEEE Intelligent Vehicles Symposium (IV)*.
- Monahan, George E. 1982. « State of the Art—A Survey of Partially Observable Markov Decision Processes: Theory, Models, and Algorithms ». *Management Science* 28(1):1-16. doi: 10.1287/mnsc.28.1.1.
- Parhizkar, Thierry. 2021. *Online Probabilistic Risk Assessment of Complex Marine Systems*.
- Rosique, Francisca, Pedro J. Navarro, Carlos Fernández, et Antonio Padilla. 2019. « A Systematic Review of Perception System and Simulators for Autonomous Vehicles Research ». *Sensors* 19(3):648. doi: 10.3390/s19030648.
- Temizer, Selim, Mykel Kochenderfer, Leslie Kaelbling, Tomas Lozano-Perez, et James Kuchar. s. d. « Collision Avoidance for Unmanned Aircraft using Markov Decision Processes* ». in *AIAA Guidance, Navigation, and Control Conference*. American Institute of Aeronautics and Astronautics.
- Tran, Duy Quang, et Sang-Hoon Bae. 2021. « Improved Responsibility-Sensitive Safety Algorithm through a Partially Observable Markov Decision Process Framework for Automated Driving Behavior at Non-Signalized Intersection ». *International Journal of Automotive Technology* 22(2):301-14. doi: 10.1007/s12239-021-0029-z.
- Trentesaux, Damien, Rudy Dahyot, Abel Ouedraogo, Diego Arenas, Sébastien Lefebvre, Walter Schön, Benjamin Lussier, et Hugues Chéritel. 2018. « The Autonomous Train ». P. 514-20 in *2018 13th Annual Conference on System of Systems Engineering (SoSE)*.