

Titre: Partage de tâches adaptatif dans une équipe humain-machine basé sur des modèles quantitatifs de performance et de confiance
Title:

Auteur: Clémence Dubois
Author:

Date: 2020

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Dubois, C. (2020). Partage de tâches adaptatif dans une équipe humain-machine basé sur des modèles quantitatifs de performance et de confiance [Mémoire de maîtrise, Polytechnique Montréal]. PolyPublie.
Citation: <https://publications.polymtl.ca/5568/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/5568/>
PolyPublie URL:

Directeurs de recherche: Jérôme Le Ny
Advisors:

Programme: Génie électrique
Program:

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

**Partage de tâches adaptatif dans une équipe humain-machine basé sur des
modèles quantitatifs de performance et de confiance**

CLÉMENCE DUBOIS

Département de génie électrique

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
Génie électrique

Décembre 2020

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé :

Partage de tâches adaptatif dans une équipe humain-machine basé sur des modèles quantitatifs de performance et de confiance

présenté par **Clémence DUBOIS**

en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

a été dûment accepté par le jury d'examen constitué de :

Roland MALHAMÉ, président

Jérôme LE NY, membre et directeur de recherche

Heather NEYEDLI, membre

REMERCIEMENTS

Mes remerciements vont en premier lieu vers l'institution qui a financé mes travaux à travers le programme «*Innovation for Defence Excellence and Security (IDEaS) Micro-Net program*» proposé par le Ministère canadien de la Défense nationale.

Je remercie ensuite l'ensemble des personnes qui font partie de ce programme et avec qui j'ai pris du plaisir à échanger des idées et à travailler.

Merci à mes professeurs de Polytechnique Montréal pour leur qualité d'enseignement et leur disponibilité.

Merci à mes camarades du laboratoire MRASL pour leur présence, leur soutien et leurs conseils au quotidien.

Merci à mon directeur de recherche, Jérôme le Ny, d'abord de m'avoir proposé ce sujet de recherche et ensuite de m'avoir accompagnée pendant ces deux ans. Merci à lui pour m'avoir encouragée à présenter notre travail à la conférence internationale SMC (*2020 IEEE International Conference on Systems, Man, and Cybernetics*) (Dubois et Le Ny, 2020).

Je remercie les personnes qui m'ont entourées pendant ces deux ans à Montréal. Parmi elles, merci à Estelle, Marco et mes colocataires. Finalement, merci à ma famille sans qui rien n'aurait été possible. Merci pour leur soutien matériel et moral indéfectible.

RÉSUMÉ

Lorsqu'une équipe humain-machine est formée dans le but d'accomplir des tâches de prise de décision, certains facteurs humains peuvent avoir un impact important sur les performances de l'équipe. C'est le cas en particulier de la charge de travail et de la confiance que place un opérateur dans les capacités de la machine. Il peut alors être nécessaire de prendre en compte ces facteurs dans le design de la stratégie de collaboration. Cependant, la confiance et la charge de travail varient dans le temps en fonction des interactions entre l'humain et la machine. Une façon de prendre en compte ces éléments est d'opter pour une stratégie de collaboration adaptative, c'est-à-dire qui varie en fonction de l'état cognitif de l'opérateur. Dans ce mémoire on propose une stratégie de collaboration adaptative sous la forme de suggestions automatiques et dynamiques de partage de tâche. Régulièrement, une proposition de partage de tâche est suggérée à l'opérateur en prenant en compte sa charge de travail ainsi que sa confiance. Cette stratégie est issue de la résolution d'un Processus Décisionnel Markovien Partiellement Observable (POMDP). Pour cela des modèles quantitatifs des performances humaines et de la dynamique de la confiance ont été sélectionnés. Des simulations permettent de montrer le potentiel de la méthode en comparant les performances de la stratégie adaptative proposée à celles d'une stratégie statique plus simple. Les résultats à long terme de l'équipe humain-machine sont en moyenne meilleurs de 24% lorsque la stratégie adaptative est appliquée plutôt que la stratégie statique. L'utilisation de modèles quantitatifs dont certains paramètres doivent être identifiés pose la question de la robustesse de la stratégie aux erreurs de calibration. On montre, toujours en simulation, que malgré certaines erreurs de modèles, la stratégie proposée conserve son avantage.

ABSTRACT

In mixed-initiative systems where human and automation collaborate in order to complete a decision-making task, some human factors can have an impact on the team performance. For instance, the cognitive workload and the trust placed by the operator on the automation capabilities can be determining factors. Hence it could be relevant to take into account these cognitive variables in the design of the collaboration strategy. However, both workload and trust fluctuate with the history of past interactions. One way to include these dynamic variables is to opt for an adaptive collaboration strategy. In this work, we propose an adaptive task allocation suggestion which dynamically allocate task according to the operator's trust level. This adaptive strategy is computed by solving a Partially Observable Markovian Decision Process (POMDP). The POMDP is defined using quantitative models of human performance and trust dynamic. We study this method's potential by comparing, in simulation, the performance results collected when the adaptive strategy is applied and those when a static strategy is applied. The long term mean team reward is 24% higher with the adaptive strategy than with the static strategy. Moreover we study the impact of model calibration errors on the strategy performance. The proposed method seems to bring benefits even in the presence of errors in the models.

TABLE DES MATIÈRES

REMERCIEMENTS	iii
RÉSUMÉ	iv
ABSTRACT	v
TABLE DES MATIÈRES	vi
LISTE DES TABLEAUX	ix
LISTE DES FIGURES	x
LISTE DES SIGLES ET ABRÉVIATIONS	xii
CHAPITRE 1 INTRODUCTION	1
1.1 La collaboration humain machine : les problématiques	1
1.1.1 Le contexte	1
1.1.2 Les particularités du «système» humain	2
1.2 Le problème de partage de tâches en classification	3
1.2.1 Description du scénario	3
1.2.2 Justification du scénario	4
1.3 Objectifs de recherche	5
1.3.1 Développer une solution adaptative au problème de partage de tâches	5
1.3.2 Sélectionner des modèles quantitatifs de performance et de la confiance	
humaine	5
1.3.3 Évaluer la solution adaptative	5
1.4 Plan du mémoire	6
CHAPITRE 2 REVUE DE LITTÉRATURE	7
2.1 La charge de travail	7
2.1.1 Définitions et justification	7
2.1.2 Modéliser la charge de travail	8
2.2 La confiance	9
2.2.1 Définitions et justifications	9
2.2.2 Prédire le comportement d'un humain en modélisant la confiance . . .	11

2.3	Conception de stratégies de collaboration humain-machine basées sur des modèles quantitatifs	14
2.3.1	Des stratégies d'allocation de contrôle (contrôle manuel ou automatique)	14
2.3.2	Des stratégies de collaboration influençant le comportement de la machine via des états dégradés	14
2.3.3	Des stratégies de collaboration influençant le comportement de la machine avec un impact continu de la confiance	15
2.3.4	Des stratégies de collaboration promouvant une charge de travail appropriée	16
2.3.5	Des cadres de conception promouvant une confiance appropriée	16
2.4	Conclusion de la revue de littérature	16
CHAPITRE 3 SOLUTION STATIQUE		18
3.1	Un modèle de performance humaine dépendant de la charge de travail	18
3.1.1	Performance d'un classificateur binaire	18
3.1.2	Impact de la charge de travail	22
3.2	Calibration du modèle	25
3.2.1	Récolte des données	26
3.2.2	Identification des paramètres	27
3.3	Une solution statique au problème de partage de tâches	29
3.3.1	Le problème d'optimisation statique	29
3.3.2	Résultats du problème d'optimisation	31
3.3.3	Conclusion	34
CHAPITRE 4 SOLUTION ADAPTATIVE		35
4.1	Un modèle dynamique de la confiance	35
4.1.1	Définition et dynamique de la confiance	35
4.1.2	Définition de la dépendance à la machine	37
4.1.3	Remarques sur le modèle de la confiance	38
4.2	Formulation du problème en MDP	39
4.2.1	Description	39
4.2.2	Implémentation et résolution numérique	40
4.2.3	Résultats de simulation	42
4.3	Formulation du problème en POMDP	45
4.3.1	Description	45
4.3.2	Implémentation et résolution numérique	47
4.3.3	Résultats de simulation	48

4.4	Expérience 1 : Modèles parfaits	49
4.4.1	Comportement en régime permanent	49
4.4.2	Comportement en régime transitoire	51
4.4.3	Effet de l'observation partielle	52
4.5	Expérience 2 : Modèles imparfaits	54
4.5.1	Comparaison entre les politiques statique et adaptative	54
4.5.2	Effets du bruit sur le suivi de l'état	54
4.6	Expérience 3 : Erreurs d'identification des modèles	55
4.6.1	Erreur constante dans le modèle de performance	56
4.6.2	Erreur dans le modèle de confiance : constantes de temps	57
4.6.3	Erreur dans le modèle de la confiance : fonction de capacité réelle de la machine	58
	CHAPITRE 5 CONCLUSION	61
5.1	Synthèse des travaux	61
5.2	Limitations de la solution	62
5.2.1	Variété des scénarios	62
5.2.2	Difficulté de calibration des modèles	63
5.3	Améliorations futures	64
5.3.1	Validation expérimentales	64
5.3.2	Étudier les effets négligés dans les modèles	64
5.3.3	Comparer la solutions à d'autres approches	64
	RÉFÉRENCES	66

LISTE DES TABLEAUX

Tableau 1.1	Résultats possibles d'une tâche de classification binaire	4
Tableau 1.2	Notations	5
Tableau 3.1	Résultats d'une période de travail	31
Tableau 4.1	Valeurs numériques des paramètres des différents modèles	41
Tableau 4.2	Comparaison de la durée du régime transitoire lorsque le comportement de l'opérateur est déterministe	51

LISTE DES FIGURES

Figure 3.1	Exemple de courbes ROC	19
Figure 3.2	Schéma du modèle Gaussien à Variances Égales de la Théorie de la Détection du Signal	21
Figure 3.3	Exemple de courbes ROC tracées avec la loi logistique	22
Figure 3.4	Famille de courbes ROC indexées par la charge de travail	24
Figure 3.5	Courbes de performance humaine isobiais	25
Figure 3.6	Courbes de performance humaine isocritère tracées avec le modèle Gaussien à Variances Égales (GVE) (3.5) complété par la dépendance à la charge de travail W donnée par (3.8)	26
Figure 3.7	Données expérimentales	28
Figure 3.8	Données expérimentales dans le plan gaussien	28
Figure 3.9	Modèle de performance identifié	30
Figure 3.10	Maximisation de \tilde{f} (3.19)	34
Figure 4.1	Modèle de dépendance de l'humain à la machine avec $\nu = 5$, $\bar{T} = 0.75$ et $\epsilon = 0.1$	38
Figure 4.2	Modèle de performance de classification de l'humain et de la machine	42
Figure 4.3	Simulation appliquant la politique du MDP	43
Figure 4.4	Simulation appliquant une politique statique	44
Figure 4.5	Récompenses cumulées amorties récoltées pendant les simulations des figures 4.3 et 4.4	44
Figure 4.6	Simulations appliquant la politique du POMDP (figures du haut) et la politique statique (figures du bas)	49
Figure 4.7	Récompenses cumulées amorties récoltées pendant deux simulations, l'une appliquant la stratégie du <i>Partially Observed Markov Decision Process</i> (POMDP), l'autre appliquant la stratégie statique	50
Figure 4.8	Simulations du régime permanent avec $\bar{T} = C^* = 0.8$	51
Figure 4.9	Simulations du régime permanent avec $\bar{T} = 0.6 \neq C^*$	52
Figure 4.10	Simulations avec $\bar{T} = C^* = 0.8$ et $x_0^h = (0, 0)^T$	53
Figure 4.11	Simulations où l'état de croyance initial de la machine est exact . . .	53
Figure 4.12	Récompenses totales amorties collectées dans 5000 simulations de 50 périodes de travail	55
Figure 4.13	Erreur de suivi de l'état de croyance, telle que défini par (4.15) et (4.16)	56
Figure 4.14	Robustesse à une erreur constante dans le modèle de performance . .	57

Figure 4.15	Robustesse à une erreur constante dans le modèle de confiance	59
Figure 4.16	Robustesse à une erreur de définition de C	60

LISTE DES SIGLES ET ABRÉVIATIONS

GVE Gaussien à Variances Égales.

LoA *Levels of Automation.*

MDP *Markov Decision Process.*

POMDP *Partially Observed Markov Decision Process.*

ROC *Receiver Operating Characteristic.*

TDS Théorie de la Détection du Signal.

CHAPITRE 1 INTRODUCTION

Dans ce chapitre, on commencera par évoquer les diverses problématiques rencontrées lorsqu'on conçoit une machine destinée à travailler en collaboration avec un humain. Ensuite, un scénario de collaboration précis sera introduit comme support de travail. Enfin on formulera les objectifs qui ont orienté les travaux présentés dans ce mémoire.

1.1 La collaboration humain machine : les problématiques

1.1.1 Le contexte

Lorsqu'on décide qu'une machine¹ doit collaborer avec un humain, c'est que l'équipe performe mieux que chacun des partenaires s'ils devaient réaliser la tâche seuls. Une approche possible pour concevoir cette équipe est d'automatiser le maximum de fonction possible et de laisser le reste à l'humain. Or, aujourd'hui, on est capable d'automatiser des fonctions de plus en plus complexes et dans des environnements de plus en plus incertains. Les ingénieurs doivent donc repenser la façon dont ils conçoivent les équipes humain-machine (Cummings, 2014). D'abord on peut s'intéresser à l'allocation des tâches, c'est-à-dire quelle tâche doit être confiée à quel partenaire. En introduisant le concept de niveau d'automatisation (ou *Levels of Automation* (LoA)) (Parasuraman *et al.*, 2000), on diversifie les types de collaboration entre humain et machine. L'échelle des niveaux d'automatisation liste les façons dont peut être partagée l'autorité. Autrement dit, une tâche n'a pas à être réalisée soit entièrement par un humain soit entièrement par la machine. Ensuite, on peut poser la question de savoir comment la machine et l'humain doivent collaborer. Une des réponses possibles est les stratégies adaptatives (Rouse, 1988), (Parasuraman *et al.*, 1992). Une stratégie de collaboration adaptative consiste à faire varier certains aspects d'une tâche en fonction de paramètres relatifs au système (par exemple la demande, les performances de la machine etc...) ou relatifs à l'humain (ses performances, sa charge mentale etc...). Trois questions se posent dans la conception de stratégies adaptatives : que faut-il adapter, en fonction de quoi on adapte et comment adapter (Wickens *et al.*, 2015)[Chapitre 12]. Évidemment les réponses à ces questions dépendent beaucoup du contexte particulier.

Dans ce mémoire, on s'intéresse à la conception de stratégies de collaboration adaptatives. En particulier on conçoit une stratégie adaptative correspondant à un scénario donné. Cette

1. Dans ce mémoire, on appelle «machine» tout système comportant une partie automatisée. Cela comprend les systèmes d'aide à la décision (e.g. aide à la conduite), des logiciels de supervision de systèmes complexes (e.g. supervision de centrale nucléaire) ou encore des robots (e.g. rover d'exploration).

stratégie doit améliorer les performances globales de l'équipe humain-machine en s'appuyant sur les caractéristiques d'un humain.

1.1.2 Les particularités du «système» humain

Lorsqu'on souhaite concevoir une machine qui prend en compte l'humain, il faut comprendre quelles sont les difficultés à surmonter pour espérer atteindre cet objectif. En particulier, plusieurs problématiques sont posées par les caractéristiques du «système» humain.

D'abord, il faut rappeler la multiplicité des facteurs impactant le comportement humain dans un contexte de collaboration avec une machine. Ces facteurs peuvent provenir de différentes sources telles que la machine (ses performances, sa lisibilité etc...), l'environnement (les risques, la présence de collaborateurs humains etc...) ou l'humain lui même (sa charge mentale, son stress, ses compétences etc...). En plus de la multiplicité des sources influençant le comportement humain, il faut ajouter l'importance de l'historique des interactions entre l'humain et la machine. En effet, l'humain se base en partie sur ses expériences passées pour prendre ses décisions, par exemple quand il fait le choix de faire confiance ou non à une machine pour réaliser une tâche.

Afin de guider des stratégies de conception qui soient précises, et pas seulement sous forme de recommandations, il est impératif de quantifier un certain nombre de phénomènes. Par exemple, si on souhaite concevoir un système qui minimise la charge mentale d'un opérateur, il faut pouvoir représenter la charge mentale par un scalaire. Or il n'y a pas de consensus sur une définition de la charge mentale comme une variable unidimensionnelle qu'il serait possible, en plus, de mesurer. En outre, on ne dispose pas de modèle physique du cerveau humain qui permettrait de prédire le comportement humain en général. Enfin, bien que des modèles du comportement humain ont été proposés dans des contextes particuliers, ils doivent répondre à la question des différences inter-personnelles. En effet, chaque humain ne réagit pas de la même façon même si les situations sont identiques. Doit-on alors concevoir des stratégies de collaboration uniques ou personnalisées ?

Même si on a défini des concepts pour aider notre compréhension de l'humain (charge mentale, confiance, confiance en soi, conscience de la situation etc...), ces éléments ne sont pas directement observables. On ne sait observer que la version subjective de ces concepts par des questionnaires. On peut également observer directement le comportement de l'humain, ses performances ou encore son utilisation de la machine sans l'interrompre dans ses activités. Il est aussi possible de mesurer des manifestations physiques un peu plus fines grâce à des capteurs physiologiques plus ou moins invasifs (rythme cardiaque, mouvement des yeux, activité électrique du cerveau etc...). Dans tous les cas, il est difficile d'observer les processus

cognitifs qui conduisent à tel ou tel comportement.

1.2 Le problème de partage de tâches en classification

Dans cette section, on introduit le scénario de collaboration humain-machine pour lequel on concevra la stratégie adaptative.

1.2.1 Description du scénario

Dans ce scénario de collaboration, l'équipe humain-machine doit réaliser des tâches de classification binaire d'évènements aléatoires sur la base d'un certain nombre d'informations incomplètes. Chaque évènement ou objet peut être classé soit par la machine soit par l'opérateur. On suppose que les évènements sont indépendants les uns des autres. De plus ces évènements surviennent de manière périodique dans le temps.

Concrètement, on peut formuler la tâche de classification sous la forme d'un test d'hypothèse. Pour chaque évènement k , la machine ou l'humain doit choisir parmi l'hypothèse $H_{1,k}$ (l'évènement est intéressant) ou $H_{0,k}$ (hypothèse nulle). Lorsque le décideur choisit l'hypothèse $H_{1,k}$ (respectivement $H_{0,k}$) on notera sa décision $D_k = 1$ (respectivement $D_k = 0$). Son choix se fait à partir d'informations Y_k disponibles, définies comme une variable aléatoire.

Étant donné les quatre issues possibles d'une classification binaire (tableau 1.1), les capacités d'un classificateur binaire peuvent être représentées par ses probabilités de vrais positifs $P_{VP} := P(D_k = 1|H_{1,k})$ et de faux positifs $P_{FP} := P(D_k = 0|H_{0,k})$. On suppose que ces probabilités ne dépendent pas du temps. À chaque issue possible de la classification (vrai positif, faux positif, faux négatif ou vrai négatif), on associe une récompense positive ou négative qu'on note respectivement $R^{VP} \geq 0$, $R^{FP} \leq 0$, $R^{FN} \leq 0$ et $R^{VN} \geq 0$. On supposera de plus que $R^{VP} \neq R^{FN}$ et $R^{FP} \neq R^{VN}$. Enfin un coût additionnel $R^h \leq 0$ est collecté lorsque la classification est réalisée par l'opérateur. Ce coût permet de prendre en compte les situations où l'opérateur a d'autres tâches à accomplir que la classification d'évènement. Ainsi, la problématique du temps perdu sur la réalisation des tâches secondaires est intégrée dans la réalisation de la tâche principale.

Dans ce problème, le temps est divisé en *périodes de travail*, indexées par t . Une période de travail correspond à N tâches de classification successives (N est constant). Au début de chaque période de travail, la machine suggère un partage de tâche à l'opérateur, noté a_t , c'est-à-dire la proportion d'objet que l'humain devrait classer pendant cette période. L'opérateur a l'autorité sur la machine et décide s'il veut compléter plus ou moins de tâches que proposé par la machine. Par exemple, un opérateur n'ayant pas confiance dans les capacités

Tableau 1.1. Résultats possibles d'une tâche de classification binaire.

Classe prédite	Classe réelle	
	Positif ($H_{1,k}$)	Négatif ($H_{0,k}$)
Positif ($D_k = 1$)	Vrai positif	Faux positif
Négatif ($D_k = 0$)	Faux négatif	Vrai négatif

du classificateur automatique aura tendance à classer manuellement tous les évènements.

Remarque. Le choix du nombre N d'évènements dans une période de travail doit être fait en fonction du scénario particulier. Si N est trop faible, on risque de brouiller l'opérateur avec des changements de suggestion trop fréquents de la part de la machine. En revanche si N est trop grand, on donne moins de chance à la suggestion automatique d'allocation de tâche de prendre en compte l'état instantané de l'opérateur.

1.2.2 Justification du scénario

En premier lieu, on choisit une tâche de classification binaire car elle est suffisamment précise pour être décrite formellement. En outre, il s'agit d'un premier exemple de collaboration relativement simple qui pourrait être complexifié par la suite selon plusieurs directions (classification en plus de deux catégories, environnement non statique etc...).

Le scénario choisi met en scène une tâche de prise de décision dans un environnement incertain. Le décideur est contraint d'adopter un comportement basé sur des connaissances ou une expertise. L'humain est connu pour être plus performant dans ce genre de tâche relativement à ce que pourrait faire une machine (Cummings, 2014). Cependant les outils d'apprentissage permettent de plus en plus aux machines d'aborder des problèmes de décision dans des environnements complexes et incertains. Il est alors pertinent d'imaginer une collaboration entre humain et machine pour des problèmes de classification d'évènements.

De plus, le cadre est suffisamment général pour englober des situations diverses. On peut imaginer des contextes industriels comme l'identification de défauts de fabrication sur des objets ou le classement de courrier selon deux zones géographiques. Ce scénario peut correspondre également à des contextes de sécurité comme l'identification d'évènements à risque sur un ensemble d'écrans de vidéo-surveillance ou la différenciation entre des véhicules alliés et des véhicules hostiles sur un écran radar.

Tableau 1.2. Notations.

p	Probabilité qu'un objet soit intéressant ($p = P(H_1)$)
P_{VP}	Probabilité qu'un classificateur réalise un vrai positif
P_{FP}	Probabilité qu'un classificateur réalise un faux positif
N	Nombre d'objet ou d'évènement à classer pendant une période de travail
W_t	Proportion d'objet classé par l'humain pendant la période de travail t
a_t	Proportion d'objet classé par l'humain suggéré par la machine

1.3 Objectifs de recherche

1.3.1 Développer une solution adaptative au problème de partage de tâches

On souhaite proposer une méthode de conception d'une stratégie de partage de tâches adaptative. Wickens, dans (Wickens *et al.*, 2015)[Chapitre 12], recommande qu'il faut adapter le paramètre qui influence le plus la charge mentale de travail de l'opérateur. On a vu dans le scénario décrit que la machine doit régulièrement proposer une répartition des tâches entre l'humain et la machine. L'objectif est donc de rendre la suggestion de partage de tâches adaptative.

1.3.2 Sélectionner des modèles quantitatifs de performance et de la confiance humaine

Étant donné le contexte du scénario, on décide que la solution adaptative doit prendre en compte deux facteurs humains susceptibles d'avoir un impact important sur les performances de l'équipe : la charge de travail de l'humain et la confiance qu'il place dans la machine. Une revue de littérature doit justifier ce choix. Enfin, la prise en compte de ces deux facteurs nécessite la définition de modèles quantitatifs liant charge mentale, confiance et performance de l'équipe.

1.3.3 Évaluer la solution adaptative

Outre la conception d'une stratégie adaptative qui prend en compte les concepts de charge de travail et de confiance, on souhaite évaluer la pertinence de la démarche. En particulier on cherche à comparer cette stratégie à une stratégie statique, i.e. dont la suggestion de partage de tâches est constante. La comparaison doit se faire en terme de performance de l'équipe à long terme mais également en terme de robustesse aux erreurs de modèle et de facilité

d'application.

1.4 Plan du mémoire

Ce mémoire s'organise de la façon suivante. Le chapitre 2 offre une revue de littérature qui, à la fois, place ces travaux dans leur contexte de recherche et apporte des justifications à un certain nombre de choix. Le chapitre 3 introduit le modèle liant charge de travail et performance humaine et apporte une première solution statique au problème de partage de tâches. Le chapitre 4, quant à lui, présente un modèle dynamique de la confiance qui sera ensuite utilisé dans la conception de la stratégie adaptative. On présente également dans ce chapitre les résultats des simulations permettant d'évaluer la pertinence de l'approche choisie. Enfin, le dernier chapitre conclut ces travaux en évoquant leurs limites et de potentielles améliorations.

CHAPITRE 2 REVUE DE LITTÉRATURE

Dans ce chapitre, on cherche à placer le sujet de ce mémoire dans son contexte. Cela permettra en outre de justifier un certain nombre de points dans la démarche adoptée. En particulier, on commencera par introduire le concept de charge de travail ainsi que des approches qui ont été adoptées dans la littérature pour modéliser l'impact de cette charge de travail sur les performances humaines. Dans un deuxième temps, on propose de définir ce qu'est la confiance que place un humain dans une machine. De même, on explore les différentes façons de modéliser ce concept psychologique. Enfin, la dernière section tache de donner une idée au lecteur des solutions proposées dans les dernières années pour prendre en compte, dans le design de stratégies de collaboration humain-machine, ces deux facteurs humains que sont la confiance et la charge de travail.

2.1 La charge de travail

2.1.1 Définitions et justification

Définitions

Lorsqu'on s'intéresse au facteur humain, on a l'habitude de distinguer la charge de travail objective de la charge mentale de travail. La charge de travail objective correspond à la quantité de tâche qu'on demande à un humain alors que la charge mentale reflète les ressources cognitives mobilisées par l'humain pour réaliser les tâches demandées. Évidemment, la charge de travail objective impacte la charge mentale mais ne la caractérise pas à elle seule. Par exemple, la personnalité, les compétences et l'attention d'un opérateur sont d'autres facteurs influençant la charge mentale d'un individu.

Impact de la charge de travail sur les performances humaines

L'article (Young *et al.*, 2015) offre une revue des études abordant le concept de charge mentale en ergonomie depuis leur début, dans les années 80. La charge mentale est un des concepts le plus étudié en facteur humain. La principale motivation pour chercher à définir et mesurer la charge mentale est l'étude de son impact sur les performances humaines. En particulier, on cherche à prévenir les situations dans lesquelles un individu serait surchargé ou sous-chargé. En effet, lorsqu'un opérateur est surchargé mentalement, il sera susceptible de commettre des erreurs. À l'inverse, lorsque sa charge mentale est trop basse, cela peut provoquer des

problèmes de perte d'attention, de perte de compétences ou encore d'ennui.

Mesurer la charge mentale de travail

De nombreuses études ont été proposées pour mesurer la charge mentale d'un individu lorsqu'il accomplit une tâche. La première solution consiste à questionner ce dernier sur son ressenti. On accède alors à la vision subjective de la charge mentale de l'opérateur (Moray, 1982). Cette mesure ne reflète pas exactement la quantité des ressources cognitives dépensées par l'individu mais permet tout de même des études comparatives. L'un des questionnaires les plus répandus semble être le *NASA-Task Load index (NASA-TLX)* (Hart et Staveland, 1988). Celui-ci, originellement développé pour le domaine de l'aviation, est aujourd'hui très largement utilisé pour évaluer la charge mentale d'un individu (Hart, 2006). La deuxième façon de mesurer la charge mentale consiste à observer un certain nombre de manifestations physiologiques causées par celle-ci. Il est possible d'observer, par exemple, la dilatation des pupilles, le rythme cardiaque ou encore l'activité électrique du cerveau. (Charles et Nixon, 2019) offre une revue récente des techniques de mesure de charge mentale via l'utilisation de capteurs physiologiques.

2.1.2 Modéliser la charge de travail

On souhaite montrer ici que divers modèles quantitatifs de la charge mentale de travail ont été proposés dans la littérature. On peut classer ces modèles en deux catégories : les modèles qui sont capables de prédire la charge mentale d'un individu et les modèles qui, en plus, vont prédire les performances de cet individu.

Modèles prédisant la charge mentale

Les modèles de charge mentale varient d'abord en fonction de leur domaine d'application. Les auteurs de (Loft *et al.*, 2007) offrent une revue des modèles prédisant la charge mentale dans le contexte de contrôle aérien et proposent à leur tour un autre modèle dans (Loft *et al.*, 2009). Par exemple, dans (Manning *et al.*, 2002), des modèles construits par régression, sont capables de prédire la charge mentale subjective de contrôleurs aériens à partir du nombre, de la durée et du type de contenu des communications orales. D'autres méthodes, listées dans (da Silva, 2014), sont utilisées dans le domaine de la conduite d'un véhicule. Wickens a, quant à lui, proposé un modèle qui traite spécifiquement des situations où l'opérateur doit accomplir plusieurs tâches et qui fait la distinction entre les types de ressources mobilisées (perception visuelle, auditive, analyse, réponse verbale, manuelle etc...) (Wickens, 2008).

Modèles prédisant les performances humaines

On cite ici quelques travaux qui ont pour objet de prédire les performances humaines en prenant en compte l'impact de la charge de travail. Par exemple, dans (Sorkin et Woods, 1985), Sorkin et Woods proposent de modéliser les performances de classification d'un humain en utilisant un modèle de la Théorie de la Détection du Signal et en faisant varier l'un des paramètres du classifieur humain avec la fréquence d'apparition des éléments à traiter. Ici cette fréquence peut s'interpréter comme la charge de travail objective de l'humain. Dans (Rouse *et al.*, 1993), les auteurs avancent un modèle linéaire capable de prédire la charge mentale subjective de l'humain ainsi que son comportement. Plus récemment, les auteurs de (Wu et Liu, 2007) proposent un modèle capable de prédire les performances de conduite d'un individu en estimant à chaque instant leur charge mentale grâce à l'observation de son comportement et de ses réponses à un questionnaire. Ce modèle prend également en compte l'âge du conducteur. Enfin, une méthode estimant la charge mentale et la fiabilité d'un opérateur grâce à ses interactions avec une machine est utilisée dans (Gregoriades et Sutcliffe, 2008).

2.2 La confiance

2.2.1 Définitions et justifications

Un exemple

Lorsqu'une équipe de collaboration humain-machine est formée, il est fréquent que l'opérateur ait l'autorité sur la machine. Cette autorité peut lui permettre de reprendre le contrôle manuel en cas de doutes sur les capacités de la machine. Typiquement, un opérateur qui viendrait d'assister à plusieurs échecs successifs d'une machine préférera, par manque de confiance, réaliser les tâches manuellement plutôt que de continuer à les déléguer à la machine. Il est possible alors que l'opérateur se surcharge volontairement le conduisant à faire des erreurs alors que la machine est toujours capable de réaliser une partie des tâches. Dans cet exemple il apparaît d'une part que le choix de l'opérateur d'avoir recours partiellement, totalement ou pas du tout à la machine (on appellera cela la **dépendance** ou *reliance*), a un impact significatif sur les performances de l'équipe humain-machine. D'autre part, la dépendance de l'opérateur à la machine semble directement liée à sa confiance.

Les facteurs de la dépendance de l'homme à la machine

Il est connu qu'une dépendance inappropriée peut avoir un impact négatif allant de la performance sous-optimale (Robinette *et al.*, 2016) à l'accident. Parasuraman et Riley classent cet impact en deux catégories : la sous-exploitation (*disuse*) et la sur-exploitation (*misuse*) (Parasuraman et Riley, 1997). Dans le premier cas, le problème vient du fait que l'opérateur n'utilise pas suffisamment la machine, dans le second, l'opérateur se repose à tort sur elle. Ce sont deux situations qu'il est préférable d'éviter lorsqu'on conçoit une stratégie de collaboration. Plusieurs facteurs ont été identifiés comme influant le comportement de dépendance de l'humain à la machine. La confiance que place un opérateur dans les capacités de la machine apparaît comme un facteur très important (Lee et See, 2004), (Dzindolet *et al.*, 2003). Cependant, des travaux ont mis en évidence le rôle de la confiance en soi (Lee et Moray, 1994), la charge mentale de travail (Parasuraman et Riley, 1997), la pression temporelle (Rice et Keller, 2009), la difficulté apparente de la tâche (Schwark *et al.*, 2010) ou encore le style de comportement (van den Brule *et al.*, 2014). Les importances relatives de chacun de ces facteurs dépendent du type de tâche demandé et de l'environnement. Certains ont proposé des outils pour les identifier (Van Dongen et Van Maanen, 2013), (Dzindolet *et al.*, 2001), (Inagaki et Itoh, 2010).

Les facteurs de la confiance

Dans la littérature, on définit la confiance que place un individu dans un agent comme la propension de cet agent à aider un individu dans la réalisation de ses objectifs dans un environnement incertain et comportant des risques (Lee et See, 2004). Il s'agit d'un processus psychologique qui permet de guider facilement le comportement d'un individu car il lui serait extrêmement coûteux en ressources cognitives de prendre des décisions sans cette aide devant des machines très complexes.

La confiance est un processus cognitif interne à un individu et est, par conséquent, difficile à mesurer. Le principal moyen d'évaluer la confiance passe par des mesures subjectives. C'est-à-dire qu'on demande au participant de remplir un questionnaire. On peut citer par exemple l'échelle *Trust Perception Scale-HRI* spécialisée dans le domaine des interactions entre humain et robot (Schaefer, 2016), ou encore (Yagoda et Gillan, 2012a), (Park *et al.*, 2008), (Yagoda et Gillan, 2012b) ou (Madsen et Gregor, 2000). Le principal reproche fait aux mesures subjectives est leur caractère invasif et incompatible avec des tâches pour lesquelles le temps est compté. D'autres travaux tentent de mesurer la confiance par des mesures comportementales. En particulier, dans (Kaniarasu *et al.*, 2012), les auteurs sont capables de détecter des changements de confiance en observant le nombre d'alarmes lancées avant

que l'opérateur baisse le niveau d'autonomie de la machine ainsi que le temps écoulé entre la dernière alarme et le moment où il agit ainsi. De même, dans (Freedy *et al.*, 2007), on propose une mesure de la confiance dite *rationnelle* à partir de l'observation des comportements de l'humain et de la machine.

À la difficulté de mesurer la confiance s'ajoute celle de la complexité de ses facteurs. Hancock *et al* (Hancock *et al.*, 2011) classent ces facteurs affectant la confiance en trois catégories : les facteurs relatifs au robot, à l'humain et à l'environnement. Parmi les éléments émanant de la machine, on trouve la fiabilité (Sanchez *et al.*, 2014), la lisibilité (Oduor et Wiebe, 2008), la prédictabilité (van den Brule *et al.*, 2014) ou encore les types d'erreurs (Dixon *et al.*, 2007), (Madhavan *et al.*, 2006) . Le niveau d'automatisation a également un impact sur le poids de ces paramètres sur la confiance (Chavaillaz *et al.*, 2016) ainsi que le type des informations partagées (Gao et Lee, 2006a). Parmi les facteurs environnementaux, on peut trouver entre autre la complexité de la tâche et les risques encourus (Perkins *et al.*, 2010). Les facteurs liés à l'humain comprennent les particularités culturelles (Sanchez *et al.*, 2014), la personnalité de l'opérateur (Szalma et Taylor, 2011), son âge (McBride *et al.*, 2010). On peut également citer son expertise représentée par la qualité du modèle mental qu'il se fait de la machine et de l'environnement (Wilkison, 2008) ou encore sa confiance initiale (Zhou, 2011).

2.2.2 Prédire le comportement d'un humain en modélisant la confiance

Remarques sur les modèles quantitatifs du comportement humain

Avant de concevoir une stratégie de collaboration entre humain et machine il est nécessaire de comprendre le comportement des deux partenaires dont celui de l'humain en particulier. Cette compréhension peut passer par des études qualitatives desquelles résultent des recommandations appliquées par les professionnels. Cette compréhension peut également passer par l'élaboration de modèles quantitatifs capable de prédire le comportement humain. Ces modèles pourraient alors guider, avec davantage de précision que des recommandations, la conception de stratégies de collaboration. Cependant, on a vu dans la section précédente la diversité des facteurs influençant la dépendance d'un humain à une machine. À celle-ci s'ajoute la diversité des facteurs impactant la confiance. De plus ces facteurs ne sont ni facilement quantifiables ni facilement observables. Parasuraman (Parasuraman, 2000) formule donc quelques remarques sur la nature des modèles quantitatifs. D'abord il semble impossible qu'un modèle de comportement humain soit pertinent dans le cas général de toutes les situations de collaboration humain-machine. Les différents modèles qui ont été proposés dans la littérature ne sont donc pas en compétition. De plus, certains modèles sont dits *normatifs*, c'est-à-dire qu'ils se basent sur une certaine théorie de la façon dont un humain prend une

décision. Le fait que cette théorie soit en partie fausse ou trop simpliste ne rend pas obsolète un modèle quantitatif. Celui-ci peut tout de même guider une stratégie de collaboration qui va dans le sens d'une amélioration des performances de l'équipe.

Des approches de modélisation du comportement de dépendance humain

La première étape pour espérer prédire la confiance d'un humain envers une machine a été d'en comprendre tous les facteurs. Ainsi les premiers modèles de la confiance sont qualitatifs. Par de multiples expériences réalisées dans divers contextes, on tente de dresser la liste exhaustive des paramètres à prendre en compte si on souhaite modéliser la confiance. Ce travail a été fait dans (Desai, 2012) dans un contexte de téléopération d'un robot mobile, dans (Boubin *et al.*, 2017) ou encore dans (Sanders *et al.*, 2011).

Les premiers modèles quantitatifs ont vu le jour dans les années 90. Dans (Lee et Moray, 1992), Lee et Moray proposent un modèle linéaire dynamique de la confiance. La confiance d'un opérateur au temps t dépend de sa confiance au temps $t - 1$ ainsi que des performances et des erreurs de la machine aux temps t et $t - 1$. Dans (Gao et Lee, 2006b) les auteurs proposent également un modèle explicite mais celui-ci trouve ses bases dans la théorie de la décision humaine. Ce modèle, en plus de prédire si l'opérateur utilisera le contrôle manuel ou automatique, décrit les processus psychologiques derrière la construction de la confiance. L'entrée de ce modèle est une valeur, non précisée, décrivant les capacités réelles de la machine. Ce modèle a été étendu, dans (Gao *et al.*, 2006), pour simuler une tâche où plusieurs humains travailleraient avec une machine. Les autres modèles quantitatifs ne cherchent pas forcément à fonder leur structure sur les théories du comportement humain.

Parmi les modèles linéaires, on peut trouver également (Xu et Dudek, 2016) qui observe le comportement humain en plus de celui de la machine pour prédire la confiance. Dans (Li, 2020), l'état de confiance d'un humain est prédit par un filtre de Kalman dans le contexte du contrôle d'un essaim de robot. Les entrées de ce modèle sont les interventions de l'humain et les performances de l'algorithme de contrôle des robots. De même, (Azevedo-Sa *et al.*, 2020) avance un modèle estimant la confiance dans un véhicule à conduite autonome. Cette estimation se fait grâce aux observations du mouvement des yeux du participant, de son temps d'utilisation de la conduite automatique et de ses performances sur une tâche secondaire à la supervision de la conduite.

Étant donné la nature aléatoire d'un certain nombre d'éléments dans un scénario de collaboration humain-machine (l'environnement, le comportement de la machine, le comportement humain, les observations etc...) et le faible nombre d'observations, des auteurs ont choisi d'utiliser des outils de statistique bayésienne. La confiance estimée d'un humain est représentée

par une distribution de probabilité. Cette distribution s'affine au fur et à mesure des interactions. Dans (Liu, 2020), Liu fait une revue des modèles bayésiens de la confiance humaine. On peut citer par exemple (Guo et Yang, 2020), (Van Maanen *et al.*, 2007) et le modèle OPTIMo (*Online Probabilistic Trust Inference Model*) (Xu et Dudek, 2015) qui prédit la confiance et les interventions d'un humain à partir des performances du robot, des interventions passées de l'opérateur et de quelques retours sous forme de questionnaires.

Enfin une dernière approche consiste à construire un modèle de la confiance par l'apprentissage d'un réseau de neurones artificiels. C'est le cas dans (Farrell et Lewandowsky, 2000) par exemple. Les outils en apprentissage profond et apprentissage par renforcement semblent ouvrir la voie vers ces modèles de la confiance basés largement sur les données. Par exemple dans (Nam *et al.*, 2017), les auteurs cherchent à modéliser le comportement humain lors du contrôle d'un essaim de robot. En utilisant l'apprentissage par renforcement inversé ils parviennent à identifier les paramètres qui impactent la confiance ainsi que de modéliser quantitativement cet impact. Les auteurs de (Soh *et al.*, 2019) proposent et comparent des modèles bayésiens et neuronaux. Ils concluent entre autre que ces modèles peuvent être complémentaires et qu'une approche hybride devrait être investiguée.

Critères de choix d'un modèle

Dans (Hiatt *et al.*, 2017), les auteurs proposent une revue des différentes techniques de modélisation du comportement humain dans les contextes de collaboration humain robot. Ces techniques sont classées en fonction de leur niveau de modélisation. Le niveau *computational* modélise ce que l'humain fait, le niveau *algorithmique*, comment il le fait et les processus psychologiques sous-jacents et le niveau *implémentational* la façon dont ces processus psychologiques se réalisent physiquement. En outre, les auteurs identifient plusieurs critères qui doivent être pris en compte dans le choix d'un modèle. Parmi ceux-ci on trouve la part du modèle qui est apprise à partir de données, la quantité de données nécessaire à la convergence du modèle, la présence d'une structure spécifiée «à la main», l'échelle de temps sur laquelle le modèle réalise ses prédictions ou encore la façon dont le modèle prend en compte ses erreurs.

Ces différents critères confirment que les modèles quantitatifs proposés dans la littérature sont aujourd'hui variés et complémentaires. La diversification des outils utilisés pour aborder le concept de confiance humain machine montre le dynamisme de ce sujet de recherche tant sur l'aspect de la mesure que sur celui, peut-être plus délicat, de sa modélisation.

2.3 Conception de stratégies de collaboration humain-machine basées sur des modèles quantitatifs

On s'intéresse ici aux diverses solutions qui ont été apportées au problème d'amélioration de la collaboration humain-machine. Plutôt que de proposer une liste exhaustive des stratégies de collaboration qui ont été publiées, on choisit dans cette section de citer une dizaine d'entre elles, parmi les plus récentes, démontrant leur diversité. On s'intéresse particulièrement aux solutions qui cherchent à maintenir une charge de travail et/ou un niveau de confiance appropriés. En outre, chacune des solutions citées sont basées sur un modèle quantitatif du comportement humain. Ces modèles sont utilisés pour prédire les performances, les décisions ou le comportement de dépendance d'un humain à la machine.

2.3.1 Des stratégies d'allocation de contrôle (contrôle manuel ou automatique)

Une des façons d'éviter qu'un opérateur sous-exploite ou sur-exploite une machine est de guider son choix de passer du contrôle manuel au contrôle automatique et vice versa. Certaines stratégies consistent à avertir l'opérateur lorsque la machine considère qu'il serait mieux de changer le contrôle d'un système. Dans (Saeidi et Wang, 2019), les auteurs utilisent des modèles dits objectifs de la confiance et de la confiance en soi ainsi que des modèles de performance humaine pour proposer une stratégie d'allocation dynamique de tâche. Ils testent leur solution dans un contexte de téléopération d'un robot mobile et montrent que leur stratégie est plus avantageuse (meilleures performances et charge mentale réduite) que celle consistant à laisser le choix du mode de contrôle à l'opérateur seul. Dans (Wang *et al.*, 2018), une stratégie de contrôle est proposée concernant le choix de la trajectoire de plusieurs robots. Cette stratégie intègre également un modèle dynamique de la confiance humaine.

2.3.2 Des stratégies de collaboration influençant le comportement de la machine via des états dégradés

Influencer ou contraindre le choix du mode de contrôle (manuel ou automatique) permet d'éviter directement des situations de sur-exploitation ou de sous-exploitation. C'est un exemple d'allocation de niveau d'automatisation adaptatif. En effet, le niveau d'automatisation s'adapte ici à la confiance de l'humain en la machine, la confiance de la machine en l'humain et même de la confiance de l'humain en lui-même. Cependant, comme le décrit (Muslim et Itoh, 2019), il n'est pas nécessaire d'adopter une stratégie d'allocation de contrôle adaptative pour améliorer la collaboration dans une équipe humain-machine. On peut également concevoir des stratégies adaptatives de partage de tâches mais qui ne changent pas le

niveau d'automatisation. Par exemple, la machine peut adapter son comportement au niveau de confiance de l'humain mais celui-ci garde toujours l'autorité sur la machine.

Parmi ces stratégies qui ne changent pas le niveau d'automatisation de la collaboration, on trouve celles qui consistent à passer d'un état de fonctionnement normal à un état de fonctionnement dégradé, et vice versa, en fonction du niveau de confiance estimé de l'opérateur. Par exemple, dans la stratégie *TACTiC* (*Trust-Aware Conservative Control*) (Xu et Dudek, 2016) (Xu, 2016), un robot mobile va altérer son comportement lorsque son superviseur humain perd confiance en lui. Cet état de fonctionnement altéré consiste à rendre les mouvements et les décisions du robot plus prédictibles et compréhensibles pour l'humain (vitesse réduite et mouvements plus lisses). La confiance du superviseur humain est estimée par le modèle *OPTIMo* évoqué plus haut. Dans le même esprit, dans (Sadrifaridpour, 2018), on cherche à améliorer la collaboration dans une tâche de manipulation d'objet partagée par un humain et un bras robotique. Lorsque la confiance de l'opérateur est au dessus d'un certain seuil, le robot est pro actif dans le choix de la trajectoire et estime la trajectoire souhaitée par son partenaire humain lui permettant ainsi de mieux prendre en charge les efforts lors de la manipulation. Lorsque la confiance est trop basse, le choix de la trajectoire est laissé à l'humain. L'estimation de la confiance de l'humain est basée sur les observations des performances du robot mais aussi de la force appliquée par l'humain sur l'objet à déplacer, interprétée comme le désaccord entre le robot et l'humain.

2.3.3 Des stratégies de collaboration influençant le comportement de la machine avec un impact continu de la confiance

On a vu des stratégies qui proposent un mode dégradé lorsque la confiance descend en-dessous d'un certain seuil. Voyons ici des stratégies qui intègrent de manière continue le niveau de confiance de l'humain dans le comportement de la machine, c'est-à-dire sans le comparer à un simple seuil. Par exemple, dans (Chen *et al.*, 2020), un certain nombre d'objets avec des récompenses et des risques différents, doivent être déplacés par un bras robotique. L'opérateur qui supervise le robot peut choisir de l'interrompre pour déplacer l'objet à sa place si il pense que le robot va échouer. Les auteurs proposent une stratégie qui indique quel objet doit être déplacé à partir d'un estimé de l'état de confiance de l'humain. Cette stratégie est le résultat de la résolution d'un processus de décision markovien partiellement observable. Dans (Saeidi *et al.*, 2017), il est question de téléopération dans laquelle la commande du robot est partagée entre l'humain et l'algorithme. La part du contrôle de l'algorithme est pondérée par la confiance estimée de l'humain. De plus, cette confiance est également prise en compte dans le retour haptique fourni à l'opérateur.

2.3.4 Des stratégies de collaboration promouvant une charge de travail appropriée

Jusque là, on a évoqué des stratégies qui promeuvent principalement une confiance appropriée de l'humain et qui, parfois, prennent en compte ses performances et l'impact de la charge de travail sur ses performances. Les deux prochains exemple de solutions adaptatives intègrent en particulier des modèles de la charge de travail. Dans (Srivastava *et al.*, 2012), la machine doit choisir le temps laissé à l'humain pour chacune des décisions qu'il doit prendre. La conception de cette stratégie utilise des modèles décrivant la prise de décision humaine, le compromis entre rapidité et précision ainsi que la conscience de la situation. Dans (Savla et Frazzoli, 2011), un modèle dynamique de la charge mentale permet de concevoir une stratégie d'affectation de tâches à un opérateur adaptée à sa charge mentale courante.

2.3.5 Des cadres de conception promouvant une confiance appropriée

Lorsqu'une solution de collaboration est proposée, elle s'inscrit en général dans un contexte relativement particulier (manipulation d'objet, téléopération, supervision etc...). Cependant, certains travaillent sur des méthodes de conception beaucoup plus générales. C'est le cas de (Zhang et Lin, 2019) où les auteurs proposent de concevoir une stratégie de collaboration humain-machine en utilisant le formalisme des POMDP. Dans un exemple, ils prennent en compte l'humeur d'un conducteur comme un des états du POMDP montrant le caractère adaptatif de leur solution. De même, dans (Floyd *et al.*, 2014), un formalisme est proposé pour permettre à une machine de choisir à chaque instant le comportement promouvant le plus la confiance parmi un ensemble de comportement disponible. Cette méthode permet, de plus, de prendre en compte les retours exprimés par l'humain (Floyd *et al.*, 2015).

2.4 Conclusion de la revue de littérature

Ce chapitre a permis de mettre en lumière les points suivants. D'abord la charge de travail et la confiance sont deux facteurs qui peuvent impacter les performances d'une équipe humain-machine. Il semble donc important de les prendre en compte si on souhaite concevoir une machine centrée sur l'humain. De plus, on a vu que beaucoup se sont attelés à modéliser ses concepts, y compris de manière quantitative, en guise de premier pas vers la conception de meilleures stratégies de collaboration. Certains de ces modèles ont d'ailleurs déjà pris part à l'élaboration de ces stratégies. La solution proposée dans ce mémoire s'inscrit dans cette démarche puisqu'il s'agit de proposer une stratégie de partage de tâches basée sur des modèles quantitatifs de confiance et de charge de travail.

L'originalité de la solution proposée dans ce mémoire réside dans la combinaison du/de

- contexte, i.e., la suggestion de partage de tâches pour des séquences de classifications,
- la prise en compte de l'impact de la charge de travail sur les performances,
- la prise en compte de la dynamique de la confiance.

En effet, ni la méthode, ni les modèles, ni le concept de partage de tâches adaptatif ne sont à eux seuls une nouveauté dans le domaine de la collaboration humain-machine. Cependant l'association de ces divers éléments appliqués à ce contexte constitue l'intérêt de ce mémoire.

CHAPITRE 3 SOLUTION STATIQUE

Dans ce chapitre, une première solution statique est proposée pour répondre au problème de partage de tâches décrit dans la section 1.2. Cette stratégie est conçue pour prendre en compte l'impact de la charge de travail de l'opérateur sur ses performances. Pour ce faire on modélise la relation entre performance et charge de travail en étendant les concepts de la Théorie de la Détection du Signal (TDS), théorie régulièrement utilisée pour aborder des problèmes de performances humaines. Il résulte du modèle de performance une stratégie statique de partage de tâches entre humain et machine. Cette stratégie servira de point de comparaison dans la suite de ce mémoire.

La section 3.1 introduit le modèle liant performance humaine et charge de travail. Ensuite on discute de la façon dont on pourrait identifier les paramètres de ce modèle dans la section 3.2. Enfin on développe dans la section 3.3 la solution statique au problème de distribution de tâche.

3.1 Un modèle de performance humaine dépendant de la charge de travail

3.1.1 Performance d'un classificateur binaire

Caractéristique de performance (ou *Receiver Operating Characteristic* (ROC))

La capacité d'un classificateur binaire peut être caractérisée par son taux de vrais positifs (P_{VP}) et celui de faux positifs (P_{FP}) espérés lors d'une séquence de plusieurs classifications. On appelle *caractéristique de performance* (ou *Receiver Operating Characteristic* (ROC)) (Fawcett, 2006) l'ensemble des points (P_{FP}, P_{VP}) obtenus en faisant varier le seuil de discrimination du classificateur. La figure 3.1 montre des exemples de courbes ROC. Le point $(0, 1)$ représente un classificateur parfait. Il détecte 100% des objets intéressants et ne lance aucune fausse alarme. Un point sur la diagonale partant de $(0, 0)$ à $(1, 1)$ représente un classificateur aléatoire, i.e., qui décide de la classe d'un objet en lançant une pièce. Les courbes de la figure 3.1 sont tracées avec l'équation (3.5) avec $d = 2$ (trait pointillé) et $d = 0.8$ (trait plein).

Les courbes ROC sont utilisées comme outil graphique pour qualifier les performances en détection de signal, classification d'évènement ou encore en diagnostic médical car il met en évidence le compromis entre bonnes détections et fausses alarmes (Egan, 1975; Swets, 1988). Plus récemment, cet outil est souvent utilisé pour visualiser les performances des classificateurs construits par apprentissage (Spackman, 1989).

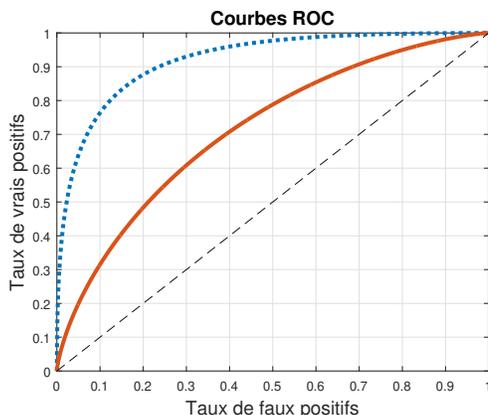


Figure 3.1 Exemple de courbes ROC

Il est possible de tracer la courbe ROC d'un classificateur directement à partir de données expérimentales. Lorsque le classificateur est un opérateur humain, il peut être pertinent d'utiliser un modèle guidant le tracé de la caractéristique de performance. L'utilisation d'un modèle sous-jacent permet de

1. faciliter le tracé de sa caractéristique de performance en réduisant cette étape à l'identification de quelques paramètres seulement,
2. interpréter l'impact de certains paramètres comme la charge de travail de l'opérateur sur ses performances.

Modèle proposé par la Théorie de la Détection du Signal (TDS)

Le modèle de performance humaine proposé dans ce mémoire est basé sur la Théorie de la Détection du Signal (TDS) (Peterson *et al.*, 1954). Cette théorie est un moyen de modéliser et d'évaluer la capacité à détecter un signal dans un environnement bruité. Initialement les concepts de la Théorie de la Détection du Signal (TDS) ont été appliqués à la détection automatique de signal mais rapidement des travaux ont montré les apports de cet outil en psychologie (Green et Swets, 1966). Les principes fondamentaux peuvent être trouvés dans (Wickens, 2002) et (Wickens *et al.*, 2015, Chapitre 2).

Le modèle le plus simple proposé dans la TDS est appelé modèle GVE. Reprenons le problème de classification énoncé en introduction 1.2. Pour plus de clarté on omettra l'indice k représentant l'index d'un objet parmi une séquence de classification. On a défini Y comme une variable aléatoire représentant les informations disponibles pour permettre au classificateur de faire son choix. Supposons que les mécanismes d'acquisition et de traitement de l'information d'un humain transforme ce signal Y en une autre variable aléatoire I continue

et unidimensionnelle appelée *réponse interne*.

Le modèle Gaussien à Variances Égales fait l'hypothèse que la réponse interne I suit la loi $\mathcal{N}(0, 1)$ quand l'hypothèse H_0 est vraie ou la loi $\mathcal{N}(d, 1)$ dans le cas de l'hypothèse H_1 , avec $\mathcal{N}(\mu, \sigma^2)$ la loi normale de moyenne μ et de variance σ^2 . On notera d , un réel positif, la *sensibilité*. La décision de l'humain est basée sur un seuil unique λ . Si $I > \lambda$ alors la décision sera $D = 1$, sinon l'humain décide $D = 0$ (voir la figure 3.2). Plutôt que de parler du seuil λ dont la valeur dépend de la variable abstraite qu'est la réponse interne I , on préférera définir le *biais* β représentant la propension d'un classificateur à choisir une hypothèse plutôt qu'une autre. On définit le biais comme le rapport des probabilités que la réponse interne corresponde au seuil λ sachant que le signal vienne des hypothèses H_1 ou H_0 ,

$$\beta = \frac{P(I = \lambda | H_1)}{P(I = \lambda | H_0)}. \quad (3.1)$$

Le biais β grandit avec le seuil λ . Dans le cas du modèle GVE, il s'agit du rapport des densités de probabilités des variables $I|H_1$ et $I|H_0$ suivants des lois normales, évaluées en λ . Après calcul, on obtient

$$\beta = e^{\lambda d - \frac{1}{2}d^2}. \quad (3.2)$$

Le modèle GVE a pour particularité de distinguer la sensibilité du décideur de son biais. Typiquement ce modèle est capable de déterminer si des erreurs de classifications sont causées par une faible sensibilité ou un biais inapproprié. Le biais est fixé par l'opérateur mais plusieurs facteurs peuvent l'influencer comme la perception de la fréquence de présence des signaux et celle de l'importance relative des erreurs. Une formation peut également impacter le biais. La sensibilité, quant à elle, peut être interprétée comme la capacité du mécanisme de détection à distinguer les stimuli provenant du signal de ceux provenant du bruit. Plus la sensibilité est grande, plus on peut espérer différencier facilement les signaux. Si $d = 0$ alors aucun indice ne permet de guider la classification.

Tracé de la courbe ROC avec le modèle GVE

On note Φ la fonction de répartition d'une variable aléatoire X suivant la loi normale centrée réduite, définie pour tout $x \in \mathbb{R}$ par

$$\begin{aligned} \Phi(x) &= P(X \leq x) \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt \end{aligned} \quad (3.3)$$

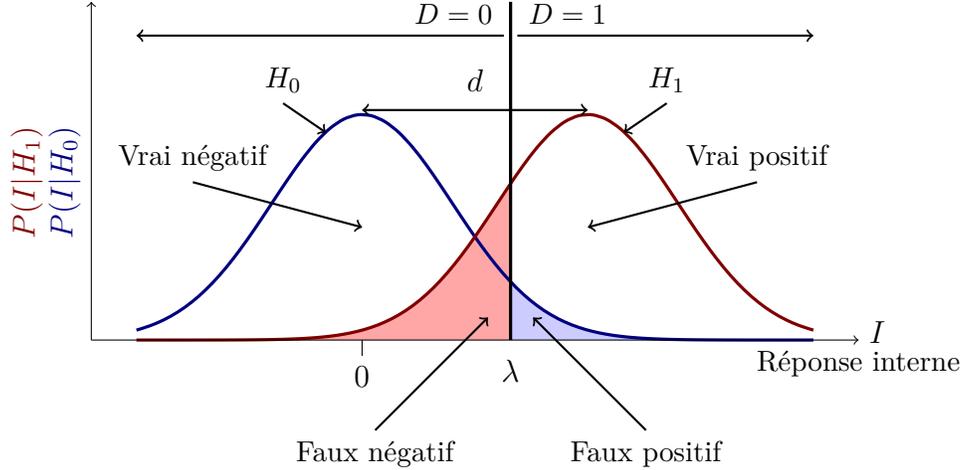


Figure 3.2 Schéma du modèle Gaussien à Variances Égales de la Théorie de la Détection du Signal

Dans le modèle GVE la probabilité de réaliser un vrai positif est donnée par

$$\begin{aligned}
 P_{VP} &= P(I \geq \lambda | H_1) \\
 &= 1 - P(X + d < \lambda) \quad \text{avec } X \sim \mathcal{N}(0, 1) \\
 &= \Phi(d - \lambda)
 \end{aligned} \tag{3.4}$$

Or on a également la probabilité de réaliser un faux positif $P_{FP} = P(I \geq \lambda | H_0) = \Phi(-\lambda)$. D'où l'expression de la courbe ROC résultant du modèle GVE :

$$P_{VP} = \Phi(d + \Phi^{-1}(P_{FP})). \tag{3.5}$$

La figure 3.1 montre deux courbes ROC tracées avec l'équation (3.5) pour deux différentes valeurs de sensibilité d . Plus la sensibilité est faible plus la courbe s'aplatit contre la diagonale (où $d = 0$), c'est-à-dire que le classificateur devient aléatoire.

Remarque. On pourrait supposer que la réponse interne I suive une loi logistique plutôt qu'une loi gaussienne comme dans le modèle GVE. En remplaçant Φ par la fonction de répartition de la loi logistique standard dans l'équation (3.5), on obtient une courbe ROC décrite par

$$P_{VP} = \frac{P_{FP}}{P_{FP} + (1 - P_{FP})e^{-d}}. \tag{3.6}$$

La figure 3.3 montre un exemple de courbes ROC tracées en supposant que la réponse interne suive des lois logistiques. Les courbes sont tracées avec l'équation (3.6) avec $d = 2$ (trait

pointillé) et $d = 0.8$ (trait plein). Ce choix a l'avantage d'offrir une formulation explicite des courbes ROC contrairement au cas gaussien. Les démonstrations mathématiques de la fin de ce chapitre en seraient simplifiées. On décide cependant de s'en tenir au modèle gaussien car, tel que justifié dans la TDS (Wickens, 2002), d'un point de vue empirique, les données semblent souvent révéler des distributions gaussiennes. Ceci est appuyé par la théorie avec le théorème central limite qui garantit que la somme de variables aléatoires indépendantes et identiquement distribuées converge en loi vers une loi normale.

3.1.2 Impact de la charge de travail

On souhaite étendre le modèle de performance humaine donné par la TDS en prenant en compte l'impact de la charge de travail de l'humain.

Définition de la charge de travail W

On définit la charge de travail $W \in [0, 1]$ représentant la quantité de travail effectuée pendant une période de travail. Si N objets doivent être classés pendant une période de travail et si l'humain effectue la classification de n^h objets alors la machine doit classifier les n^m objets restants ($n^m = N - n^h$). Dans la suite on utilisera régulièrement l'exposant h (respectivement m) pour désigner une variable se rapportant à l'humain (respectivement à la machine). On définit la charge de travail par :

$$W := \frac{n^h}{N}. \quad (3.7)$$

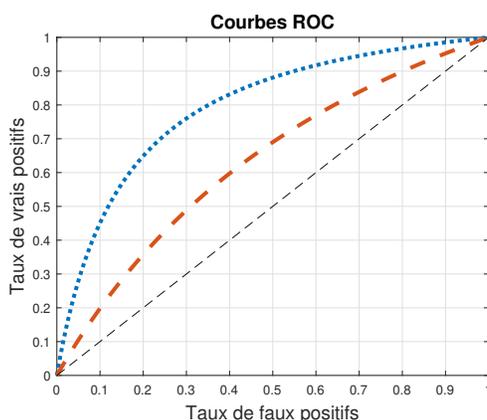


Figure 3.3 Exemple de courbes ROC tracées avec la loi logistique

Impact de la charge de travail W sur la sensibilité d

Dans le cadre du modèle Gaussien à Variances Égales, trois paramètres peuvent être impactés par la charge de travail : la sensibilité d , le critère λ et le biais β . On note que ces paramètres sont reliés entre eux par l'équation (3.2). On cherche à introduire dans ces paramètres une dépendance à W .

Dans (Putri *et al.*, 2016), les auteurs évaluent l'impact de la fréquence d'apparition d'un signal sur la performance de détection dans un environnement multi-tâche. Dans cette étude les participants ont comme tâche primaire de détecter des alarmes dans une interface présentant 52 jauges de niveau. Deux scénarios de complexité différente sont testés (le nombre d'alarmes est multiplié par deux entre les scénarios). Les données recueillies servent à identifier le modèle gaussien de la TDS. Les résultats montrent que la sensibilité diminue lorsque la quantité d'alarme à détecter augmente tandis que le biais des participants ne semble pas significativement impacté. Bien que la tâche ainsi que l'environnement de cette étude soient différents de ceux énoncés dans notre problème de classification binaire d'évènement, on décide d'inclure la charge de travail dans le paramètre de sensibilité.

Dans (Sorkin et Woods, 1985), les auteurs proposent des modèles de performance humaine dans le cas où l'opérateur doit valider ou non des alarmes lancées par un système automatique. En particulier, ils proposent deux modèles reliant sensibilité de l'humain, au sens de la TDS, et la fréquence des alarmes. Dans les deux cas, la sensibilité humaine est une fonction décroissante de la fréquence des alarmes. Dans ce contexte, on peut supposer que la fréquence d'apparition des alarmes s'apparente à la charge de travail telle qu'on l'a défini.

On propose le modèle linéaire suivant :

$$d = d(W) = d_0 + d_1 W. \quad (3.8)$$

Ce modèle très simple est choisi pour sa facilité d'identification. Il est linéaire et ne dispose que de deux paramètres $d_0 \geq 0$ et $d_1 \leq 0$. De plus il traduit la diminution de la sensibilité lorsque la charge de travail augmente.

Ainsi on obtient une famille de courbes ROC indexées par la charge de travail W (figure 3.4). Plus la charge de travail est grande plus les points (P_{VP}, P_{FP}) se rapprochent de la diagonale.

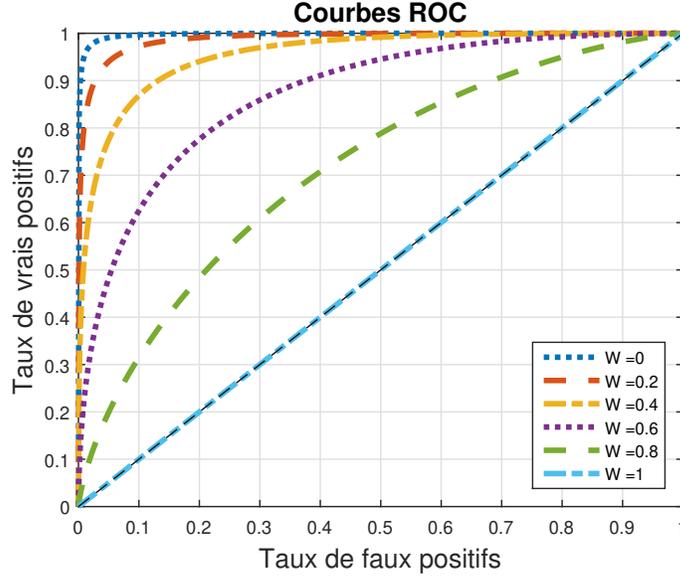


Figure 3.4 Famille de courbes ROC (équation (3.5)) indexées par la charge de travail W (équation 3.8) avec $d_0 = -d_1 = 4$

Impact de la charge de travail W sur le biais β et le critère de décision λ

Si on considère que le biais β du classificateur humain est indépendant de la charge de travail W comme le suggère (Putri *et al.*, 2016), alors de l'équation (3.2) on déduit

$$\lambda(W)d(W) - \frac{1}{2}d^2(W) = \ln \beta \quad (3.9)$$

D'où, en supposant $d(W) \neq 0$,

$$P_{FP}(W) = \Phi \left(- \left(\frac{\ln \beta}{d(W)} + \frac{d(W)}{2} \right) \right), \quad (3.10)$$

avec $d(W)$ donné par l'équation (3.8). Le point (P_{VP}, P_{FP}) décrivant les performances d'un opérateur se déplacerait alors sur les courbes isobiais en fonction de sa charge de travail (figure 3.5). Quand la charge de travail devient grande, l'opérateur tendrait vers l'un des deux comportements suivants : soit il sélectionne toujours l'hypothèse H_1 (si $\ln \beta < 0$), soit il sélectionne toujours l'hypothèse nulle H_0 (si $\ln \beta > 0$). Si son biais est exactement égal à 1 ($\ln(\beta) = 0$), alors à chaque classification, il aura autant de chance de prendre la bonne décision que de se tromper.

Supposons que l'opérateur ait appris une règle de classification pendant une formation. Ainsi on peut penser que sa règle de classification reste identique quelque soit sa charge de travail.

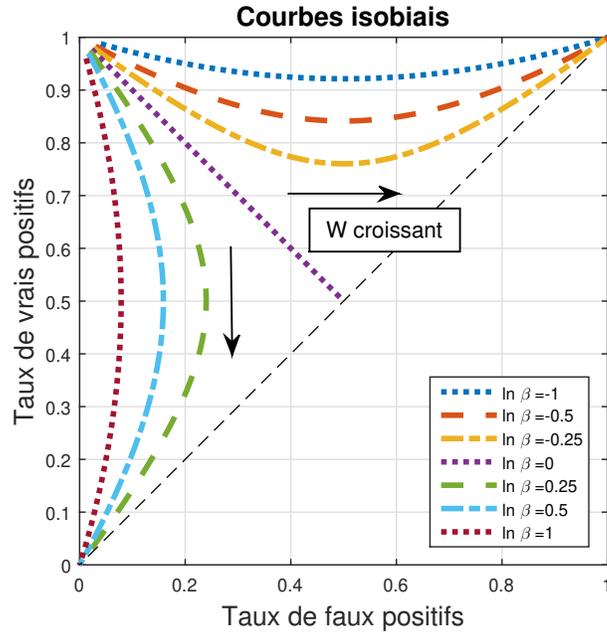


Figure 3.5 Courbes de performance humaine isobiais

En revanche sa capacité à appliquer cette règle, elle, est susceptible d'être amoindrie quand l'opérateur est surchargé. Du point de vue du modèle GVE, cela signifie que le critère de décision λ est indépendant de W . Supposer λ constant est cependant incompatible avec supposer que le biais β est constant.

On choisit finalement de supposer que la sensibilité d varie avec la charge de travail selon (3.8) et que le critère λ reste constant. Étant donné que le taux de faux positifs ne dépend que de λ ($P_{FP} = \Phi(-\lambda)$), cela revient à supposer que le taux de faux positifs est fixé. Le point (P_{VP}, P_{FP}) décrivant les performances d'un opérateur se déplacerait alors sur les courbes isocritère en fonction de sa charge de travail (figure 3.6).

3.2 Calibration du modèle

Dans cette section, on souhaite aborder la question de l'identification du modèle liant performance et charge de travail défini dans la section précédente. Pour ce faire, on commence par décrire une expérience théorique permettant de récolter les données nécessaires à la calibration du modèle. Ensuite, à partir de données expérimentales fictives, on montre comment les paramètres du modèle peuvent être identifiés facilement.

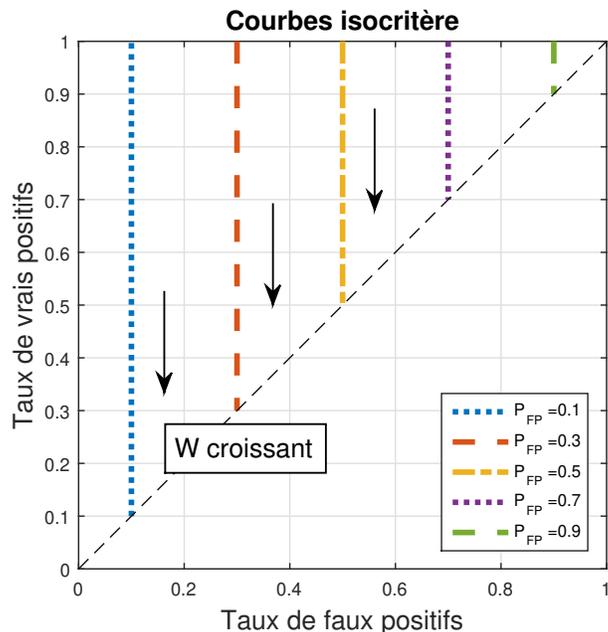


Figure 3.6 Courbes de performance humaine isocritère tracées avec le modèle GVE (3.5) complété par la dépendance à la charge de travail W donnée par (3.8)

3.2.1 Récolte des données

Dans un double objectif de calibrer les paramètres du modèle de performance humaine et d'en évaluer la pertinence, on cherche à récolter des données expérimentales. Le modèle que l'on cherche à évaluer a une entrée, la charge de travail W , et deux sorties, la performance de classification caractérisée par P_{VP}^h et P_{FP}^h . Les paramètres à identifier sont au nombre de deux : d_0 et d_1 .

Le but de l'expérience est de mesurer les performances de l'opérateur tout en faisant varier sa charge de travail. Cependant, dans le but de tester la validité du modèle, il est nécessaire d'introduire de la variabilité dans le comportement de l'opérateur. On cherche en particulier, pour une charge de travail donnée, à parcourir le plus possible la courbe ROC de l'opérateur. Pour cela il est nécessaire de l'inciter à changer son critère de classification tout en gardant la sensibilité constante. Plusieurs méthodes sont possibles pour expérimentalement faire varier le critère de décision d'un humain, entre autre (Wickens, 2002, Chapitre 3),

1. faire varier la fréquence d'apparition de l'évènement intéressant ($p = P(H_1)$) : si un évènement intéressant est très rare, l'humain aura tendance à baisser son critère et donc augmenter son taux de fausses alarmes. À l'inverse, si l'évènement semble très fréquent, le taux de fausses alarmes devrait être plus faible.

2. faire varier la perception des coûts relatifs de chaque réponse (R^{VP} , R^{FP} , R^{VN} et R^{FN}).

Les participants non experts seront davantage sensibles à la première méthode alors que les opérateurs plus entraînés peuvent plus facilement altérer leur critère de décision suite à une demande plus explicite.

3.2.2 Identification des paramètres

Supposons qu'une expérience a permis de récolter les points tracés sur la figure 3.7.

On rappelle que le modèle à identifier est le suivant :

$$P_{VP} = \Phi(d_0 + d_1W + \Phi^{-1}(P_{FP})). \quad (3.11)$$

Pour simplifier l'identification des paramètres, on représente les points expérimentaux dans le plan gaussien, c'est-à-dire qu'on applique la transformation $Z_x := \Phi^{-1}(x)$ (figure 3.8). On note qu'il faut retirer les points égaux à 0 et 1 car Φ^{-1} est défini sur $(0, 1)$. L'équation (3.11) devient alors

$$Z_{P_{VP}} = d_0 + d_1W + Z_{P_{FP}}. \quad (3.12)$$

Supposons qu'on dispose de $n > 0$ mesures, indexées par i , sous la forme $(Z_{P_{VP},i}, Z_{P_{FP},i}, W_i)$. Supposons de plus que ces mesures proviennent d'un même modèle entaché d'un bruit :

$$\begin{bmatrix} Z_{P_{VP},1} - Z_{P_{FP},1} \\ \vdots \\ Z_{P_{VP},n} - Z_{P_{FP},n} \end{bmatrix} = \begin{bmatrix} 1 & W_1 \\ \vdots & \vdots \\ 1 & W_n \end{bmatrix} \begin{bmatrix} d_0 \\ d_1 \end{bmatrix} + \boldsymbol{\omega}, \quad (3.13)$$

où $\boldsymbol{\omega}$ est une erreur aléatoire de moyenne 0 et d'écart-type σ . Sur la figure 3.7, on a $\mathbf{d} = [4, -4]^T$ et $\sigma = 0.3$.

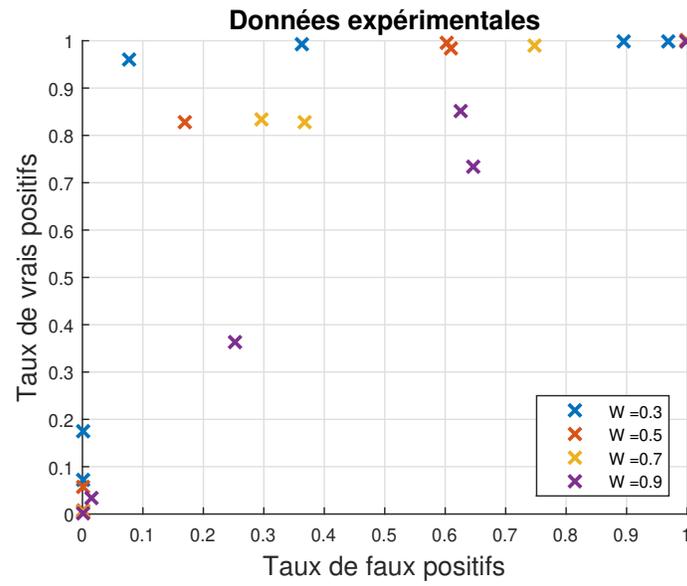


Figure 3.7 Données expérimentales fictives

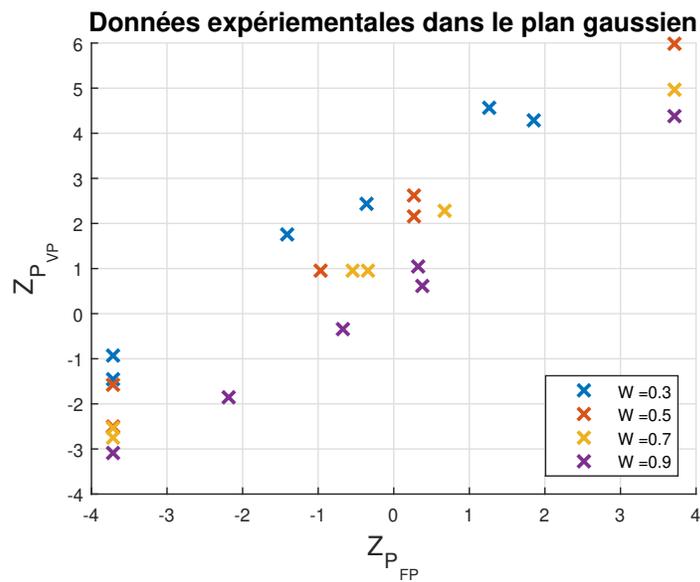


Figure 3.8 Données expérimentales fictives dans le plan gaussien

En notant

$$\mathbf{y} = \begin{bmatrix} Z_{PVP,1} - Z_{PFP,1} \\ \vdots \\ Z_{PVP,n} - Z_{PFP,n} \end{bmatrix},$$

$$\mathbf{H} = \begin{bmatrix} 1 & W_1 \\ \vdots & \vdots \\ 1 & W_n \end{bmatrix},$$

$$\mathbf{d} = \begin{bmatrix} d_0 \\ d_1 \end{bmatrix},$$

on formule le problème d'estimation suivant

$$\mathbf{y} = \mathbf{H}\mathbf{d} + \boldsymbol{\omega}. \quad (3.14)$$

Ce problème peut être résolu par moindres carrés avec en particulier,

$$\hat{\mathbf{d}} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{y}. \quad (3.15)$$

On obtient avec le jeu de données fictives $\hat{\mathbf{d}} = [3.09, -3.78]^T$ (voir figure 3.9).

Remarque. La méthode d'identification présentée ici est simple et certaines précautions devraient être prises si on devait l'appliquer à un vrai jeu de données. Par exemple, la transformation Z_x (3.12) peut être très sensible aux bruits de mesure et introduire des erreurs au moment de l'identification des paramètres.

3.3 Une solution statique au problème de partage de tâches

3.3.1 Le problème d'optimisation statique

Maintenant que le modèle liant charge de travail et performances humaines est défini, on formule un problème de partage de tâches exploitant ce modèle. Dans ce chapitre, on est intéressé par une stratégie de partage de tâches statique, c'est-à-dire que le partage de tâches ne change pas en fonction des périodes de travail.

Soit $p := P(H_1)$ ($p \neq 0$) la probabilité que l'objet ou l'évènement soit intéressant (hypothèse H_1). On suppose que l'opérateur réalise une proportion W de tâche pendant une période de travail. La répartition des résultats de classification observée pendant cette période est détaillée dans le tableau 3.1.

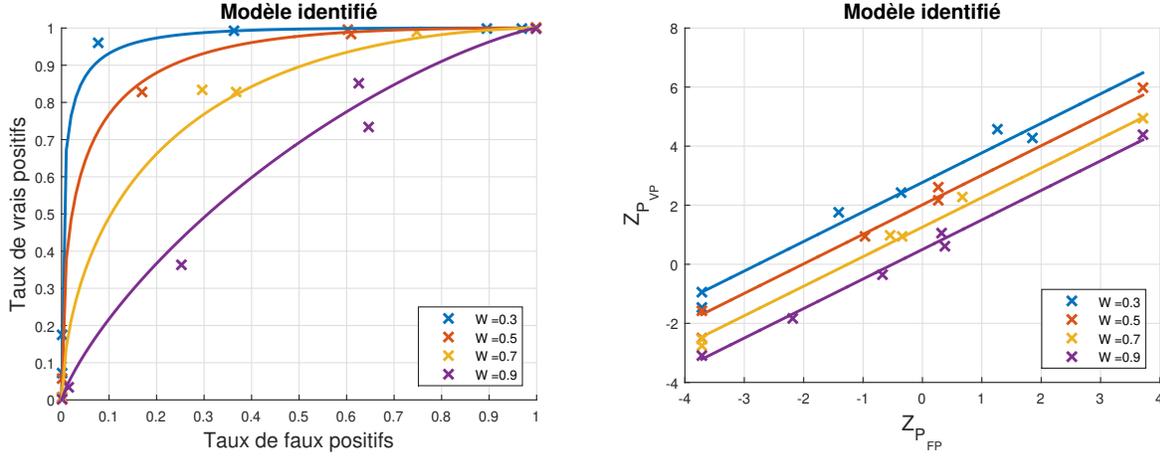


Figure 3.9 Modèle de performance identifié

Étant donné les récompenses $(R^{VP}, R^{FP}, R^{FN}, R^{VN}, R^h)$ associées à chaque issue d'une classification (section 1.2), on définit la fonction de récompense f suivante

$$\begin{aligned}
 f(W) = & \\
 & (1 - W) \left[p \left(P_{VP}^m R^{VP} + (1 - P_{VP}^m) R^{FN} \right) + (1 - p) \left(P_{FP}^m R^{FP} + (1 - P_{FP}^m) R^{VN} \right) \right] + \\
 & W \left[p \left(P_{VP}^h(W) R^{VP} + (1 - P_{VP}^h(W)) R^{FN} \right) + (1 - p) \left(P_{FP}^h(W) R^{FP} + (1 - P_{FP}^h(W)) R^{VN} \right) \right] + \\
 & W R^h
 \end{aligned} \tag{3.16}$$

La stratégie de partage de tâches statique proposée dans ce chapitre consiste à suggérer à l'opérateur d'accomplir une proportion $a_t = W^*$ de classification quelque soit la période de travail t . La charge de travail optimale W^* est le résultat de la maximisation de la fonction de récompense instantanée f :

$$\boxed{W^* = \underset{W \in [0,1]}{\operatorname{argmax}} f(W).} \tag{3.17}$$

Tableau 3.1. Résultats d'une période de travail.

Classificateur	Classe prédite	Classe réelle		Total
		Positif	Négatif	
Machine	Positif	$(1 - W)pP_{VP}^m$	$(1 - W)(1 - p)P_{FP}^m$	$(1 - W)$
	Négatif	$(1 - W)p(1 - P_{VP}^m)$	$(1 - W)(1 - p)(1 - P_{FP}^m)$	
Humain	Positif	$WpP_{VP}^h(W)$	$W(1 - p)P_{FP}^h(W)$	W
	Négatif	$Wp(1 - P_{VP}^h(W))$	$W(1 - p)(1 - P_{FP}^h(W))$	
	Total	p	$(1 - p)$	1

3.3.2 Résultats du problème d'optimisation

Nous nous intéressons ici à la résolution du problème de maximisation statique (3.17). On note avant tout, en développant (3.17), que le problème est équivalent à

$$W^* = \underset{W \in [0,1]}{\operatorname{argmax}} \tilde{f}(W) \quad (3.18)$$

où

$$\tilde{f}(W) = W \left[p \left(P_{VP}^h(W) - P_{VP}^m \right) \left(R^{VP} - R^{FN} \right) + \right. \\ \left. (1 - p) \left(P_{FP}^h(W) - P_{FP}^m \right) \left(R^{FP} - R^{VN} \right) + R^h \right]. \quad (3.19)$$

Intérêt d'une équipe humain-machine

Un premier résultat donne la condition suivant laquelle l'opérateur humain est bénéfique à l'équipe.

Théorème 1. Supposons que P_{VP}^h et P_{FP}^h sont des fonctions continues de W . Alors \tilde{f} atteint son maximum W^* dans $[0, 1]$.

Supposons de plus que P_{VP}^h et P_{FP}^h sont différentiables en $W = 0$ et $W = 1$. Si

$$p(R^{VP} - R^{FN})(P_{VP}^h(0) - P_{VP}^m) + (1-p)(R^{FP} - R^{VN})(P_{FP}^h(0) - P_{FP}^m) > |R^h|, \quad (3.20)$$

alors le maximum W^* est dans $]0, 1]$. Si, de plus,

$$p(R^{VP} - R^{FN})(P_{VP}^h(1) - P_{VP}^m) + (1-p)(R^{FP} - R^{VN})(P_{FP}^h(1) - P_{FP}^m) < |R^h|, \quad (3.21)$$

alors le maximum W^* est dans $]0, 1[$.

Remarque. Puisqu'on a supposé que l'humain a un taux de faux positifs indépendant de sa charge de travail, les conditions (3.20) et (3.21) deviennent respectivement :

$$\begin{aligned} (P_{VP}^h(0) - P_{VP}^m) &> C, \\ (P_{VP}^h(1) - P_{VP}^m) &< C, \end{aligned} \quad (3.22)$$

où

$$C := \frac{|R^h| - (1-p)(R^{FP} - R^{VN})(P_{FP}^h - P_{FP}^m)}{p(R^{VP} - R^{FN})}. \quad (3.23)$$

Ce résultat montre que former une équipe humain-machine devient utile lorsque les performances de l'opérateur lorsqu'il est peu chargé (quand $W = 0$) sont suffisamment grandes devant celles de l'algorithme et que ses performances lorsqu'il est très chargé (quand $W = 1$) sont suffisamment petites devant celles de la machine.

Démonstration. Dans le but de simplifier les notations, on réécrit (3.19) :

$$\tilde{f}(W) = \alpha W P_{VP}^h(W) + \delta W P_{FP}^h(W) - \gamma W, \quad (3.24)$$

avec

$$\begin{aligned} \alpha &:= p(R^{VP} - R^{FN}), \\ \delta &:= (1-p)(R^{FP} - R^{VN}), \\ \gamma &:= p(R^{VP} - R^{FN})P_{VP}^m + (1-p)(R^{FP} - R^{VN})P_{FP}^m - R^h. \end{aligned}$$

La fonction \tilde{f} est continue sur un intervalle fermé réel donc atteint son maximum dans $[0, 1]$. De plus,

$$\tilde{f}'(W) = \alpha P_{VP}^h(W) + \delta P_{FP}^h(W) + W(\alpha P_{VP}^{h'}(W) + \delta P_{FP}^{h'}(W)) - \gamma.$$

Or il suffit d'avoir $\tilde{f}'(0) > 0$ pour s'assurer que le maximum W^* n'est pas zéro, i.e., il suffit

d'avoir

$$\alpha P_{VP}^h(0) + \delta P_{FP}^h(0) > \gamma,$$

ce qui correspond à l'hypothèse (3.20). De même, il suffit d'avoir $\tilde{f}'(1) < 0$ pour s'assurer que le maximum W^* n'est pas 1, i.e., il suffit d'avoir

$$\alpha P_{VP}^h(1) + \delta P_{FP}^h(1) < \gamma,$$

ce qui correspond à l'hypothèse (3.21). □

Résolution numérique du problème

On s'intéresse ici à la résolution numérique du problème (3.17) lorsque qu'on utilise le modèle de performance décrit plus haut. Le résultat suivant montre que le problème de recherche de maximum peut être résolu facilement par dichotomie ou par résolution de l'équation $\tilde{f}'(W) = 0$. La figure 3.10 montre un exemple de résolution où $P_{FP}^h = P_{FP}^m = 0.1$, $d_0 = -d_1 = 4$ et $P_{VP}^m = \Phi(1.5 + \Phi^{-1}(P_{FP}^m))$. On obtient alors $W^* \approx 0.38$.

Théorème 2. Supposons que $P_{VP}^h(W)$ soit donné par (3.5) et (3.8) et que P_{FP}^h soit indépendant de W . Sous les conditions du théorème 1, \tilde{f} a un unique maximum global W^* dans $]0,1[$, et W^* est l'unique solution de l'équation $\tilde{f}'(W) = 0$.

Démonstration. Étant donné que $P_{FP}^{h'}(W) = 0$, l'expression de \tilde{f}' devient :

$$\tilde{f}'(W) = \alpha P_{VP}^h(W) + \alpha W P_{VP}^{h'}(W) - \gamma + \delta P_{FP}^h, \quad (3.25)$$

où $P_{VP}^h(W) = \Phi(d_1 W + K)$, en notant $K := d_0 + \Phi^{-1}(P_{FP}^h)$.

On cherche à montrer que \tilde{f}' change de signe qu'une seule fois sur $]0,1[$.

On sait du théorème 1 que $\tilde{f}'(0) > 0$ et $\tilde{f}'(1) < 0$. De plus, \tilde{f}' est dérivable et

$$\begin{aligned} \tilde{f}''(W) &= \alpha \left(2P_{VP}^{h'}(W) + W P_{VP}^{h''}(W) \right), \\ &= -\frac{d_1 \alpha}{\sqrt{2\pi}} \exp\left(\frac{-(d_1 W + K)^2}{2}\right) (d_1^2 W^2 + d_1 K W - 2), \\ &:= g(W) (d_1^2 W^2 + d_1 K W - 2), \end{aligned} \quad (3.26)$$

où $g(W)$ est strictement positif (si $d_1 < 0$). Ainsi \tilde{f}'' a le même signe que le polynôme $l(W) := (d_1^2 W^2 + d_1 K W - 2)$. Ce polynôme possède deux racines réelles W_1 et W_2 ($W_1 < W_2$) car son discriminant $\Delta = d_1^2(K^2 + 8)$ est positif. Le coefficient d_1^2 est positif donc le polynôme

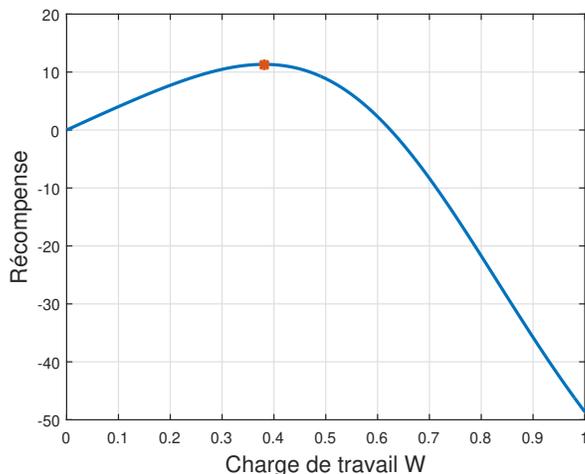


Figure 3.10 Maximisation de \tilde{f} (3.19)

est d'abord positif, puis négatif puis de nouveau positif. Or $l(0) = -2 < 0$ donc la première racine W_1 est négative. En outre,

- si $W_2 \geq 1$, alors \tilde{f}'' est toujours négative, donc \tilde{f}' est décroissante,
- si $W_2 < 1$, alors \tilde{f}' est décroissante puis croissante sur $[0, 1]$.

Dans les deux cas, étant donné que $\tilde{f}'(0) > 0$ et $\tilde{f}'(1) < 0$, \tilde{f}' ne peut passer qu'une fois par 0 sur $]0, 1[$. W^* est donc l'unique solution de $\tilde{f}'(W) = 0$.

□

3.3.3 Conclusion

Dans ce chapitre, un modèle liant, à chaque période de travail, charge de travail W et performance humaine a été introduit. Ce modèle a permis de proposer une première solution au problème de partage de tâches dans l'équipe humain-machine. Cette stratégie simple est la solution d'un problème statique de maximisation de récompense instantanée. On peut espérer de cette stratégie qu'elle trouve un équilibre entre charge de travail et performance humaine. Cependant le problème formulé pour aboutir au partage de tâches statique repose sur une hypothèse importante : l'opérateur suit toujours les recommandations de la machine. Autrement dit on suppose que jamais il ne choisira de réaliser plus ou moins de classifications que suggéré par la machine. Or il a été mis en évidence (voir chapitre 2) que la confiance entre humain et machine pouvait avoir un impact sur le choix d'un opérateur de réaliser une tâche manuellement plutôt que de la déléguer à une machine. Cette remarque conduit à explorer de nouvelles stratégies de partage de tâches.

CHAPITRE 4 SOLUTION ADAPTATIVE

On a déjà mis en évidence la nécessité de prendre en compte le comportement de dépendance à la machine de l'opérateur dans le design d'une stratégie de partage de tâches. Pour ce faire, on a décidé de s'intéresser à l'un des principaux facteurs de la dépendance : la confiance que place un humain dans les capacités d'une machine. La nature dynamique de la confiance justifie la nécessité d'adopter une stratégie de collaboration adaptative, i.e., qui s'adapterait à l'état changeant de la confiance humaine. Dans ce chapitre, on propose et teste une telle solution.

En premier lieu, on présente dans la section 4.1 le modèle dynamique de la confiance qui est utilisé dans la conception de la stratégie d'allocation de tâches présentée dans la section 4.2. Ensuite, dans la section 4.3, on cherche à prendre en compte la nature non observable de la confiance pour proposer une stratégie plus réaliste. Le reste du chapitre tente d'explorer en simulation les bénéfices et les limitations de la solution avancée.

4.1 Un modèle dynamique de la confiance

4.1.1 Définition et dynamique de la confiance

Le modèle présenté ci-dessous est une simplification du modèle proposé par Gao et Lee dans (Gao et Lee, 2006b). Dans cet article, les auteurs proposent un modèle dynamique et quantitatif capable de prédire à chaque instant si un opérateur va préférer le mode manuel au mode automatique pour réaliser une tâche de supervision de processus industriel. La dépendance de l'humain à la machine est vu comme la différence entre la confiance que place l'humain en les capacités de la machine et celle qu'il place en ses propres capacités.

Plusieurs simplifications ont été faites pour aboutir au modèle présenté dans ce mémoire :

1. On suppose que la confiance en soi de l'opérateur reste constante au cours du temps. Cette hypothèse peut être valable si l'opérateur est expérimenté. Ceci étant dit, il est tout à fait possible d'étendre le modèle proposé en prenant en compte la confiance en soi.
2. On suppose que les performances de classification de la machines sont observables par l'opérateur à chaque fin de périodes de travail.
3. Dans (Gao et Lee, 2006b), la prédiction du modèle est binaire (contrôle manuel ou automatique) alors que nous cherchons plutôt à prédire le pourcentage de tâche que

l'opérateur va compléter lui-même. Ainsi les modèles de dépendance seront un peu différents.

Définition de l'état humain

L'état de confiance de l'opérateur au début de la période t est représenté par un couple de réel

$$x_t^h = (T_t, B_t)^T, \quad (4.1)$$

où T_t est appelé confiance dans la machine et B_t est l'estimation faite par l'opérateur des capacités de la machine.

Dynamique de l'état

La dynamique de l'état x_t^h a pour entrées l'état décrivant l'environnement à la période t , noté x_t^e , ainsi que la suggestion de partage de tâches de la machine a_t . La dynamique est donnée par

$$\begin{aligned} B_{t+1} &= B_t + \eta \left(C(x_t^e, x_t^h, a_t) - B_t \right) + w_t^B, \\ T_{t+1} &= (1 - \mu)T_t + \mu B_{t+1} + w_t^T, \\ &= (1 - \mu)T_t + \mu \left(B_t + \eta \left(C(x_t^e, x_t^h, a_t) - B_t \right) + w_t^B \right) + w_t^T. \end{aligned} \quad (4.2)$$

où

w_t^B et w_t^T sont des variables aléatoires de moyenne nulle et d'écart type, respectivement, σ_B et σ_T ,

$\eta \in [0, 1]$ représente la **transparence** de l'interface. Plus η est grand, plus l'opérateur a une bonne estimation des capacités que la machine vient de démontrer,

$\mu \in [0, 1]$ représente l'**inertie** de la confiance. Plus μ est proche de 1, plus l'opérateur donne de l'importance aux interactions récentes par rapport aux interactions plus anciennes,

$C(x_t^e, x_t^h, a_t) \in \mathbb{R}$ est la fonction de **capacités réelles** de la machine. Cette fonction représente les performances de la machine démontrées si la machine propose un partage de tâches a_t alors que l'état de l'environnement est x_t^e et celui de l'humain, x_t^h .

La définition de la fonction représentant les capacités réelles de la machine vu comme l'entrée du modèle de la dynamique de la confiance dépend du scénario étudié. Il peut s'agir de la précision avec laquelle la machine exécute sa tâche, d'une rapidité d'exécution, d'un taux d'erreur ou encore d'un taux de succès. Dans le cas du scénario décrit dans la section 1.2, il

est possible que l'humain prenne en compte non seulement les performances de classification de la machine mais également sa capacité à suggérer un bon partage de tâches, comme c'est le cas avec la définition suivante :

$$C(x_t^h, a_t) = (1 - W_t)P_s^m + W_tP_s^h(W_t). \quad (4.3)$$

P_s^m et P_s^h sont les pourcentages de succès des classifications réalisées respectivement par la machine et l'humain. Ils sont définis par :

$$\begin{aligned} P_s^m &= pP_{VP}^m + (1 - p)(1 - P_{FP}^m) \\ P_s^h &= pP_{VP}^h(W_t) + (1 - p)(1 - P_{FP}^h) \end{aligned} \quad (4.4)$$

où $P_{VP}^h(W_t)$ est donné par le modèle de performance humaine décrit dans le chapitre précédent (section 3.1). La fonction (4.3) représente les performances de classification de l'équipe humain-machine au complet mais les poids relatifs à certaines erreurs (R^{VP} , R^{VN} , R^{FP} , R^{FN} et R^h) ne sont pas pris en compte. Cette définition sera utilisée dans le reste du mémoire. D'autres candidats seront introduits dans la section 4.6.3.

La définition de C peut être délicate car, même si le système est complexe, elle requiert de réduire les performances de la machine à une variable unidimensionnelle. Une étude préliminaire peut être nécessaire pour identifier le facteur qui représente le mieux les capacités instantanées de la machine vues par un opérateur. On pourrait même imaginer que cette fonction soit un peu différente suivant les opérateurs.

4.1.2 Définition de la dépendance à la machine

On cherche désormais à prédire le comportement de dépendance de l'opérateur étant donné son état de confiance x_t^h . Autrement dit on veut prédire le pourcentage de tâches que l'humain va réaliser pendant une période de travail, c'est-à-dire W_t , autrement appelé la charge de travail dans ce mémoire.

On choisit le modèle non linéaire suivant :

$$W_t(x_t^h, a_t) = \frac{W^{\bar{T}}(a_t)}{W^{\bar{T}}(a_t) + (1 - W^{\bar{T}}(a_t))e^{\nu(T_t - \bar{T})}}, \quad (4.5)$$

avec

$$W^{\bar{T}}(a_t) = sat(a_t), \quad (4.6)$$

où la fonction saturation est définie par $sat(x) = \max(0 + \epsilon, \min(1 - \epsilon, x))$.

On compte trois paramètres dans ce modèle :

$\nu \geq 0$ représente la vitesse à laquelle la charge de travail imposée par l'opérateur lui-même, va grandir au fur et à mesure que sa confiance diminue.

$\bar{T} \in \mathbb{R}$ est la valeur de la confiance à laquelle l'opérateur va respecter la suggestion de la machine si celle-ci est dans $a_t \in [0 + \epsilon, 1 - \epsilon]$.

$\epsilon \in]0, 1]$, lorsque a_t est au-delà de l'intervalle $[0 + \epsilon, 1 - \epsilon]$, on estime que l'opérateur ne va pas respecter une suggestion trop extrême (trop petite ou trop grande).

Ce modèle, illustré à la figure 4.1, traduit bien le fait que l'opérateur va davantage utiliser la machine (i.e., W_t est faible) que sa confiance T_t est grande. De plus cette relation est non linéaire comme le suggèrent les données expérimentales dans (Gao et Lee, 2006b; Lee et Moray, 1994). Le modèle présenté a la particularité d'avoir introduit l'impact de la suggestion de partage de tâches a_t dans la charge de travail W_t .

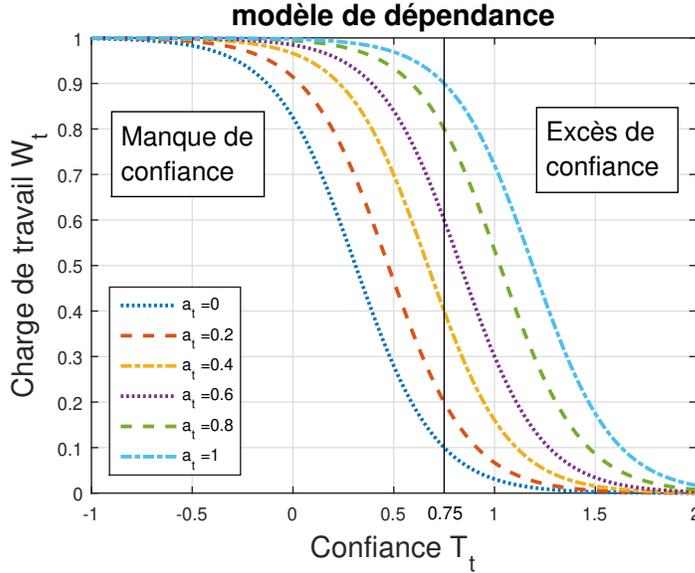


Figure 4.1 Modèle de dépendance de l'humain à la machine avec $\nu = 5$, $\bar{T} = 0.75$ et $\epsilon = 0.1$

4.1.3 Remarques sur le modèle de la confiance

Le modèle proposé pour prédire le comportement de dépendance de l'opérateur à la machine nécessite de garder en mémoire un état bidimensionnel $x_t^h = (T_t, B_t)^T$. La connaissance de cet état et celle de la suggestion de partage de tâches de la machine a_t permettent de prédire quelle sera la charge de travail que s'imposera l'opérateur pendant la période de travail t (équation (4.5)). De plus, à partir de x_t^h et a_t nous sommes capable de prédire l'évolution de l'état humain, i.e., x_{t+1}^h (équation (4.2)).

Les paramètres du modèle présenté sont au nombre de 5 seulement. De plus chacun de ces paramètres a une interprétation. En revanche il est probable que ce modèle ne soit pas simple à identifier dans des scénarios complexes (par exemple la définition de la fonction C). En outre, la valeur de certains paramètres peut dépendre de l'opérateur et de sa personnalité. En effet, dans (Gao et Lee, 2006b), certains mauvais scores de prédiction sont expliqués par les différences de comportement individuel. Dans ce mémoire on ne propose pas de méthode pour identifier les paramètres du modèle de la confiance.

4.2 Formulation du problème en MDP

On rappelle qu'on cherche à concevoir une stratégie de partage de tâches qui s'adapterait à l'état de confiance de l'opérateur x_t^h . On peut s'aider pour cela du modèle quantitatif de la dynamique de la confiance identifié dans la section précédente ainsi que du modèle liant charge de travail et performance humaine. En concevant cette stratégie, on cherche à maximiser les performances de l'équipe humain-machine sur le long terme.

Ce problème de partage de tâches adaptatif peut en fait se formuler sous la forme d'un processus de décision markovien (ou *Markov Decision Process* (MDP)).

4.2.1 Description

L'état du MDP comprend uniquement l'état de l'opérateur humain $x_t^h = (T_t, B_t)^T$. L'espace d'état est alors $\mathbf{S} = \mathbb{R}^2$. Le modèle de transition de l'état est donné par les équations (4.2).

Remarque. Si le scénario est plus complexe que celui présenté dans la section 1.2, on pourrait augmenter l'état du MDP avec une description de l'environnement si celui-ci doit être gardé en mémoire pour décrire les performances de l'équipe. On devrait alors obtenir un modèle séquentiel markovien de cet état additionnel. L'état du MDP deviendrait alors $s_t = (x_t^h, x_t^e)^T$.

L'action du MDP a_t correspond à la suggestion de partage de tâches formulée par la machine à chaque début de période de travail. a_t est la proportion de tâche que l'opérateur devrait accomplir selon la machine. L'espace d'action est alors $\mathbf{A} = [0, 1]$.

La fonction de récompense $R(x_t^h, a_t)$ immédiate est récoltée en fin de période t . Elle est

donnée par

$$\begin{aligned}
R(x_t^h, a_t) = & \\
& (1 - W_t) \left[p \left(P_{VP}^m R^{VP} + (1 - P_{VP}^m) R^{FN} \right) + (1 - p) \left(P_{FP}^m R^{FP} + (1 - P_{FP}^m) R^{VN} \right) \right] + \\
& W_t \left[p \left(P_{VP}^h(W_t) R^{VP} + (1 - P_{VP}^h(W_t)) R^{FN} \right) + (1 - p) \left(P_{FP}^h R^{FP} + (1 - P_{FP}^h) R^{VN} \right) \right] + \\
& W_t R^h,
\end{aligned} \tag{4.7}$$

conformément à la description du scénario. La charge de travail W_t dépend de x_t^h et a_t selon l'équation (4.5).

Dans un premier temps, on va supposer que l'état du MDP x_t^h est complètement observable mais avec une période de retard. En effet, on peut observer de manière certaine quelle a été la charge de travail W_t que s'est imposé l'opérateur à la fin de la période t . Sachant la suggestion de la machine a_t , on peut déduire T_t car la fonction $T_t \mapsto W_t((T_t, B_t)^T, a_t)$ est inversible (équation (4.5)). Par ailleurs, on peut imaginer que B_t soit observé directement à partir de questionnaires à la fin de la période t . Ce retard d'une période dans l'observation de l'état implique que la machine va suggérer un nouveau partage de tâches au début de la période t qui dépendra de l'état x_{t-1}^h , et pas de l'état x_t^h .

On peut maintenant formuler le problème sous la forme suivante. On cherche une politique, c'est-à-dire des fonctions π_t qui à l'état x_{t-1}^h associe un partage de tâches a_t . On souhaite trouver la politique qui maximise la somme amortie des récompenses espérées sur un horizon infini, i.e.,

$$\max_{\pi=(\pi_0, \pi_1, \dots)} \mathbb{E} \left[\sum_{t=1}^{\infty} \alpha^t R(x_t^h, \pi_t(x_{t-1}^h)) \right], \tag{4.8}$$

où $\alpha \in [0, 1[$ est le facteur d'atténuation.

4.2.2 Implémentation et résolution numérique

Résolution numérique du problème

Le package, codé en Julia, nommé *POMDPs.jl* (Egorov *et al.*, 2017) disponible ici <https://github.com/JuliaPOMDP/POMDPs.jl>, a été utilisé pour la résolution numérique du problème d'optimisation. Ce package permet de définir un MDP ou un POMDP et réalise l'interface avec divers algorithmes de résolution.

Parmi ces algorithmes, on a fait le choix d'utiliser le solveur *SARSOP* (Kurniawati *et al.*,

2008) dont l’interface en Julia est disponible ici <https://github.com/JuliaPOMDP/SARSOP.jl>. Cet algorithme permet de calculer une politique qui approxime la politique optimale d’un POMDP ou d’un MDP dont les espaces d’états peuvent être relativement grands. Il s’agit d’un solveur hors ligne dont les espaces d’état et d’action sont discrets. En conséquence ces espaces sont discrétisés de la façon suivante : $\mathbf{S} = (-1, -0.95, \dots, 2)^2$ et $\mathbf{A} = (0, 0.05, \dots, 1)$. Ainsi on a $Card(\mathbf{S}) = 3721$ et $Card(\mathbf{A}) = 21$.

Dans son calcul, l’algorithme *SARSOP* maintient une borne supérieure \bar{V} et une borne inférieure \underline{V} de la fonction de valeur optimale V^* . Plus l’algorithme réalise d’itérations, plus \bar{V} et \underline{V} se rapprochent de V^* . La précision d’une politique approximative est définie comme la différence entre \bar{V} et \underline{V} . Lorsqu’on résoudra un MDP ou un POMDP, l’algorithme arrêtera sa recherche lorsque cette précision deviendra inférieure à 1.

Valeurs des paramètres

Dans la suite, sauf mention contraire, les valeurs des paramètres utilisées sont récapitulées dans le tableau 4.1.

Tableau 4.1. Valeurs numériques des paramètres des différents modèles.

Modèle de performance	$P_{FP}^m = 0.1, P_{FP}^h = 0.1, P_{VP}^m = \Phi(1.5 + \Phi^{-1}(P_{FP}^m)), d_0 = -d_1 = 4$
Modèle de confiance	$\mu = 0.5, \eta = 0.5, \sigma_T = 0.1, \sigma_B = 0.1$
Modèle de dépendance	$\nu = 5, \bar{T} = 0.8, \epsilon = 0.1$
Fonction de récompense	$p = 0.5, R^{VP} = R^{VN} = 100, R^{FP} = R^{FN} = -100, R^h = 0, \alpha = 0.95$

Justifications des valeurs numériques des paramètres

Les valeurs des paramètres du modèle de performance ont été choisies de sorte à ce que l’humain soit plus performant que la machine lorsqu’il a peu de charge de travail et moins performant lorsqu’il est surchargé. Pour illustration, la figure 4.2 montre la relation entre charge de travail et taux de vrais positifs lorsqu’on utilise les valeurs du tableau 4.1.

Concernant le modèle de confiance, les paramètres η et μ sont choisis égaux à 0.5 car il s’agit de la moitié de leur intervalle de définition. Les écart-type des erreurs aléatoires sont relativement grands car on s’attend à ce que le modèle de confiance ne soit pas très précis.

Une discussion sur un paramètre du modèle de la dépendance est offerte dans la section 4.4.1.

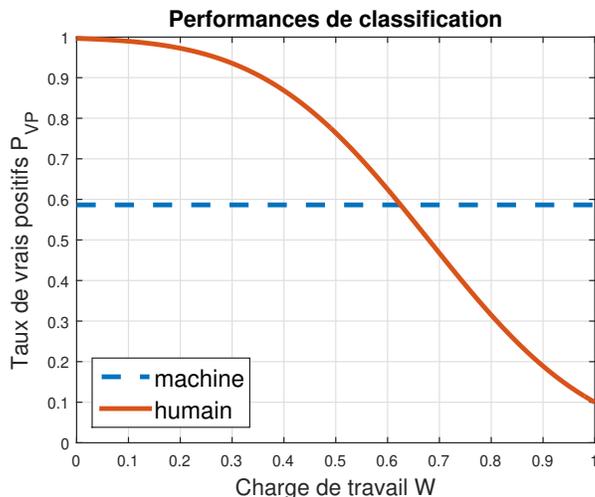


Figure 4.2 Modèle de performance de classification de l’humain et de la machine

Enfin, on a choisi, pour le modèle de récompense, de donner la même importance respectivement aux succès (vrais positifs et vrais négatifs) et aux erreurs (faux positifs et faux négatifs) de classification. De plus on n’ajoute pas de coût artificiel au travail de l’humain ($R^h = 0$). Ainsi les pénalités engendrées par le travail humain ne passent que par une augmentation de charge de travail.

4.2.3 Résultats de simulation

Application de la politique adaptative en simulation

Une fois la politique approximativement optimale calculée, on souhaite la voir en action en simulant des interactions entre machine et humain selon notre scénario.

On simule ainsi la collaboration de l’humain et de la machine par 50 périodes de travail successives. Le comportement de l’humain est simulé par le modèle décrit dans la section 4.1, c’est-à-dire qu’on tire au hasard, selon la distribution de probabilité dictée par le modèle, la transition de l’état x_t^h à l’état x_{t+1}^h à chaque période.

La figure 4.3 montre un exemple de 50 périodes de travail lorsque la machine suggère à chaque période un partage de tâches donné par la politique calculée en résolvant le MDP. L’état initial de l’opérateur est $x_0^h = (0, 0)^T$, c’est-à-dire qu’il a très peu confiance en les capacités de la machine. Sur la figure de gauche, la machine suggère, à chaque début de période, la proportion d’objet que devrait accomplir l’opérateur (pointillés noirs). Cette suggestion est dictée par la politique solution du MDP. Pendant la période, l’opérateur choisit d’accomplir plus ou moins de tâche que suggéré (trait magenta). Sur la figure de droite sont tracés l’évolution de

l'état de l'opérateur (trait plein bleu pour T_t et tirets oranges pour B_t) ainsi que les capacités réelles de la machine C en pointillés verts.

Un point important à retenir ici est que la politique optimale selon la résolution du MDP formule une suggestion de partage de tâches différente en fonction de l'état de l'opérateur. Autrement dit, la politique qui maximise les récompenses de l'équipe au long terme est bien adaptative.

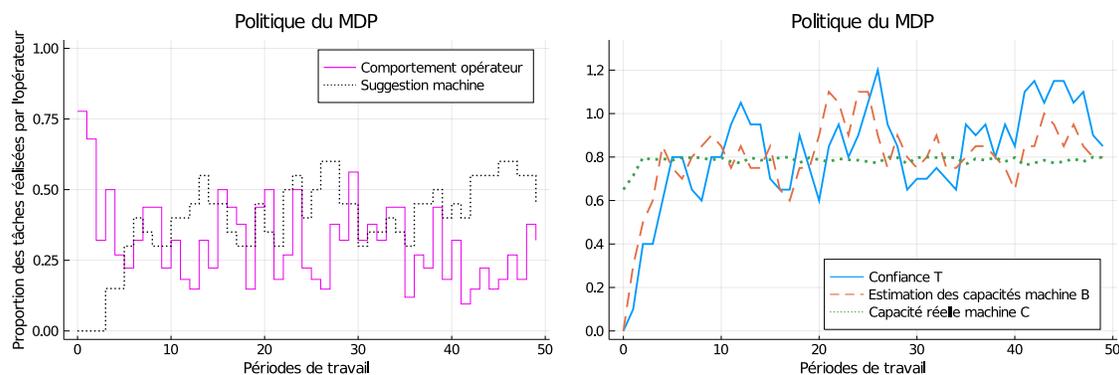


Figure 4.3 Simulation appliquant la politique du MDP

Remarques sur le choix du nombre N de classifications à réaliser pendant une période de travail

On a déjà évoqué en introduction (section 1.2) certaines précautions à prendre lorsqu'on choisit le nombre N de classifications que doit réaliser l'équipe pendant une période de travail. Si N est trop petit, alors, en pratique, les capacités réelles de la machine et de l'humain seraient susceptibles de beaucoup varier d'une période de travail à une autre car les taux de succès seraient calculés à partir d'un faible nombre de classifications. On peut alors imaginer qu'on observerait des grandes variations sur la figure 4.3, remettant potentiellement en cause la pertinence du partage de tâche adaptatif. À l'inverse, si N est trop grand, on perdrait la capacité du partage de tâches adaptatif à réagir lorsque la confiance chute ou bondit.

Comparaison des politiques adaptative et statique

On souhaite maintenant voir en quoi la stratégie adaptative est davantage bénéfique pour l'équipe humain-machine. La figure 4.4 montre une simulation de 50 périodes de travail lorsque la suggestion de partage de tâches est la même pour chaque période, ici $a_t = 0.3808$. Ce partage statique des tâches est calculé selon la méthode décrite dans le chapitre 3. Il ne

prend pas en compte la dynamique de la confiance de l'opérateur mais seulement sa charge de travail. L'état initial de l'opérateur est aussi ici $x_0^h = (0, 0)^T$.

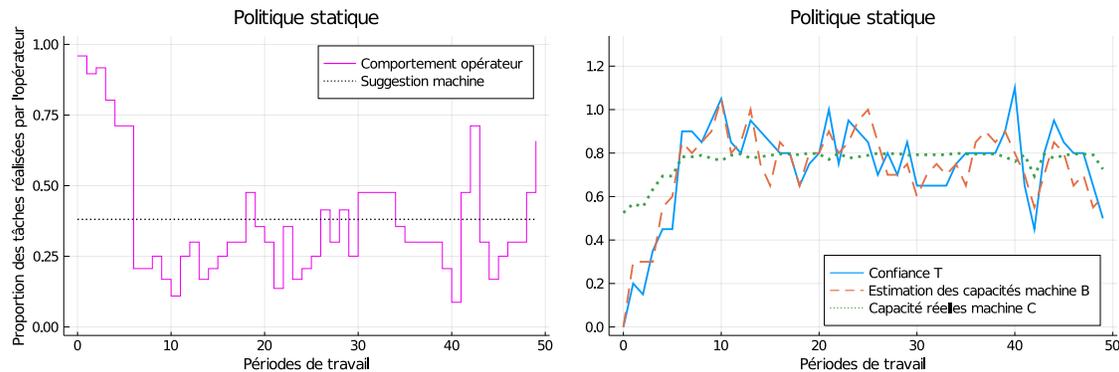


Figure 4.4 Simulation appliquant une politique statique

Lorsqu'on regarde les récompenses récoltées pendant les simulations précédentes, on remarque qu'en effet la stratégie adaptative du MDP semble apporter un bénéfice par rapport à la stratégie statique (figure 4.5). On précise que pour que la comparaison soit juste, le calcul des récompenses obtenues est le même pour les deux stratégies (on utilise la somme atténuée des récompenses instantanées données par (4.7)).

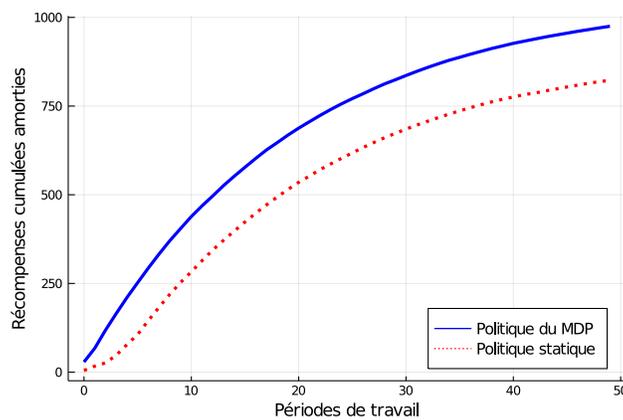


Figure 4.5 Récompenses cumulées amorties récoltées pendant les simulations des figures 4.3 et 4.4

Bien que cette simulation semble montrer un avantage certain à l'utilisation de la stratégie issue de la résolution du MDP, une hypothèse pourrait remettre en question ces résultats préliminaires. Cette hypothèse est le caractère observable de l'état de confiance x_t^h de l'opérateur.

4.3 Formulation du problème en POMDP

On est parvenu à concevoir une stratégie de partage de tâches adaptative, qui à l'état de confiance x_t^h associe la suggestion d'allocation de tâche. Jusque là, on a supposé que l'état de l'humain était complètement observable par la machine, avec une période de retard. L'observation d'un état cognitif tel que la confiance, peut se faire à travers des observations comportementales, des questionnaires ou des mesure physiologiques. La moins invasive et la plus transparente pour l'opérateur de ces méthodes est sans doute l'observation comportementale. Cependant de telles observations ne permettent pas de connaître avec certitude l'état interne humain.

Le cadre des MDP dispose d'une extension permettant d'inclure les situations où l'état n'est pas complètement observé : les processus de décisions markoviens partiellement observables (ou *Partially Observed Markov Decision Process* (POMDP)). Le but de cette section est de proposer une stratégie adaptative d'allocation de tâches qui ne s'appuierait que sur des observations comportementales de l'état de confiance de l'opérateur en formulant le problème comme un POMDP.

4.3.1 Description

Les états ainsi que les actions du POMDP sont les même que ceux du MDP, voir la section 4.2. La fonction de transition reste également la même (équation (4.2)), tout comme la fonction de récompense (équation (4.7)).

Observations

À chaque fin de période de travail, la machine fait une observation de l'opérateur. Dans le cas d'un MDP, cette observation est directement l'état x_t^h , mais dans celui d'un POMDP, on observe o_t . Puisqu'on s'est fixé comme objectif de n'utiliser que des observations comportementales de l'opérateur, o_t est défini comme étant la proportion de tâches qu'a effectivement réalisé l'humain pendant la période t . L'espace des observations est donc $\mathbf{O} = [0, 1]$.

Dans un POMDP il faut également définir la fonction d'observation, qui donne la relation entre une observation o_t et l'état réel x_t^h . Cette fonction est donnée par

$$o_t(x_t^h, a_t^a) = \frac{W^{\bar{T}}(a_t)}{W^{\bar{T}}(a_t) + (1 - W^{\bar{T}}(a_t))e^{\nu((T_t + w_t^o) - \bar{T})}}, \quad (4.9)$$

où w_t^o est une variable aléatoire de moyenne nulle et d'écart type σ_o . $W^{\bar{T}}(a_t)$ est défini par

(4.6). Ce modèle des observations est cohérent avec le modèle estimant la charge de travail de l'opérateur, utilisé dans le modèle de la dynamique de l'état (équation (4.5)). On aura pour autant ajouté un bruit aléatoire w_t^o , représentant les erreurs de modélisation du comportement humain.

État de croyance

Dans un POMDP, la machine ne connaît pas parfaitement l'état x_t^h de l'opérateur à chaque instant mais dispose seulement d'une distribution de probabilité sur l'espace d'état, appelé état de croyance (ou *belief state*), noté b_t , pour choisir son action a_t . On a $b_t \in \Pi(\mathbf{S})$, où $\Pi(X)$ désigne l'ensemble des distributions de probabilité sur un ensemble X .

À chaque fois que la machine fait une observation o_t de l'opérateur après avoir pris l'action a_t , l'état de croyance b_t est mis à jour de sorte à ce que la probabilité que l'opérateur ait évolué dans l'état $x^{h'} \in \mathbf{S}$ soit

$$\begin{aligned}
 b_{t+1}(x^{h'}) &= P(x^{h'} | o_t, a_t, b_t) \\
 &= \sum_{x^h \in \mathbf{S}} P(x^{h'} | x^h, o_t, a_t, b_t) P(x^h | o_t, a_t, b_t), \\
 &= \sum_{x^h \in \mathbf{S}} P(x^{h'} | x^h, a_t) \frac{P(o_t | x^h, a_t, b_t) P(x^h | a_t, b_t)}{P(o_t | a_t, b_t)}, \quad \text{avec le théorème de Bayes} \\
 &= \frac{1}{P(o_t | a_t, b_t)} \sum_{x^h \in \mathbf{S}} P(x^{h'} | x^h, a_t) P(o_t | x^h, a_t) b_t(x^h),
 \end{aligned} \tag{4.10}$$

où

- $P(x^{h'} | x^h, a_t)$ est la probabilité que l'état de l'opérateur devienne $x^{h'}$ à la période $t+1$ sachant qu'à la période t son état était x^h et que la machine a pris la décision a_t ,
- $P(o_t | x^h, a_t)$ est la probabilité d'observer o_t à la fin de la période t sachant que l'état au début de la période est x^h et l'action prise par la machine est a_t ,
- $b_t(x^h)$ est la probabilité que l'état de l'opérateur à la période t soit x^h selon l'état de croyance b_t .

Remarque. On note que le dénominateur $P(o_t | a_t, b_t)$ dans (4.10) peut être vu comme un facteur de normalisation permettant en tout temps t d'avoir $\sum_{x^h \in \mathbf{S}} b_t(x^h) = 1$.

Ainsi, la machine garde un état de croyance aussi proche possible de l'état réel de l'opérateur.

Lorsque l'opérateur et la machine se rencontrent pour la première fois, on peut aider la machine à avoir une idée de l'état de confiance de l'opérateur, i.e., initialiser son état de

croissance. À défaut d'indices, la machine peut définir b_0 comme une distribution uniforme sur l'ensemble des états possibles de l'opérateur.

Désormais, la résolution du POMDP doit permettre de trouver la politique, i.e., les fonctions π_t qui à chaque état de croyance b_t associent une suggestion de partage de tâches a_t . Cette politique devra maximiser l'espérance de la somme amortie des récompenses sur un horizon infini :

$$\max_{\pi=(\pi_0,\pi_1\dots)} \mathbb{E} \left[\sum_{t=1}^{\infty} \alpha^t R(x_t^h, \pi_t(b_t)) \right], \quad (4.11)$$

où $\alpha \in [0, 1[$ est le facteur d'atténuation.

Boucle de simulation

La simulation d'une succession de périodes de travail est décrite par l'algorithme 1. La fonction *observation* : $\mathbf{S} \times \mathbf{A} \rightarrow \Pi(\mathbf{O})$ associe à un état de l'opérateur x_t^h et une suggestion d'allocation de tâche a_t , la distribution de probabilité de l'observation o_t de la machine. Cette distribution est calculée à partir de (4.9). La fonction *transition* : $\mathbf{S} \times \mathbf{A} \rightarrow \Pi(\mathbf{S})$ associe à un état de l'opérateur x_t^h et une suggestion d'allocation de tâche a_t , la distribution de probabilité du nouvel état x_{t+1}^h de l'opérateur. Cette distribution est calculée à partir de (4.2). La fonction *rand*(Y) où $Y \in \Pi(X)$ est une distribution de probabilité sur l'espace X , tire au hasard un élément de l'espace X selon la distribution de probabilité Y . Enfin la fonction *update* : $\Pi(\mathbf{S}) \times \mathbf{O} \times \mathbf{A} \rightarrow \Pi(\mathbf{S})$ donne la distribution de probabilité correspondant au nouvel état de croyance b_{t+1} connaissant l'état de croyance précédent b_t , l'observation o_t et l'action de la machine a_t . Cette fonction est donnée par (4.10).

À chaque nouvelle période de travail, la machine suggère le partage de tâches correspondant à son état de croyance actuel b selon la politique π . Ensuite la période de travail est réalisée et des récompenses sont collectées. Pendant ce temps, la machine observe le comportement de l'opérateur, c'est-à-dire la portion de tâche qu'il va décider d'accomplir lui-même. La machine va alors mettre à jour son état de croyance en prenant en compte cette dernière observation ainsi que son état de croyance précédent. À la fin de la période, l'état réel de l'opérateur a évolué en fonction de l'interaction qu'il vient d'avoir avec la machine. Une nouvelle période de travail peut alors débuter.

4.3.2 Implémentation et résolution numérique

Les outils utilisés pour résoudre ce POMDP sont les mêmes que pour le MDP (voir la section 4.2.2).

Algorithme 1 : Simulation de n périodes de travail lorsque la machine suggère un partage de tâches donné par la politique π en se basant sur l'observation du comportement de l'opérateur.

```

1  $x^h \leftarrow x_0^h$  // état réel (caché pour la machine)
2  $r \leftarrow 0$  // récompenses cumulées amorties
3  $b \leftarrow b_0$  // initialisation de l'état de croyance
4 pour  $t \leftarrow 1$  à  $n$  faire // on souhaite simuler  $n$  périodes de travail
5    $a \leftarrow \pi(b)$  // la machine choisit un partage de tâches à suggérer selon la
   politique  $\pi$ 
6    $r \leftarrow r + \alpha^t R(x^h, a)$  // une récompense est collectée
7    $o \leftarrow \text{rand}(\text{observation}(x^h, a))$  // la machine observe le comportement de
   l'opérateur pendant la période
8    $b \leftarrow \text{update}(b, o, a)$  // la machine met à jour son état de croyance en prenant en
   compte sa dernière observation
9    $x^h \leftarrow \text{rand}(\text{transition}(x^h, a))$  // l'état réel de l'opérateur évolue
10 retourner  $r$ 

```

Tout comme les espaces d'état et d'action, il a été nécessaire de discrétiser l'espace des observations. On a alors $\mathbf{O} = (0, 0.05, \dots, 1)$ et $\text{Card}(\mathbf{O}) = 21$.

Les valeurs numériques des paramètres des modèles sont données, sauf mention contraire, dans le tableau 4.1. Un paramètre supplémentaire est nécessaire pour le modèle d'observation, l'écart type de la variable aléatoire w_i^o . Dans la suite on choisira $\sigma_0 = 0.05$.

4.3.3 Résultats de simulation

Après calcul de la stratégie approximativement optimale, on décide de réaliser une simulation de 50 périodes de travail (voir figure 4.6). Encore une fois, on remarque la stratégie optimisant la récompense à long terme est bien une stratégie adaptative.

Dans cette simulation en particulier, la politique du POMDP a récolté une récompense totale supérieure à celle de la politique statique (figure 4.7). Ce résultat préliminaire nous encourage à chercher à savoir dans quels cas on obtient un tel bénéfice. On aimerait également explorer les limites de cette approche. Les prochaines sections s'attellent à ces questions en proposant divers expériences de simulation.

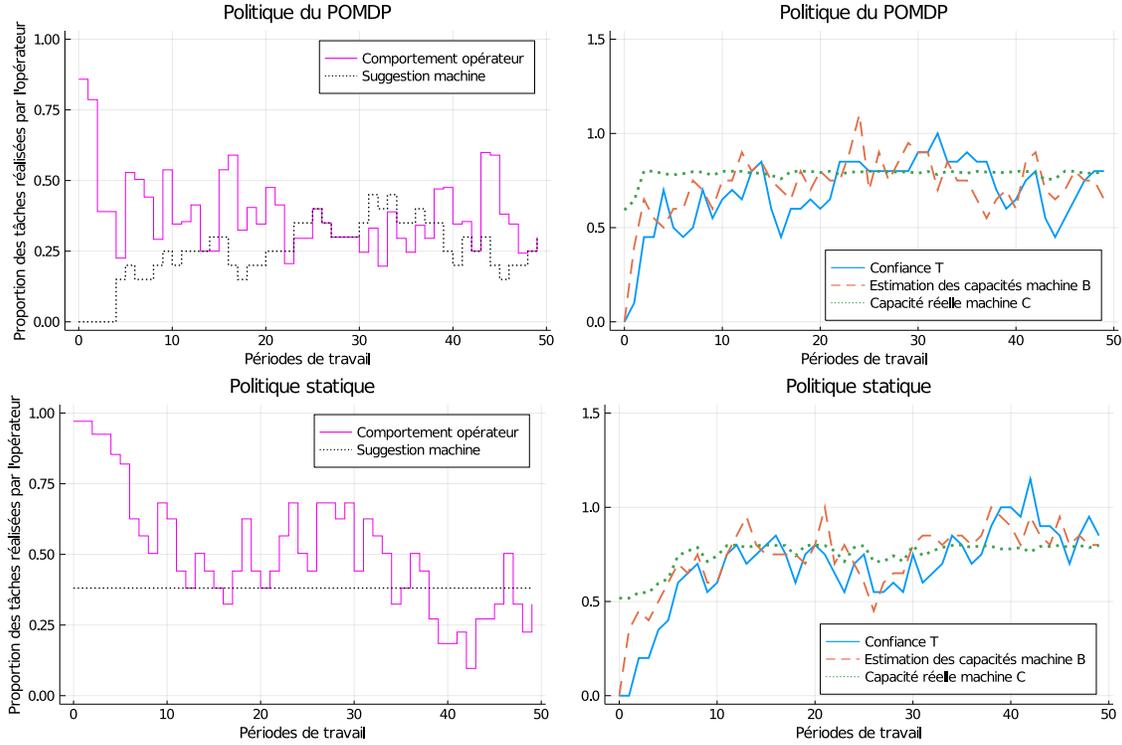


Figure 4.6 Simulations appliquant la politique du POMDP (figures du haut) et la politique statique (figures du bas)

4.4 Expérience 1 : Modèles parfaits

Dans cette section, on cherche à identifier les bénéfices de la stratégie du POMDP. Pour cela, on commence par étudier le comportement de cette politique lorsque les décisions de l'opérateur sont déterministes. Autrement dit, on retire la part d'aléatoire dans les modèles de comportement humain (il n'y a plus de variables aléatoires w_t^B , w_t^T et w_t^o dans, respectivement, (4.2) et (4.9)).

4.4.1 Comportement en régime permanent

La figure 4.8 compare les comportements des stratégies statique et dynamique lorsque le comportement de l'opérateur est déterministe. L'état initial est, pour les deux simulations, $x_0^h = (1.5, 1.5)^T$. On constate qu'en régime permanent les deux politiques ont un comportement proche. La politique issue de la résolution du POMDP est constante avec $a_\infty = 0.40$ contre $a = 0.38$ pour la politique statique. On a des récompenses totales similaires pour les deux politiques : 1027 pour la politique du POMDP et 1022 pour la politique statique.

Si la politique statique est en effet optimale en régime permanent, on devrait avoir $a_\infty = W^*$,

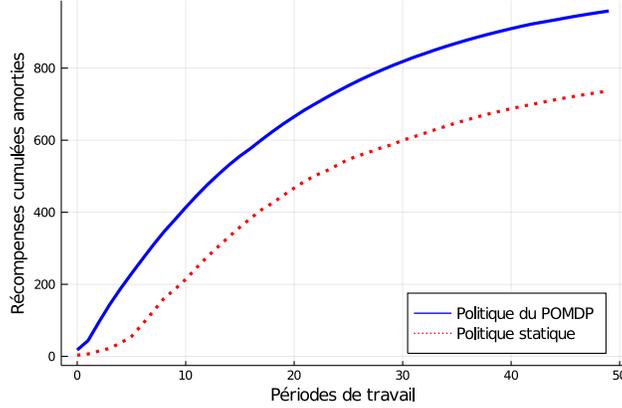


Figure 4.7 Récompenses cumulées amorties récoltées pendant deux simulations, l'une appliquant la stratégie du POMDP, l'autre appliquant la stratégie statique

où W^* est la charge de travail optimale lorsque le problème est statique. De plus il faudrait que l'opérateur réalise toujours la proportion de tâche suggérée par la machine, i.e., $W_\infty = a_\infty$. Selon (4.5), c'est le cas lorsque $a_\infty \in [0.1, 0.9]$ et $\bar{T} = T_\infty$. Or, on note qu'en régime permanent, on doit avoir

$$\begin{aligned} B_\infty &= C_\infty, \\ T_\infty &= B_\infty. \end{aligned} \tag{4.12}$$

Il faut donc avoir $\bar{T} = C_\infty$, pour que la stratégie statique soit optimale. Or $C_\infty = (1 - W_\infty)P_s^m + W_\infty P_s^h(W_\infty)$. D'où la stratégie statique est optimale si $W^* \in [0.1, 0.9]$ et $\bar{T} = C^* = (1 - W^*)P_s^m + W^* P_s^h(W^*)$. Il se trouve qu'ici on a choisi $\bar{T} = 0.8$ alors que $C^* = 0.8$. C'est pour cela qu'en régime permanent les deux stratégies sont équivalentes. La différence entre $a_\infty = 0.4$ et $W^* = 0.38$ s'explique par la discrétisation de l'espace des actions.

On peut voir que, sur la figure 4.9, lorsque $\bar{T} \neq C^*$, alors l'opérateur ne va pas suivre exactement la recommandation de la machine en régime permanent. La stratégie issue du POMDP prend en compte cela et adapte sa suggestion pour que toujours la charge de travail de l'opérateur soit le plus proche possible de la charge optimale, ici $W_\infty = 0.36$, contrairement à la stratégie statique. La politique du POMDP obtient une récompense totale d'environ 1023 alors que celle de la politique statique descend à 956.

Cette observation peut s'interpréter de la façon suivante. Si le jugement de l'opérateur n'est pas biaisé, c'est-à-dire si il suit les recommandations de la machine lorsque celle-ci démontre des capacités maximales ($\bar{T} = C^*$), alors la stratégie statique est optimale en régime permanent. La stratégie issue du POMDP rajoute une constante à la suggestion de tâche par

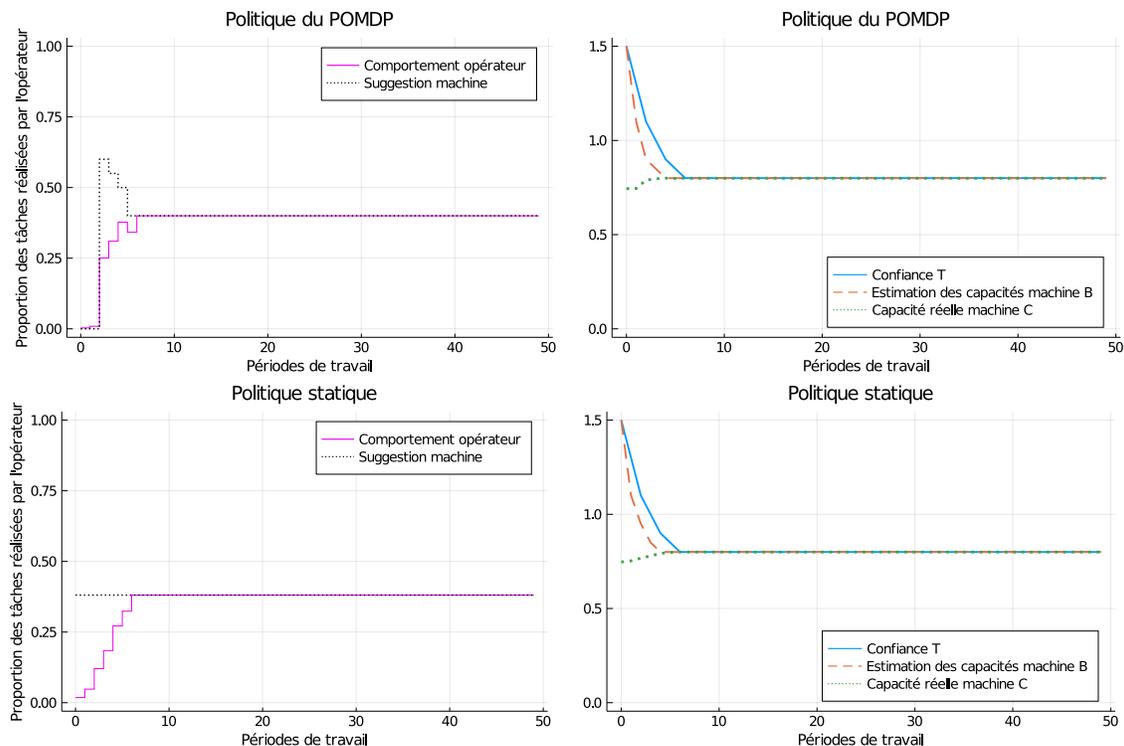


Figure 4.8 Simulations du régime permanent avec $\bar{T} = C^* = 0.8$

rapport à la stratégie statique pour prendre en compte le biais quand il existe.

4.4.2 Comportement en régime transitoire

Comparons le régime transitoire des deux stratégies sur les figures 4.8, 4.9, 4.10 et 4.11. On s'intéresse à la rapidité avec laquelle la charge de travail de l'opérateur s'approche de sa valeur finale. Appelons Tr la période à partir de laquelle W_t rentre dans l'intervalle $[0.85W_\infty, 1.15W_\infty]$ et n'en sort plus.

Tableau 4.2. Comparaison de la durée du régime transitoire lorsque le comportement de l'opérateur est déterministe.

	Figure 4.8	Figure 4.9	Figure 4.10	Figure 4.11
Politique du POMDP	$Tr = 4$	$Tr = 3$	$Tr = 3$	$Tr = 2$
Politique statique	$Tr = 5$	$Tr = 4$	$Tr = 8$	$Tr = 5$

On compare la durée des régimes transitoires pour les deux stratégies dans le tableau 4.2. On

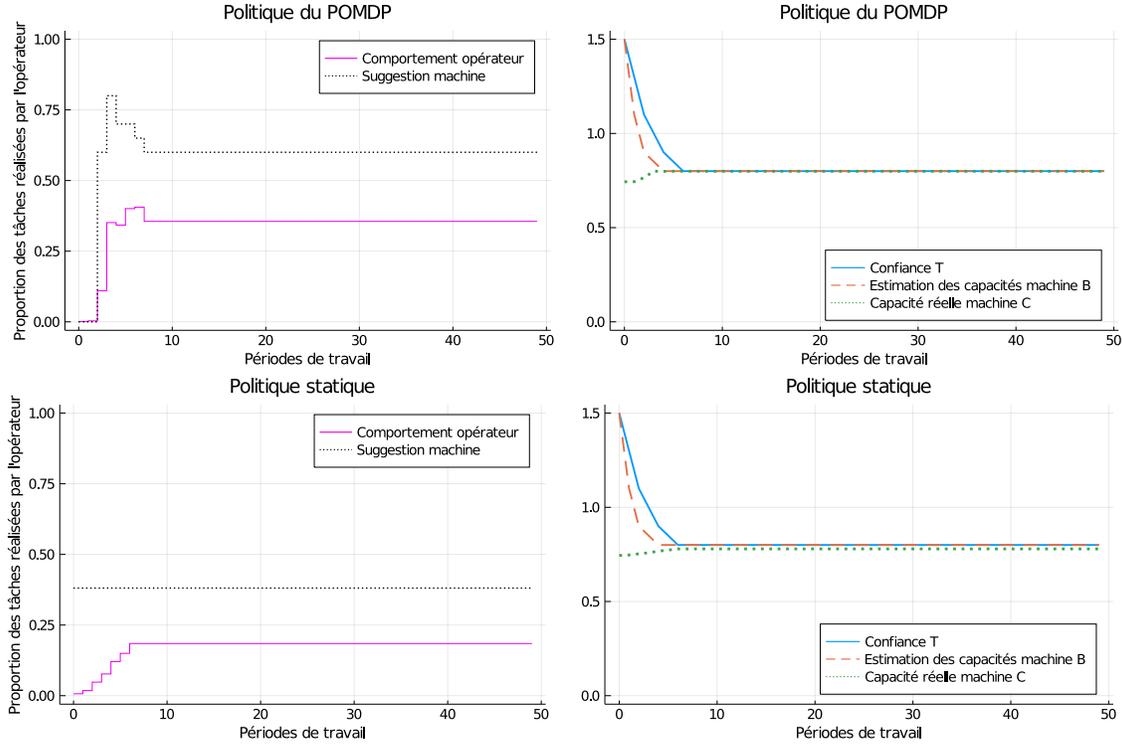


Figure 4.9 Simulations du régime permanent avec $\bar{T} = 0.6 \neq C^*$

remarque que bien que les conditions soient différentes (valeur du paramètre \bar{T} et état initial x_0^h), la charge de travail se rapproche systématiquement plus vite de sa valeur finale lorsque la politique du POMDP est appliquée. De plus, dans le cas des figures 4.8 et 4.10, les régimes permanents sont identiques donc la différence des récompenses accumulées s'expliquent par un régime transitoire plus rapide.

4.4.3 Effet de l'observation partielle

Jusque là, l'état de croyance initial qui guide les premières suggestions de la machine, est le plus flou qui puisse être puisque qu'il s'agit d'une distribution uniforme sur l'ensemble des états possibles, i.e., pour tout $x^h \in \mathbf{S}$,

$$b_0(x^h) = \frac{1}{\text{Card}(\mathbf{S})}. \quad (4.13)$$

On peut se demander à quel point l'initialisation de cette distribution a un impact sur les premières périodes de travail. La figure 4.11 montre la même politique dans les mêmes conditions que la figure 4.8 mis à part que l'état de croyance initial est exact, i.e., pour tout

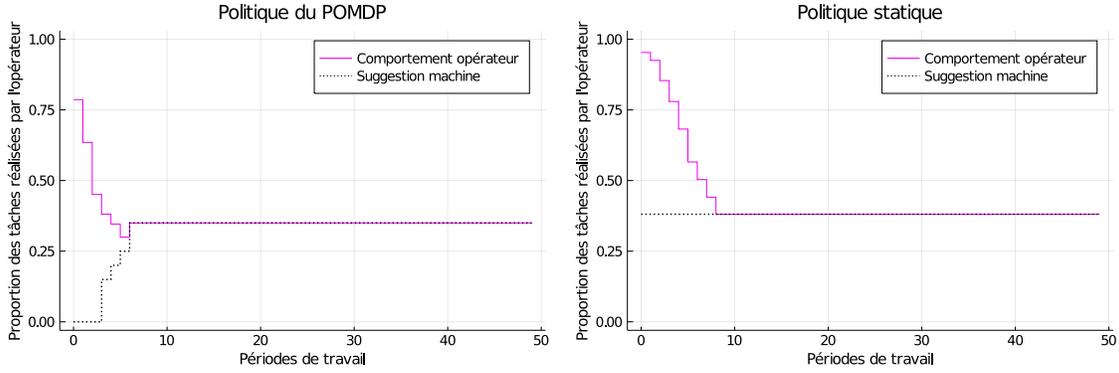


Figure 4.10 Simulations avec $\bar{T} = C^* = 0.8$ et $x_0^h = (0, 0)^T$

$x^h \in \mathbf{S}$,

$$\begin{cases} b_0(x^h) = 1 & \text{si } x^h = x_0^h, \\ b_0(x^h) = 0 & \text{sinon.} \end{cases} \quad (4.14)$$

On note que les premières suggestions de la machine sont alors différentes. On remarque également que la récompense totale obtenue monte à 1041 pour la stratégie du POMDP, contre 1027 lorsque l'état de croyance était une distribution uniforme. La stratégie du POMDP devient alors encore plus avantageuse que la stratégie statique.

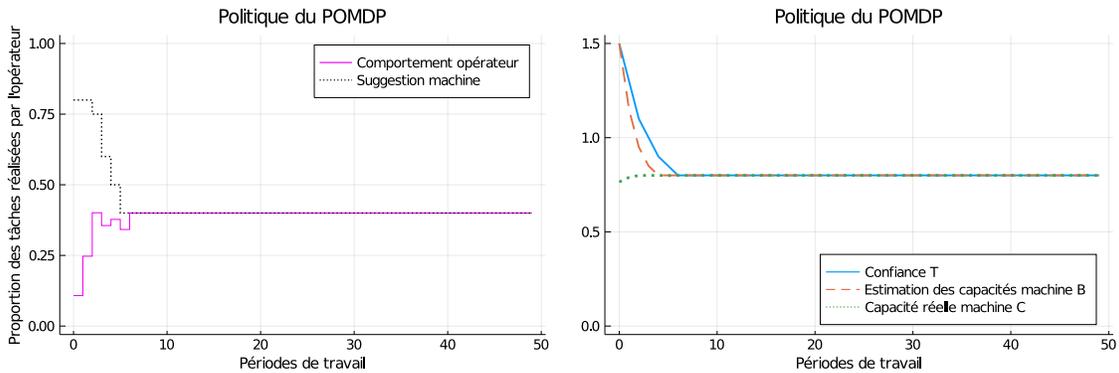


Figure 4.11 Simulations où l'état de croyance initial de la machine est exact

On conclut que la qualité de l'état de croyance initial peut affecter la durée du régime transitoire et ainsi les récompenses de l'équipe.

Pour conclure, on a vu dans cette section que, si le modèle de l'humain est parfait, la stratégie du POMDP n'a un intérêt que pendant le régime transitoire, à condition que l'opérateur ne soit pas biaisé dans son jugement (voir le paragraphe 4.4.1). On a vu que dans ce cas la stratégie du POMDP est avantageuse même en régime permanent. On note toutefois que les

différences de récompenses récoltées ne sont pas suffisamment importantes (moins de 10% de récompenses en plus) pour que l'approche par POMDP soit vraiment pertinente.

4.5 Expérience 2 : Modèles imparfaits

On vient de discuter des cas dans lesquels la politique issue de la résolution du POMDP a un bénéfice par rapport à la stratégie statique lorsque l'opérateur se comporte exactement comme les modèles de performance et de confiance le prédisent. Puisque les performances de la machine sont constantes, alors la répartition des tâches converge vers un état permanent. Dans cette situation, les différences de récompenses récoltées ne semblent pas suffisamment grandes pour valider l'approche par POMDP. Dans cette section, on va comparer les stratégies statiques et adaptatives lorsque l'opérateur ne se comporte pas toujours comme le modèle de la confiance le prédit. Cette hypothèse est beaucoup plus réaliste puisque ce modèle néglige naturellement un certain nombre de phénomènes difficilement modélisables. Désormais on stipule dans la définition du POMDP la partie aléatoire de la dynamique de la confiance.

4.5.1 Comparaison entre les politiques statique et adaptative

Maintenant que le comportement de l'opérateur a une part d'aléatoire, visualiser la politique en action sur seulement une simulation ne suffit pas à conclure sur l'avantage d'une stratégie sur une autre. Il faut pour cela, réaliser un très grand nombre de simulations.

La figure 4.12 compare les récompenses totales moyennes acquises pendant 5000 simulations appliquant la stratégie statique ou la stratégie calculée à partir du POMDP. On voit que la politique adaptative dispose d'un avantage certain sur la politique statique. Cet avantage s'exprime à la fois par une moyenne des récompenses plus haute (946 contre 763 pour la politique statique, soit près de 24% de récompenses supplémentaires) ainsi que par une plus faible dispersion (écart-type de 21 contre 68 pour la politique statique).

4.5.2 Effets du bruit sur le suivi de l'état

Dans un POMDP, une des difficultés est de construire un état de croyance suffisamment proche de l'état réel. Afin de visualiser la qualité de l'état de croyance, on définit la métrique suivante :

$$d^T(b_t, x_t^h) = \sqrt{\sum_{T \in [0, \dots, 1]} \sum_{B \in [0, \dots, 1]} b_t(x^h = (T, B))(T - T_t)^2} \quad (4.15)$$

$$d^B(b_t, x_t^h) = \sqrt{\sum_{B \in [0, \dots, 1]} \sum_{T \in [0, \dots, 1]} b_t(x^h = (T, B))(B - B_t)^2} \quad (4.16)$$

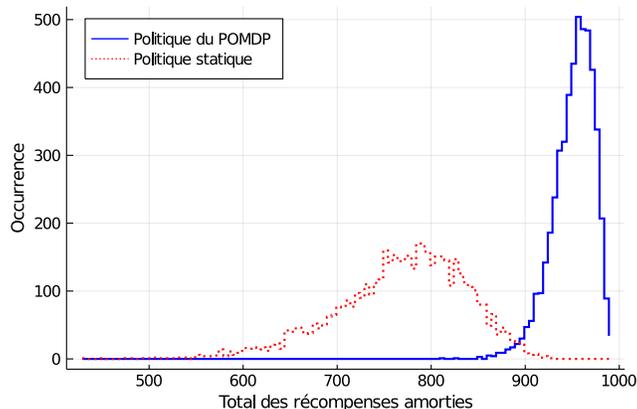


Figure 4.12 Récompenses totales amorties collectées dans 5000 simulations de 50 périodes de travail

Selon cette métrique, pour l'état de confiance T_t par exemple, la différence entre une confiance supposée T et la confiance réelle T_t compte d'autant plus que la probabilité de T selon l'état de croyance b_t est grande. Ainsi plus d^T et d^B sont grands plus la distribution de l'état de croyance est éloignée de l'état réel. Aussi l'état de croyance est exact $b_t(x_t^h) = 1$ si et seulement si $d^T(b_t, x_t^h) = 0$ et $d^B(b_t, x_t^h) = 0$.

La figure 4.13 montre l'évolution des erreurs de l'état de croyance pendant la simulation de la figure 4.6. Dans cette simulation l'état de croyance était initialisé par une distribution uniforme sur l'espace d'état, expliquant les fortes erreurs initiales. Pendant les premières périodes de travail, la machine construit un état de croyance plus précis. Elle n'atteint pour autant, jamais l'état de croyance parfait du fait de la part d'aléatoire dans les observations ainsi que dans la dynamique de x_t^h .

4.6 Expérience 3 : Erreurs d'identification des modèles

On a vu l'intérêt d'adopter une stratégie dynamique par le gain en performance à long terme que prédisent les simulations. La stratégie du POMDP repose essentiellement sur la validité des modèles prédisant le comportement de l'opérateur. Cette validité passe entre autre par la capacité à calibrer les paramètres de ces modèles. Cette étape peut poser un certain nombre de difficultés tant au niveau de la récolte de données expérimentales qu'à l'identification numérique des modèles. Dans cette section, on souhaite explorer jusqu'à quel point ces inévitables erreurs d'identification peuvent affecter les bénéfices de l'approche adaptative par rapport à l'approche statique. Pour cela on réalise des séquences de 5000 simulations de 50 périodes de travail avec $x_0^h = (0, 0)^T$. Dans certaines de ces séquences, on supposera

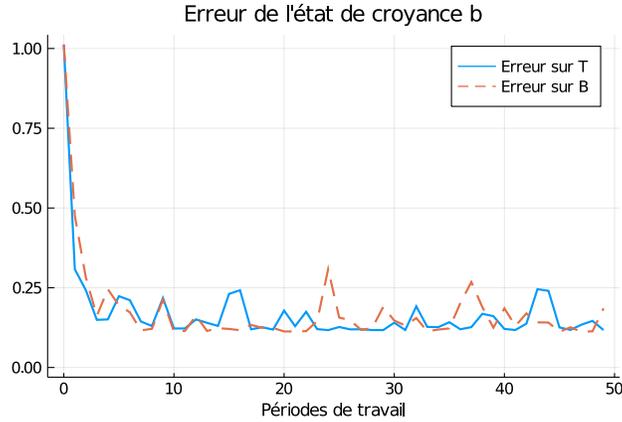


Figure 4.13 Erreur de suivi de l'état de croyance, telle que défini par (4.15) et (4.16)

que certains paramètres ont mal été identifiés, introduisant une différence entre le modèle utilisé dans le calcul de la politique optimale et celui utilisé pour simuler le comportement de l'opérateur.

4.6.1 Erreur constante dans le modèle de performance

On s'intéresse d'abord à la robustesse de la stratégie du POMDP face à des erreurs dans le modèle liant charge de travail et performances humaines. En particulier, on va supposer qu'en simulation la sensibilité de l'opérateur suit

$$d(W) = d_0 + (d_1 + kd_1)W, \quad (4.17)$$

au lieu de (3.8). Lorsque l'erreur k est négative, l'opérateur est plus performant que prévu, lorsque k est positive, il l'est moins (car $d_1 < 0$). Les courbes ROC utilisées en simulation seront donc un peu différentes des courbes ROC utilisées pour résoudre le POMDP.

La figure 4.14 montre les récompenses moyennes et leur écart-type calculées après 5000 simulations pour différentes valeurs de l'erreur kd_1 , où k est exprimé comme un pourcentage d'erreur sur d_1 . On note que la stratégie statique est aussi impactée par cette erreur de calibration car elle utilise aussi le modèle liant performance et charge de travail. Sur cette figure, et les trois prochaines, les traits indiquent un écart type de chaque côté des moyennes.

Les résultats montrent que la stratégie du POMDP garde un avantage sur la stratégie statique tant que k ne passe pas sous les -25% d'erreur. En deçà de cette limite les deux stratégies semblent donner des résultats très similaires. Autrement dit, si le modèle sur estime trop l'impact négatif de la charge de travail sur les performances humaines, alors la stratégie

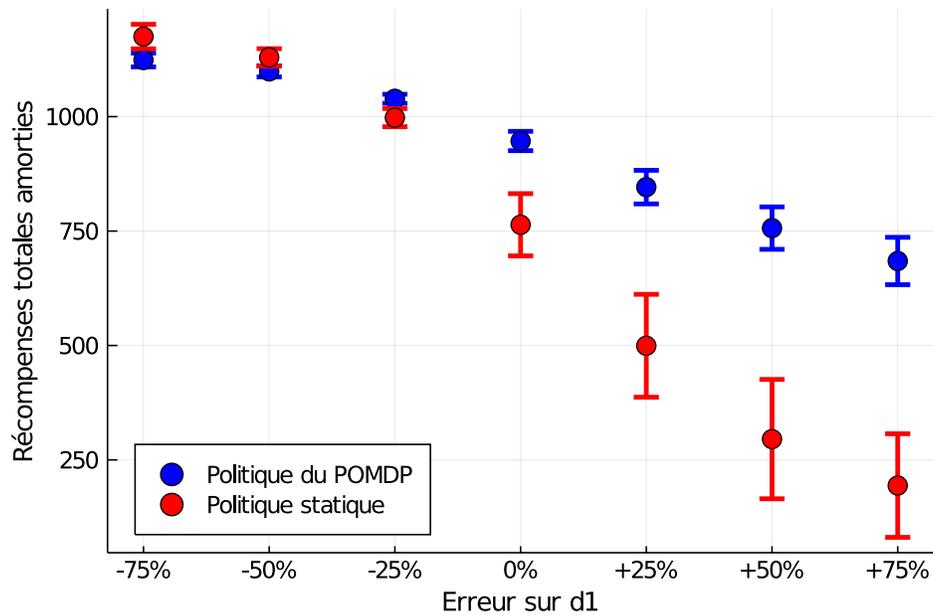


Figure 4.14 Robustesse à une erreur constante dans le modèle de performance

adaptative perd son intérêt devant une simple stratégie statique. Par ailleurs, lorsque le modèle sous estime l'impact de la charge de travail, l'écart entre les deux stratégies s'agrandit en faveur de la stratégie issue du POMDP.

4.6.2 Erreur dans le modèle de confiance : constantes de temps

On cherche ici à tester la robustesse de la stratégie adaptative à des erreurs dans le modèle de la confiance humaine. Supposons ici que dans (4.2), le paramètre μ devienne $\mu + k\mu$. De même on supposera qu'en fait η vaut $\eta + k\eta$. k est exprimé par un pourcentage d'erreur.

La figure 4.15 affiche les récompenses moyennes amorties et leur écart-type obtenus. Sur la figure du haut (resp. du bas), on introduit en simulation une valeur de μ (resp. de η) différente de celle utilisée dans le calcul du POMDP.

La figure 4.15 montre que les récompenses récoltées par la politique adaptative sont toujours supérieures à celles récoltées par la politique statique, et ce, malgré des erreurs d'identification des paramètres η ou μ .

4.6.3 Erreur dans le modèle de la confiance : fonction de capacité réelle de la machine

On a déjà évoqué la difficulté potentielle à définir la fonction représentant par un scalaire la capacité réelle démontrée par la machine pendant une période. Voyons si une erreur de définition de cette fonction affecte l'avantage de la stratégie dynamique. La fonction utilisée dans la résolution du POMDP est toujours donnée par 4.3. D'autres fonctions sont utilisées en simulations. Les candidats testés sont les suivants :

- Performances de l'équipe non biaisée : l'opérateur prend en compte la récompense obtenue lors de la période de travail (il a intégré la relative importance des erreurs)

$$C_1(x_t^h, a_t) = \frac{R(x_t^h, a_t) - (\min(R^{FN}, R^{FP}) + R^m)}{\max(R^{VP}, R^{VN}) - (\min(R^{FN}, R^{FP}) + R^m)} \quad (4.18)$$

- Performances de la machine seule : l'opérateur mesure les capacités de la machine par le pourcentage de bonnes classifications de la machine seule.

$$C_2(x_t^h, a_t) = (1 - W_t)P_s^m, \quad (4.19)$$

où P_s^m est le pourcentage de succès (vrais positifs ou vrais négatifs) de la machine (voir l'équation (4.4)).

- Performances moyennes de la machine seule : l'opérateur ne prend en compte que les performances moyennes de la machine, quelque soit le pourcentage de classifications réalisées par la machine pendant la période t .

$$C_3(x_t^h, a_t) = P_s^m \quad (4.20)$$

Encore une fois, on trace les récompenses moyennes obtenues en utilisant les différentes définitions de C . On voit sur la figure 4.16 que malgré des définitions erronées de la fonction décrivant les capacités réelles de la machine, la stratégie adaptative garde son bénéfice.

Remarque. On remarque que dans le cas de la fonction C_2 , les performances de l'équipe sont grandement diminuées, surtout lorsque la politique statique est adoptée. Ceci s'explique par le fait que la confiance de l'opérateur est faible, le conduisant ainsi à réaliser presque toutes les tâches. Cette surcharge de travail lui fait commettre beaucoup d'erreurs. L'opérateur a peu confiance car, selon C_2 , il ne juge la machine que par la quantité de tâches réussies par elle sans prendre en compte que lui aussi participe à accomplir les tâches. Autrement dit, il considère les bonnes classifications qu'il réalise lui-même comme des erreurs de la machine.

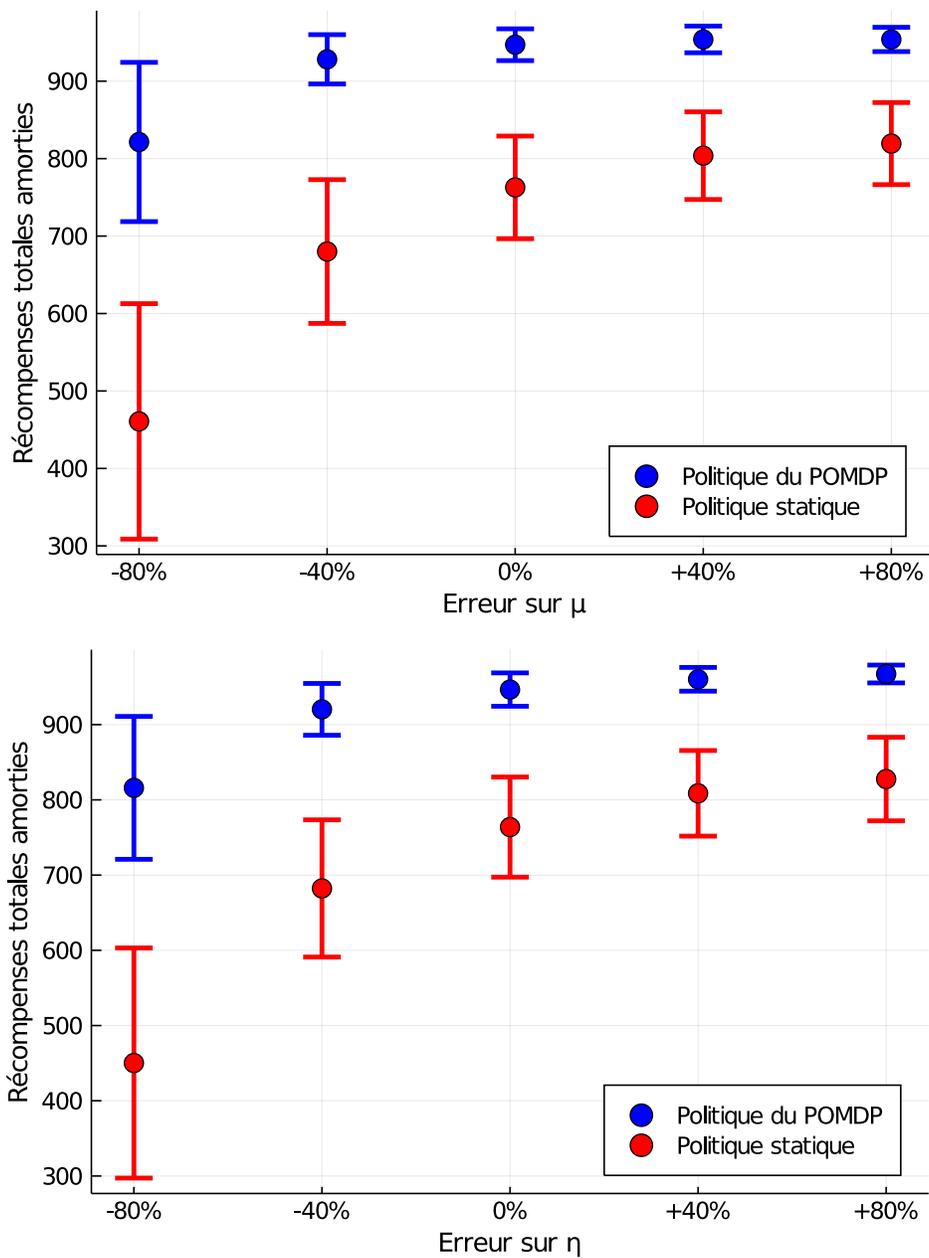


Figure 4.15 Robustesse à une erreur constante dans le modèle de confiance

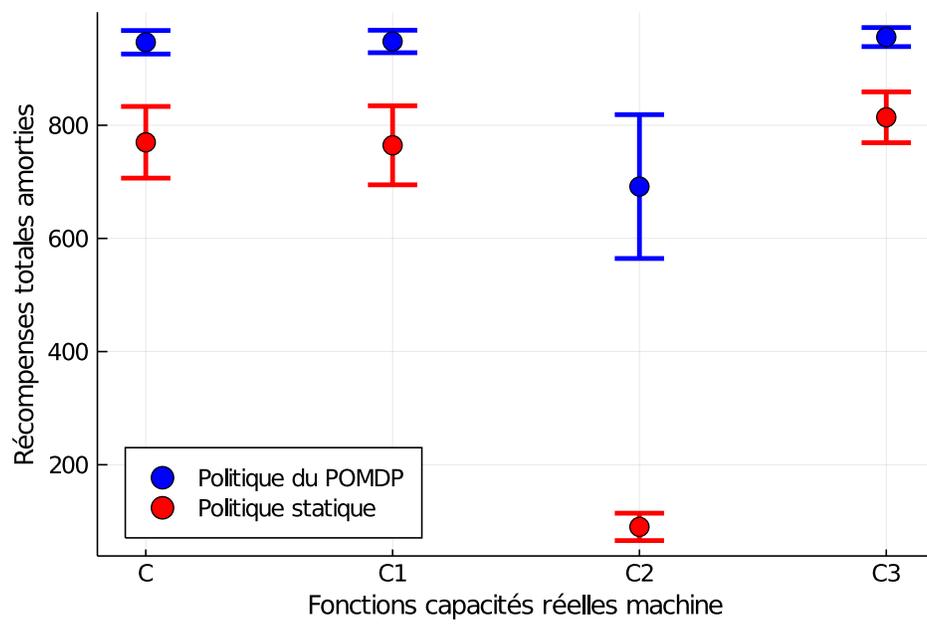


Figure 4.16 Robustesse à une erreur de définition de C

CHAPITRE 5 CONCLUSION

Ce dernier chapitre réalise une synthèse du travail proposé dans ce mémoire. On évoque les limitations de l’approche avancée. Enfin, plusieurs pistes de réflexions sont apportées pour guider de potentielles améliorations futures.

5.1 Synthèse des travaux

L’objectif de recherche était le suivant : améliorer les performances d’une équipe humain-machine en adoptant une stratégie de collaboration adaptative prenant en compte certaines caractéristiques propres à un humain.

La première étape fut la définition d’un scénario type d’interaction entre humain et machine susceptible de servir de support au développement d’une méthode de collaboration adaptative. Il a été choisi d’étudier un scénario de classifications binaires répétées réalisées soit par la machine soit par l’humain. Cette décision de choisir qui de la machine ou de l’humain réalise chaque classification revient à l’opérateur lui-même mais la machine peut lui suggérer un partage de tâches. Au regard de cette tâche, deux principaux facteurs, relatifs à l’humain, sont susceptibles d’impacter les performances de l’équipe. Ces facteurs sont d’une part l’influence de la charge de travail de l’opérateur sur ses performances de classification et d’autre part sa propension à utiliser la machine.

Afin de guider une stratégie de partage de tâches améliorant les performances de l’équipe, on a choisi de se baser sur des modèles quantitatifs permettant de prédire le comportement de l’opérateur, tout en gardant en tête que ces modèles pouvaient être approximatifs. La relation entre performance de classification et charge de travail a été modélisée par des courbes ROC dont la forme s’inspire des modèles de la Théorie de la Détection du Signal. La dépendance à la machine est, quant à elle, prédite par un modèle dynamique de la confiance proposé par Gao et Lee (Gao et Lee, 2006b).

La nature dynamique du modèle de la confiance nous a conduit à formuler le problème de suggestion de partage de tâches comme un Processus Décisionnel Markovien. De plus cet outil a l’avantage de prendre en compte la part d’aléatoire du comportement de l’humain. En outre, afin de prendre en compte le caractère non observable de l’état de confiance de l’humain, on a reformulé le problème sous la forme d’un POMDP. Grâce à cette méthode, la machine est capable de suggérer un partage de tâches à l’opérateur en prenant en compte son état présumé de confiance.

La présence de modèles qui peuvent être compliqués à définir et à identifier suivant l’environnement nous a conduit à nous questionner sur le réel bénéfice d’une telle approche par rapport à des stratégies plus naïves mais plus simple à mettre en place. Les simulations montrent que la stratégie issue de la résolution du POMDP apporte un bénéfice non négligeable sur le total des récompenses acquises sur le long terme par rapport à une stratégie d’allocation de tâche statique (récompense moyenne augmentée de 24%) . Cet avantage s’explique par plusieurs comportements. D’abord, la machine est capable de prendre en compte que l’opérateur peut avoir un a priori biaisé sur les capacités de la machine. De plus, lorsque l’opérateur change soudainement d’état de confiance (chute ou hausse non prédite), la machine est capable de produire une suggestion de partage de tâches ajustée, maintenant ainsi une charge de travail adaptée à l’opérateur pour que les performances de l’équipe n’en pâtissent pas trop.

Enfin nous avons exploré, toujours en simulation, la robustesse de la stratégie adaptative face à certaines erreurs d’identification des paramètres des modèles. Les résultats montrent que des erreurs trop importantes dans certains paramètres peuvent remettre en question la pertinence de la stratégie adaptative. Cependant la méthode semble être plus robuste lorsque les erreurs concernent d’autres paramètres.

5.2 Limitations de la solution

5.2.1 Variété des scénarios

On a évoqué dans les justifications du scenario de collaboration (section 1.2.2) le caractère général de la tâche choisie. En effet, on peut imaginer beaucoup de contextes différents dans lesquels une tâche de classification binaire répétée doit être effectuée en collaboration entre une machine et un humain. En outre, souvent, l’humain a l’autorité sur la machine. On a supposé que les deux principaux facteurs affectant les performances de l’équipe étaient la charge de travail portée par l’opérateur et la confiance que ce dernier place dans les capacités de la machine. Or, dans certains contextes de travail, d’autres facteurs mériteraient d’être pris en compte si l’on souhaite prédire le comportement de l’opérateur, même de manière approximative. On énumère dans la suite certains de ces facteurs.

La confiance en soi

On pourrait citer par exemple, l’impact de la confiance en soi de l’opérateur sur son comportement de dépendance à la machine. La confiance en soi est intégrée dans les modèles et solutions de collaboration suivants (Gao et Lee, 2006b), (Saeidi et Wang, 2019), (Van Dongen et Van Maanen, 2013), (Lee et Moray, 1994), (Gao *et al.*, 2013).

La charge mentale

Dans ce mémoire, on a défini la charge de travail comme la proportion de classifications réalisées par l'opérateur. Dans la littérature, on distingue souvent la charge de travail telle que définie ici, à la charge mentale. La charge mentale est une mesure subjective du coût demandé à un humain pour compléter une tâche avec une certaine performance. Il pourrait être alors nécessaire d'intégrer une variable décrivant la charge mentale de travail en tout temps afin de mieux prédire les performances de l'opérateur. Diverses approches existent pour mesurer la charge mentale d'un opérateur notamment à partir de capteurs physiologiques, voir (Charles et Nixon, 2019) pour une revue récente.

La compréhensibilité de la machine

Imaginons que la machine soit capable d'indiquer un niveau de confiance sur chacune des classifications qu'elle réalise, comme le permettent généralement les algorithmes de classification construits par apprentissage. L'opérateur s'aiderait probablement de cette information pour décider combien et quel objet il devrait classer manuellement. La stratégie de suggestion de partage de tâches devrait alors prendre en compte également ces indices de confiance.

5.2.2 Difficulté de calibration des modèles

Les modèles utilisés dans la stratégie de partage de tâches adaptative nécessitent la calibration de certains de leurs paramètres. Cette étape d'identification requiert la collecte de données expérimentales. On a abordé cette étape dans le cas du modèle de performance humaine (section 3.2) mais pas dans le cas du modèle de la confiance. En effet la calibration de ce modèle est plus complexe pour les raisons suivantes :

1. L'entrée du modèle (la fonction C décrivant les capacités réelles de la machine comme un réel) n'est pas clairement définie. En effet elle peut dépendre fortement du contexte (voir les différents candidats proposés dans 4.6.3). La définition de cette fonction peut requérir une étude préliminaire.
2. La confiance est dynamique et les suggestions de la machine ont des conséquences sur cette dynamique.
3. Les états B_t et T_t ne sont pas directement observables. Seul le comportement de l'opérateur l'est. Le modèle liant la dépendance de l'opérateur à son état de confiance doit lui même être identifié.

5.3 Améliorations futures

5.3.1 Validation expérimentales

Concernant la méthodologie adoptée dans ces travaux, la principale limitation porte sur l'absence de validation expérimentale. Dans un premier temps, une expérience doit être imaginée pour évaluer et calibrer les modèles sélectionnés. La stratégie issue du POMDP pourra alors être calculée. Enfin, une seconde expérience doit confronter les participants à la tâche de collaboration adoptant la stratégie adaptative pour suggérer le partage de tâches à chaque période. Cette expérience devra vérifier si la stratégie adaptative augmente significativement les récompenses cumulées par rapport à la stratégie statique.

5.3.2 Étudier les effets négligés dans les modèles

L'objectif premier de la conception d'un partage de tâches adaptatif est l'amélioration des performances à long terme. D'autres facteurs peuvent être pris en compte dans l'évaluation de la méthode. Par exemple, il serait intéressant de s'assurer que la charge mentale de l'opérateur reste convenable. Il faut également vérifier que le changement régulier de distribution de tâche n'introduise pas de confusion ou d'effet de surprise pour l'opérateur. L'utilisation de cet outil optimisant la collaboration ne doit pas non plus baisser son attention ou faire naître de la complaisance. Dans le cas de notre scénario, donner une suggestion de partage de tâches automatique à l'opérateur ne doit pas lui faire penser qu'il pourrait se dispenser de prêter attention aux performances de la machine. Ceci nuirait à sa conscience de la situation. Pour synthétiser, la solution de partage de tâches adaptatif proposée doit être évaluée à la lumière de plusieurs facteurs humains en plus des performances de l'équipe.

5.3.3 Comparer la solutions à d'autres approches

Le faible nombre de modèles quantitatifs de la confiance dans la littérature montre la difficulté de leur définition. Pourtant ces modèles, une fois calibrés, ont l'avantage de permettre la conception d'une stratégie adaptative prête à l'emploi. De plus, bien que ces modèles aient une capacité de prédiction limitée, les outils de commande stochastique permettent de gérer cela.

Cependant, d'autres outils sont capables de concevoir une stratégie de partage de tâches qui s'adapte à l'état d'un opérateur. En particulier, les méthodes d'apprentissage par renforcement, peuvent apprendre une stratégie adaptative au cours des interactions avec l'opérateur. Il n'est pas nécessaire de fournir de modèle prédisant le comportement d'un humain. Seules les

données suffisent à l'optimisation de la stratégie. Néanmoins la quantité de données nécessaire pour que les algorithmes convergent vers une solution convenable peut être très importante. Il pourrait être intéressant de comparer l'approche par modèle et celle basée sur une boîte noire afin de déterminer si elles aboutissent à des politiques identiques. Si on fournit des données physiologiques aux algorithmes d'apprentissage, on pourrait s'attendre à ce que la machine soit capable de mieux connaître l'état cognitif de l'opérateur à chaque instant avant même qu'il ne se traduise en comportement.

RÉFÉRENCES

- Azevedo-Sa, H., Jayaraman, S. K., Esterwood, C. T., Yang, X. J., Robert, L. P. et Tilbury, D. M. (2020). Real-time estimation of drivers' trust in automated driving systems. *International Journal of Social Robotics*.
- Boubin, J. G., Rusnock, C. F. et Bindewald, J. M. (2017). Quantifying compliance and reliance trust behaviors to influence trust in human-automation teams. *In Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 61, pages 750–754. SAGE Publications Sage CA : Los Angeles, CA.
- Charles, R. L. et Nixon, J. (2019). Measuring mental workload using physiological measures : A systematic review. *Applied Ergonomics*, 74:221 – 232.
- Chavaillaz, A., Wastell, D. et Sauer, J. (2016). System reliability, performance and trust in adaptable automation. *Applied Ergonomics*, 52:333–342.
- Chen, M., Nikolaidis, S., Soh, H., Hsu, D. et Srinivasa, S. (2020). Trust-aware decision making for human-robot collaboration : Model learning and planning. *ACM Transactions on Human-Robot Interaction*, 9(2).
- Cummings, M. M. (2014). Man versus machine or man + machine? *IEEE Intelligent Systems*, 29(5):62–69.
- da Silva, F. P. (2014). Mental workload, task demand and driving performance : What relation. *Procedia-Social and Behavioral Sciences*, 162:310–319.
- Desai, M. (2012). *Modeling Trust to Improve Human-Robot Interaction*. Ph.d., University of Massachusetts Lowell.
- Dixon, S. R., Wickens, C. D. et McCarley, J. S. (2007). On the independence of compliance and reliance : Are automation false alarms worse than misses? *Human Factors*, 49(4):564–572.

Dubois, C. et Le Ny, J. (2020). Adaptive task allocation in human-machine teams with trust and workload cognitive models. *In 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 3241–3246.

Dzindolet, M. T., Beck, H. P., Pierce, L. G. et Dawe, L. A. (2001). A framework of automation use. Rapport technique, ARMY RESEARCH LAB ABERDEEN PROVING GROUND MD.

Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G. et Beck, H. P. (2003). The role of trust in automation reliance. *International Journal of Human-Computer Studies*, 58(6):697–718.

Egan, J. P. (1975). *Signal detection theory and ROC-analysis*. Academic Press, New York.

Egorov, M., Sunberg, Z. N., Balaban, E., Wheeler, T. A., Gupta, J. K. et Kochenderfer, M. J. (2017). POMDPs.jl : A framework for sequential decision making under uncertainty. *Journal of Machine Learning Research*, 18(26):1–5.

Farrell, S. et Lewandowsky, S. (2000). A connectionist model of complacency and adaptive recovery under automation. *Journal of Experimental Psychology : Learning, Memory, and Cognition*, 26(2):395.

Fawcett, T. (2006). Introduction to receiver operator curves. *Pattern Recognit. Lett*, 27:861–874.

Floyd, M. W., Drinkwater, M. et Aha, D. W. (2014). Adapting autonomous behavior using an inverse trust estimation. *In International Conference on Computational Science and Its Applications*, pages 728–742. Springer.

Floyd, M. W., Drinkwater, M. et Aha, D. W. (2015). Improving trust-guided behavior adaptation using operator feedback. *In International Conference on Case-Based Reasoning*, pages 134–148. Springer.

Freedy, A., DeVisser, E., Weltman, G. et Coeyman, N. (2007). Measurement of trust in human-robot collaboration. *In 2007 International Symposium on Collaborative Technologies and Systems*, pages 106–114.

Gao, F., Clare, A., Macbeth, J. et Cummings, M. (2013). Modeling the impact of operator trust on performance in multiple robot control. *In 2013 AAAI Spring Symposium Series*.

Gao, J. et Lee, J. D. (2006a). Effect of shared information on trust and reliance in a demand forecasting task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(3):215–219.

Gao, J. et Lee, J. D. (2006b). Extending the decision field theory to model operators' reliance on automation in supervisory control situations. *IEEE Transactions on Systems, Man, and Cybernetics-Part A : Systems and Humans*, 36(5):943–959.

Gao, J., Lee, J. D. et Zhang, Y. (2006). A dynamic model of interaction between reliance on automation and cooperation in multi-operator multi-automation situations. *International Journal of Industrial Ergonomics*, 36(5):511–526.

Green, D. M. et Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. Wiley New York.

Gregoriades, A. et Sutcliffe, A. (2008). Workload prediction for improved design and reliability of complex systems. *Reliability Engineering & System Safety*, 93(4):530–549.

Guo, Y. et Yang, X. J. (2020). Modeling and predicting trust dynamics in human–robot teaming : A bayesian inference approach. *International Journal of Social Robotics*.

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J. et Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human factors*, 53(5):517–527.

Hart, S. G. (2006). Nasa-task load index (nasa-tlx); 20 years later. *In Proceedings of the human factors and ergonomics society annual meeting*, volume 50, pages 904–908. Sage publications Sage CA : Los Angeles, CA.

Hart, S. G. et Staveland, L. E. (1988). Development of nasa-tlx (task load index) : Results of empirical and theoretical research. *In Advances in psychology*, volume 52, pages 139–183. Elsevier.

Hiatt, L. M., Narber, C., Bekele, E., Khemlani, S. S. et Trafton, J. G. (2017). Human modeling for human–robot collaboration. *The International Journal of Robotics Research*, 36(5-7):580–596.

Inagaki, T. et Itoh, M. (2010). Theoretical framework for analysis and evaluation of human’s overtrust in and overreliance on advanced driver assistance systems.

Kaniarasu, P., Steinfeld, A., Desai, M. et Yanco, H. (2012). Potential measures for detecting trust changes. In *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 241–242. IEEE.

Kurniawati, H., Hsu, D. et Lee, W. S. (2008). Sarsop : Efficient point-based pomdp planning by approximating optimally reachable belief spaces. In *Robotics : Science and systems*, volume 2008. Zurich, Switzerland.

Lee, J. et Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, 35(10):1243–1270.

Lee, J. D. et Moray, N. (1994). Trust, self-confidence, and operators’ adaptation to automation. *International journal of human-computer studies*, 40(1):153–184.

Lee, J. D. et See, K. A. (2004). Trust in automation : Designing for appropriate reliance. *Human factors*, 46(1):50–80.

Li, H. (2020). *A computational model of human trust in supervisory control of robotic swarms*. Thèse de doctorat, University of Pittsburgh.

Liu, B. (2020). A survey on trust modeling from a bayesian perspective. *Wireless Personal Communications*, pages 1–23.

Loft, S., Bolland, S., Humphreys, M. S. et Neal, A. (2009). A theory and model of conflict detection in air traffic control : Incorporating environmental constraints. *Journal of Experimental Psychology : Applied*, 15(2):106.

Loft, S., Sanderson, P., Neal, A. et Mooij, M. (2007). Modeling and predicting mental workload in en route air traffic control : Critical review and broader implications. *Human*

factors, 49(3):376–399.

Madhavan, P., Wiegmann, D. A. et Lacson, F. C. (2006). Automation failures on tasks easily performed by operators undermine trust in automated aids. *Human factors*, 48(2):241–256.

Madsen, M. et Gregor, S. (2000). Measuring human-computer trust. *In 11th australasian conference on information systems*, volume 53, pages 6–8. Citeseer.

Manning, C. A., Mills, S. H., Fox, C. M., Pfeleiderer, E. M. et Mogilka, H. J. (2002). Using air traffic control taskload measures and communication events to predict subjective workload. Rapport technique, FEDERAL AVIATION ADMINISTRATION OKLAHOMA CITY OK CIVIL AEROMEDICAL INST.

McBride, S. E., Rogers, W. A. et Fisk, A. D. (2010). Do younger and older adults differentially depend on an automated system? *In Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 54, pages 175–179. SAGE Publications Sage CA : Los Angeles, CA.

Moray, N. (1982). Subjective mental workload. *Human factors*, 24(1):25–40.

Muslim, H. et Itoh, M. (2019). A theoretical framework for designing human-centered automotive automation systems. *Cognition, Technology Work*, 21(4):685–697.

Nam, C., Walker, P., Lewis, M. et Sycara, K. (2017). Predicting trust in human control of swarms via inverse reinforcement learning. *In IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, Lisbon, Portugal.

Oduor, K. F. et Wiebe, E. N. (2008). The effects of automated decision algorithm modality and transparency on reported trust and task performance. *In Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 52, pages 302–306. SAGE Publications Sage CA : Los Angeles, CA.

Parasuraman, R. (2000). Designing automation for human use : empirical studies and quantitative models. *Ergonomics*, 43(7):931–951.

Parasuraman, R., Bahri, T., Deaton, J. E., Morrison, J. G. et Barnes, M. (1992). Theory

and design of adaptive automation in aviation systems. Rapport technique, Catholic Univ of America Washington DC cognitive science lab.

Parasuraman, R. et Riley, V. (1997). Humans and automation : Use, misuse, disuse, abuse. *Human factors*, 39(2):230–253.

Parasuraman, R., Sheridan, T. B. et Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on systems, man, and cybernetics-Part A : Systems and Humans*, 30(3):286–297.

Park, E., Jenkins, Q. et Jiang, X. (2008). Measuring trust of human operators in new generation rescue robots. *In Proceedings of the JFPS International Symposium on Fluid power*, volume 2008, pages 489–492. The Japan Fluid Power System Society.

Perkins, L., Miller, J. E., Hashemi, A. et Burns, G. (2010). Designing for human-centered systems : Situational risk as a factor of trust in automation. *In Proceedings of the human factors and ergonomics society annual meeting*, volume 54, pages 2130–2134. SAGE Publications Sage CA : Los Angeles, CA.

Peterson, W., Birdsall, T. et Fox, W. (1954). The theory of signal detectability. *Transactions of the IRE professional group on information theory*, 4(4):171–212.

Putri, M., Xiaonan, Y. et Jung Hyup, K. (2016). Sensitivity, bias, and mental workload in a multitasking environment. *In Engineering Psychology and Cognitive Ergonomics. 13th International Conference, EPCE 2016, held as part of HCI International 2016, 17-22 July 2016*, Engineering Psychology and Cognitive Ergonomics. 13th International Conference, EPCE 2016, held as part of HCI International 2016. Proceedings : 9736, pages 10–18, Cham, Switzerland. Springer International Publishing.

Rice, S. et Keller, D. (2009). Automation reliance under time pressure. *Cognitive Technology*.

Robinette, P., Li, W., Allen, R., Howard, A. M. et Wagner, A. R. (2016). Overtrust of robots in emergency evacuation scenarios. *In 11th Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI 2016, March 7, 2016 - March 10, 2016*, volume 2016-April de ACM/IEEE International Conference on Human-Robot Interaction, pages 101–108, Christchurch, New zealand. IEEE Computer Society.

Rouse, W. B. (1988). Adaptive aiding for human/computer control. *Human factors*, 30(4): 431–443.

Rouse, W. B., Edwards, S. L. et Hammer, J. M. (1993). Modeling the dynamics of mental workload and human performance in complex systems. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(6):1662–1671.

Sadrifaridpour, B. (2018). Trust-based control of robotic manipulators in collaborative assembly in manufacturing.

Saeidi, H., Wagner, J. R. et Wang, Y. (2017). A mixed-initiative haptic teleoperation strategy for mobile robotic systems based on bidirectional computational trust analysis. *IEEE Transactions on Robotics*, 33(6):1500–1507.

Saeidi, H. et Wang, Y. (2019). Incorporating trust and self-confidence analysis in the guidance and control of (semi)autonomous mobile robotic systems. *IEEE Robotics and Automation Letters*, 4(2):239–246.

Sanchez, J., Rogers, W. A., Fisk, A. D. et Rovira, E. (2014). Understanding reliance on automation : effects of error type, error distribution, age and experience. *Theoretical issues in ergonomics science*, 15(2):134–160.

Sanders, T., Oleson, K. E., Billings, D. R., Chen, J. Y. et Hancock, P. A. (2011). A model of human-robot trust : Theoretical model development. *In Proceedings of the human factors and ergonomics society annual meeting*, volume 55, pages 1432–1436. SAGE Publications Sage CA : Los Angeles, CA.

Savla, K. et Frazzoli, E. (2011). A dynamical queue approach to intelligent task management for human operators. *Proceedings of the IEEE*, 100(3):672–686.

Schaefer, K. E. (2016). Measuring trust in human robot interactions : Development of the trust perception scale-hri. *Robust Intelligence and Trust in Autonomous Systems*, pages 191–218. Springer US.

Schwark, J., Dolgov, I., Graves, W. et Hor, D. (2010). The influence of perceived task difficulty and importance on automation use. *In Proceedings of the Human Factors and*

Ergonomics Society Annual Meeting, volume 54, pages 1503–1507. SAGE Publications Sage CA : Los Angeles, CA.

Soh, H., Shu, P., Chen, M. et Hsu, D. (2019). Trust dynamics and transfer across human-robot interaction tasks : Bayesian and neural computational models. *In IJCAI*, pages 6226–6230.

Sorkin, R. D. et Woods, D. D. (1985). Systems with human monitors : A signal detection analysis. *Human-computer interaction*, 1(1):49–75.

Spackman, K. A. (1989). Signal detection theory : Valuable tools for evaluating inductive learning. *In Segre, A. M., éditeur : Proceedings of the Sixth International Workshop on Machine Learning*, pages 160 – 163. Morgan Kaufmann, San Francisco (CA).

Srivastava, V., Surana, A. et Bullo, F. (2012). Adaptive attention allocation in human-robot systems. *In 2012 American Control Conference (ACC)*, pages 2767–2774. IEEE.

Swets, J. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240(4857):1285–1293.

Szalma, J. L. et Taylor, G. S. (2011). Individual differences in response to automation : The five factor model of personality. *Journal of Experimental Psychology : Applied*, 17(2):71.

van den Brule, R., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J. et Haselager, P. (2014). Do robot performance and behavioral style affect human trust ? : A multi-method approach. *International Journal of Social Robotics*, 6(4):519–531.

Van Dongen, K. et Van Maanen, P.-P. (2013). A framework for explaining reliance on decision aids. *International Journal of Human-Computer Studies*, 71(4):410–424.

Van Maanen, P.-P., Klos, T. et Van Dongen, K. (2007). Aiding human reliance decision making using computational models of trust. *In 2007 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology-Workshops*, pages 372–376. IEEE.

Wang, Y., Humphrey, L. R., Liao, Z. et Zheng, H. (2018). Trust-based multi-robot symbolic

motion planning with a human-in-the-loop. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 8(4):1–33.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, 50(3):449–455.

Wickens, C. D., Hollands, J. G., Banbury, S. et Parasuraman, R. (2015). *Engineering psychology and human performance*. Psychology Press.

Wickens, T. D. (2002). *Elementary signal detection theory*. Oxford University Press, USA.

Wilkison, B. D. (2008). *Effects of mental model quality on collaborative system performance*. Thèse de doctorat, Georgia Institute of Technology.

Wu, C. et Liu, Y. (2007). Queuing network modeling of driver workload and performance. *IEEE Transactions on Intelligent Transportation Systems*, 8(3):528–537.

Xu, A. (2016). *Efficient Collaboration with Trust-Seeking Robots*. Thèse de doctorat, McGill University.

Xu, A. et Dudek, G. (2015). OPTIMo : Online probabilistic trust inference model for asymmetric human-robot collaborations. In *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Portland, OR.

Xu, A. et Dudek, G. (2016). Maintaining efficient collaboration with trust-seeking robots. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3312–3319.

Xu, A. et Dudek, G. (2016). Towards modeling real-time trust in asymmetric human-robot collaborations. In *Robotics Research*, pages 113–129. Springer.

Yagoda, R. E. et Gillan, D. J. (2012a). You want me to trust a robot? the development of a human-robot interaction trust scale. *International Journal of Social Robotics*, 4(3):235–248.

Yagoda, R. E. et Gillan, D. J. (2012b). You want me to trust a robot? the development of a human-robot interaction trust scale. *International Journal of Social Robotics*, 4(3):235–248.

Young, M. S., Brookhuis, K. A., Wickens, C. D. et Hancock, P. A. (2015). State of science : mental workload in ergonomics. *Ergonomics*, 58(1):1–17.

Zhang, X. et Lin, H. (2019). Performance guaranteed human-robot collaboration with pomdp supervisory control. *Robotics and Computer-Integrated Manufacturing*, 57:59–72.

Zhou, T. (2011). The effect of initial trust on user adoption of mobile payment. *Information Development*, 27(4):290–300.