



Titre: Renvois de décisions dans une équipe humain-machine prise en
Title: compte de la confiance

Auteur: Antoine Drolet
Author:

Date: 2023

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Drolet, A. (2023). Renvois de décisions dans une équipe humain-machine prise en
Citation: compte de la confiance [Mémoire de maîtrise, Polytechnique Montréal].
PolyPublie. <https://publications.polymtl.ca/55833/>

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

URL de PolyPublie: <https://publications.polymtl.ca/55833/>
PolyPublie URL:

**Directeurs de
recherche:** Jérôme Le Ny
Advisors:

Programme: Génie électrique
Program:

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

**Renvois de décisions dans une équipe humain-machine:
prise en compte de la confiance**

ANTOINE DROLET

Département de génie électrique

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

Septembre 2023

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé :

**Renvois de décisions dans une équipe humain-machine:
prise en compte de la confiance**

présenté par **Antoine DROLET**

en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
a été dûment accepté par le jury d'examen constitué de :

Guchuan ZHU, président

Jérôme LE NY, membre et directeur de recherche

Roland MALHAMÉ, membre

DÉDICACE

*À quoi ça sert
d'être brillant
si t'éclaires
personne*
Patrice Desbiens

REMERCIEMENTS

Je tiens à exprimer ma sincère gratitude envers les personnes et institutions qui ont joué un rôle déterminant dans la réalisation de ce mémoire :

Tout d’abord, un grand merci au programme *Innovation for Defence Excellence and Security* (IDEaS) *Micro-Net program* pour le financement qui a rendu possible la réalisation de mes travaux de recherche.

Un remerciement tout particulier à mon directeur de recherche, Jérôme Le Ny, pour son temps, sa rigueur scientifique et ses précieux conseils qui ont guidé mes recherches, ainsi qu’aux professeurs Roland Malhamé et Guchuan Zhu qui ont accepté de prendre le temps de faire parti de mon jury.

Je tiens également à remercier chaleureusement mes collègues du laboratoire de la Polytechnique pour leur soutien, leur collaboration et les échanges enrichissants tout au long de cette expérience.

Enfin, je souhaite exprimer ma profonde gratitude envers ma famille pour leur amour, leur soutien et leur encouragement constant, qui ont été les piliers de ma réussite.

RÉSUMÉ

La performance d'une équipe humain-machine dans un contexte de tâches décisionnelles est fortement influencée par des facteurs humains tels que la charge de travail de l'opérateur et sa confiance envers l'automatisation. Il devient essentiel de tenir compte de ces facteurs lors de la conception d'une stratégie de collaboration d'une telle équipe. Cependant, la confiance et la charge de travail évoluent dans le temps en fonction des interactions entre l'humain et la machine. Dans ce contexte, une approche prometteuse consiste à adopter une stratégie de collaboration adaptative qui s'ajuste en fonction de l'état cognitif de l'opérateur. Dans ce mémoire, une politique de collaboration adaptative est proposée, où l'automatisation décide quelles tâches décisionnelles spécifiques devraient être envoyées à l'humain après avoir d'abord observé l'ensemble des tâches à effectuer. On parlera donc de renvois de décisions pour décrire ces tâches qui sont envoyées à l'humain. Ce dernier, quant à lui, peut choisir de suivre ou non cette recommandation en fonction de son état de confiance. Nous posons comme hypothèse principale qu'il y a une plus grande incertitude quant à la réponse de l'opérateur, soit sa disposition à suivre la recommandation de l'automatisation, lorsque son niveau de confiance envers l'automatisation est faible. La stratégie de l'automatisation repose sur la résolution d'un processus décisionnel Markovien et utilise des modèles quantitatifs des performances humaines ainsi qu'un modèle dynamique simplifié de la confiance humaine envers l'automatisation. Nous démontrons analytiquement que, sous certaines hypothèses, la politique d'action optimale consiste en une politique myope qui vise à minimiser le coût de classification de l'équipe à chaque période. Des simulations sont réalisées afin d'évaluer l'efficacité de la politique optimale proposée, en la comparant à d'autres stratégies plus simples à implémenter. Des résultats de simulations permettent de conjecturer sur l'optimalité de la politique myope sous certaines conditions dans le cas d'un processus décisionnel Markovien partiellement observable. Cette situation correspond à un scénario plus réaliste pour lequel l'état de confiance de l'opérateur est estimé plutôt que mesuré directement.

ABSTRACT

The performance of a human-machine team in making joint decisions is strongly influenced by human factors such as workload and trust in automation. It is essential to consider these factors when designing a collaboration strategy in human-automation teams. However, trust and workload evolve over time based on interactions between humans and machines. In this context, a promising approach is to adopt an adaptive collaboration strategy that adjusts to the operator's cognitive state. In this thesis, an adaptive collaboration policy is proposed, where the automation decides which specific decision-making tasks should be delegated to humans after initially observing all the tasks. Therefore, we refer to these delegated tasks as "decision referrals". The human operator can choose to follow or disregard these recommendations based on their level of trust. Our main hypothesis is that there is greater uncertainty regarding the operator's response, i.e., their willingness to follow the automation's recommendation, when their level of confidence in the automation is low. The automation's policy is based on solving a Markov decision process and utilizes quantitative models of human performance and a simplified dynamic model of human trust in automation. Our analysis shows that under certain assumptions, the optimal automation policy is a myopic policy that aims at minimizing the team's classification cost at each period. Simulations are conducted to evaluate the effectiveness of the proposed optimal policy and compare it to other simpler implementation strategies. Simulation results allow us to conjecture about the optimality of the myopic policy under certain conditions in the case of a partially observable Markov decision process. This situation corresponds to a more realistic scenario where the operator's confidence state is estimated rather than directly measured.

TABLE DES MATIÈRES

DÉDICACE	iii
REMERCIEMENTS	iv
RÉSUMÉ	v
ABSTRACT	vi
TABLE DES MATIÈRES	vii
LISTE DES TABLEAUX	x
LISTE DES FIGURES	xi
LISTE DES SIGLES ET ABRÉVIATIONS	xii
CHAPITRE 1 INTRODUCTION	1
1.1 Collaboration humaine avec un système automatisé	1
1.1.1 Mise en contexte	1
1.1.2 Particularités et problématiques liées à l'opérateur humain	2
1.2 Partage de tâches en classification	3
1.2.1 Description du scénario	3
1.2.2 Justification du scénario	5
1.3 Objectifs de recherche	6
1.3.1 Développer une politique de renvois de tâches	6
1.3.2 Évaluer la politique d'allocation de tâches	6
1.4 Plan du mémoire	6
CHAPITRE 2 REVUE DE LITTÉRATURE	7
2.1 Charge de travail et charge mentale	7
2.1.1 Définitions	7
2.1.2 Mesure de la charge mentale	7
2.1.3 Modélisation de la charge mentale	8
2.2 Confiance envers l'automatisation	10
2.2.1 Définitions	10
2.2.2 Facteurs influençant la confiance envers l'automatisation	11

2.2.3	Sur-confiance	11
2.2.4	Modèles quantitatifs de la confiance	12
2.3	Conception de stratégies de collaboration humain-machine	14
2.3.1	Stratégie d'allocation de contrôle	14
2.3.2	Stratégies de collaboration influençant le comportement de la machine	14
2.3.3	Stratégies de renvois de tâches	15
2.4	Conclusion de la revue de littérature	16
CHAPITRE 3 FORMULATION DU PROBLÈME		18
3.1	Description formelle du scénario	18
3.1.1	Impact de la confiance	19
3.1.2	Performance humaine	20
3.2	Formulation du problème en MDP	21
3.2.1	Définition de l'état	21
3.2.2	Définition de l'action	22
3.2.3	Définition de la fonction de coût	22
3.2.4	Définition de la fonction de transition	25
3.2.5	Description du scénario en MDP	25
3.3	Formulation du problème en POMDP	26
3.4	Modèles numériques	26
3.4.1	Modèle d'observation de l'automation	27
3.4.2	Modèle de tâches supplémentaires w_t^s	27
3.4.3	Modèle de performance humaine	29
CHAPITRE 4 OBSERVATION PARFAITE DE LA CONFIANCE		31
4.1	Calcul de la politique myope	31
4.2	Optimalité de la politique myope	33
4.3	Politique myope : version alternative	40
4.4	Simulations numériques	41
4.4.1	Définitions des politiques de comparaison	41
4.4.2	Comparaison des politiques	42
4.4.3	Conditions de quasi-optimalité pour la politique MS	49
4.5	Conclusion	53
CHAPITRE 5 OBSERVATION IMPARFAITE DE LA CONFIANCE		54
5.1	Formulation du problème	54
5.1.1	Modèle dynamique de la confiance	54

5.2	Politique de renvois de tâches	56
5.2.1	Impact de la variance sur le POMDP	57
5.2.2	Impact de la moyenne sur le POMDP	60
5.3	Conclusion	65
CHAPITRE 6 CONCLUSION		67
6.1	Synthèse des travaux	67
6.2	Limitations de la solution proposée	68
6.2.1	Modèles quantitatifs utilisés	68
6.2.2	Facteurs humains négligés	68
6.3	Améliorations futures	70
6.3.1	Valider les modèles de façon expérimentale	70
6.3.2	Inclure davantage de facteurs humains	70
6.3.3	Comparer la solution à d'autres approches	70
RÉFÉRENCES		71

LISTE DES TABLEAUX

Tableau 4.1	Valeurs numériques des différents modèles	43
-------------	---	----

LISTE DES FIGURES

Figure 1.1	Arbre de décision pour l'opérateur et radar de détection de cibles. . .	5
Figure 3.1	Coût de classification de l'humain et de l'automatisme	24
Figure 3.2	Distribution des probabilités de w^s avec une moyenne fixe égale à 1. .	28
Figure 3.3	Courbes de performance de classification de l'humain en fonction de sa charge de travail w^h	30
Figure 4.1	Coût de classification pour une moyenne de w^s fixe	45
Figure 4.2	Coût moyen par période avec une variance croissante	45
Figure 4.3	Coût cumulé pour 50 périodes avec une moyenne fixe.	46
Figure 4.4	Distribution des probabilités de w^s avec différentes moyennes.	47
Figure 4.5	Coût de classifications pour 50 périodes avec différentes moyennes. . .	48
Figure 4.6	Coût moyen par période avec une moyenne croissante.	48
Figure 4.7	Coût cumulé pour 50 périodes avec une moyenne croissante.	49
Figure 4.8	Gains G à deux postérieurs pour différentes charges de travail w_1 et w_2	52
Figure 5.1	Distribution du coût réalisé $f_c(c \mid w^s, a, p^a)$	56
Figure 5.2	s_{t+1} en fonction du coût réalisé c_t	57
Figure 5.3	Gain espéré pour différentes valeurs du modèle de w^s avec une moyenne fixe	58
Figure 5.4	$\hat{G}(s_t)$ pour différentes valeurs de paramètres du modèle w^s avec une moyenne fixe.	59
Figure 5.5	Politique optimale pour un modèle w^s à moyenne fixe	61
Figure 5.6	Gain espéré $\bar{G}(w^r, \tau)$ pour différentes moyennes du modèle de w^s . . .	62
Figure 5.7	$\tilde{G}(s_t)$ pour différentes moyennes du modèle w^s	63
Figure 5.8	Probabilité de transition de s_{t+1} pour différentes moyennes de w^s . . .	64
Figure 5.9	Politique myope pour différentes moyennes de w^s	66

LISTE DES SIGLES ET ABRÉVIATIONS

MDP	<i>Markov Decision Process</i>
POMDP	<i>Partially Observable Markov Decision Process</i>
SSD	Système de support d'aide à la décision

CHAPITRE 1 INTRODUCTION

Dans ce chapitre, nous mettons en évidence les différentes problématiques liées aux équipes dans lesquelles un humain collabore avec un système de support à la décision, ainsi que la pertinence de former d’une telle équipe. Nous présentons ensuite un scénario d’application précis sur lequel nous nous basons pour mener nos travaux. Enfin, nous formulons les objectifs de recherche de ce projet de maîtrise.

1.1 Collaboration humaine avec un système automatisé

1.1.1 Mise en contexte

L’avènement récent de technologies permettant l’automatisation de tâches de plus en plus complexes a suscité un intérêt croissant pour les équipes de travail impliquant la collaboration entre un humain et un système automatisé¹. La surveillance de systèmes de commande industriels [1] ou l’identification dynamique de menaces dans un contexte de défense militaire [2] sont quelques exemples parmi plusieurs d’équipes humain-machine. La pertinence de former une telle équipe relève d’une performance globale de l’équipe collaborative qui serait supérieure aux capacités individuelles de chaque agent la composant. En effet, les systèmes automatisés peuvent efficacement effectuer des tâches répétitives simples, tandis qu’un opérateur humain peut apporter son expertise, son jugement et sa capacité d’adaptation dans des situations plus complexes. Cependant, la conception d’une équipe dans laquelle collabore un humain et un système automatisé nécessite une réflexion approfondie quant à l’allocation adéquate des tâches décisionnelles au sein de l’équipe. Une mauvaise répartition des tâches peut effectivement entraîner une sous-utilisation des compétences humaines ou une surcharge cognitive, limitant ainsi les performances globales de l’équipe. La notion de renvois de décisions (*decision referrals*) diffère d’un problème plus simple de répartition de tâches adaptative en ce sens qu’après les avoir initialement observées, l’automation choisit quelles tâches spécifiques devraient être référées à l’opérateur, estimant que la classification de ces tâches par un humain serait bénéfique pour l’équipe [3, 4].

Pour répondre à cette problématique, une approche à considérer est l’utilisation d’une stratégie de renvois de décisions adaptative, qui consiste à laisser à l’automation la possibilité de référer certaines tâches à son coéquipier humain après avoir initialement observé l’ensemble

1. Dans ce mémoire, les termes « automation », « système automatisé » et « système de support d’aide à la décision (SSD) » sont utilisés de manière interchangeable pour faire référence à la composante « machine » de l’équipe. Le terme « opérateur » sera également utilisé pour faire référence à l’humain au sein de l’équipe.

des tâches. Ces renvois de décisions dépendent des caractéristiques spécifiques de la tâche, mais également des paramètres liés soit au système automatisé (performance du système, nombre de tâches à effectuer, etc.), soit à l'opérateur humain (performances, confiance envers le système automatisé, charge mentale, etc.) [5]. L'aspect adaptatif d'une telle stratégie, soit la considération de l'état de l'opérateur humain, permet la création d'une équipe flexible, capable de s'ajuster en fonction des différentes tâches et scénarios auxquels elle fait face. Dans la conception d'une stratégie adaptative, il est essentiel de déterminer quels éléments devront être adaptés en fonction de critères spécifiques, ainsi que de préciser comment ces adaptations doivent être réalisées [6]. Il convient de souligner que les réponses à ces questions sont hautement dépendantes du contexte particulier dans lequel la stratégie adaptative est implémentée.

Ce mémoire se concentre donc sur la conception d'une stratégie de renvois de décisions spécifique à un scénario donné. L'objectif de cette stratégie est d'améliorer les performances globales de l'équipe en tirant parti des caractéristiques propres à l'opérateur humain.

1.1.2 Particularités et problématiques liées à l'opérateur humain

La conception d'un système d'aide à la décision nécessite la compréhension de diverses caractéristiques propres au comportement humain susceptibles de poser problème. En effet, il existe de multiples facteurs difficilement quantifiables qui influencent le comportement humain et ses capacités à effectuer une tâche. Ces facteurs peuvent provenir du système automatisé (performance du système, transparence de l'interface, etc.), de l'environnement (stimuli externes, autres collaborateurs humains, etc.) ou de l'humain lui-même (compétences humaines, facteur culturel, fatigue, etc.) [7], [8]. L'historique des interactions entre l'humain et l'automation est également à prendre en compte puisque l'expérience passée peut guider les décisions de l'humain, notamment en termes de confiance envers le système automatisé [9].

À ces difficultés s'ajoute également le fait que, pour concevoir une stratégie de renvois de tâches efficace, il devient nécessaire de quantifier certains phénomènes humains. Par exemple, si l'on souhaite moduler la confiance de l'humain envers un système de support automatisé afin d'assurer une utilisation appropriée du système, il est essentiel de pouvoir mesurer et quantifier cette confiance [6]. Or, la traduction d'un phénomène d'une telle complexité sous la forme d'une variable unidimensionnelle est une tâche ardue loin de faire consensus. De plus, la complexité du cerveau humain est telle qu'il est extrêmement difficile, voire impossible, de le modéliser de façon à pouvoir prédire efficacement le comportement humain de manière générale [10]. Tout cela est sans compter les différences interpersonnelles entre les différents opérateurs qui peuvent réagir de manière complètement différente face à une même situation.

Il existe tout de même certaines pistes de solutions en ce qui a trait à la compréhension du système cognitif humain. En effet, des concepts tels que la charge mentale, la confiance, la confiance en soi et la conscience de la situation font l'objet de nombreuses études et aident à mieux comprendre le comportement humain. Cependant, ces éléments ne sont pas directement observables en pratique, et il est souvent nécessaire de se fier à des questionnaires pour en obtenir une évaluation subjective. Par ailleurs, il est possible d'observer le comportement de l'humain, ses performances et son utilisation de la machine afin de mieux comprendre ses schémas cognitifs, sans perturber ses activités. Des capteurs physiologiques plus ou moins invasifs permettent également de mesurer des manifestations physiques plus fines, tels que le rythme cardiaque, les mouvements oculaires et l'activité électrique du cerveau, qui fournissent des indications sur le comportement humain. Cependant, l'observation directe des processus cognitifs humains qui conduisent à tel ou tel comportement demeure difficile. Par conséquent, ce mémoire devra poser plusieurs hypothèses afin de pouvoir modéliser le comportement de manière générale, tout en gardant à l'esprit qu'il s'agit d'un modèle incomplet.

1.2 Partage de tâches en classification

Dans cette section, nous présentons le scénario particulier pour lequel la stratégie de collaboration adaptative entre l'opérateur humain et le système automatisé sera développée. Nous formulons ici le problème de manière générale, puisque la définition formelle du problème sera donnée au chapitre 3.

1.2.1 Description du scénario

Considérons un scénario dans lequel une équipe humain-machine doit effectuer des tâches de classification. Chaque classification peut être effectuée soit par l'automatisation ou soit par l'opérateur humain. Les tâches peuvent être considérées comme étant des objets ou des événements aléatoires indépendants les uns des autres qui prennent un état binaire. À chaque période, l'équipe reçoit donc un ensemble de tâches binaires à classer. Prenons par exemple de la détection d'objets dans un ensemble d'images, où l'objectif est de classer si chaque image contient ou non l'objet à identifier.

Nous considérons que l'automatisation observe d'abord l'ensemble des tâches. Ensuite, l'automatisation détermine quelles tâches devraient être effectuées par l'humain. Cette décision se base à la fois sur sa certitude de pouvoir classer correctement ou non une tâche observée, mais également sur sa connaissance du modèle de performance humaine en fonction de sa charge de travail. On parlera dans ce mémoire de « renvois de décisions » pour référer au fait que

l'automation décide de laisser à l'humain une tâche spécifique de classification.

Le problème se complexifie dans la mesure où l'humain peut décider de suivre ou non la décision d'allocation de tâches prise par le système automatisé. Sa volonté de suivre cette recommandation dépend de son niveau de confiance envers l'automation. Nous considérons qu'à la fin de chaque période de travail, un oracle permet de déterminer quelles tâches ont été correctement réalisées de sorte que l'équipe a conscience de ses performances globales. C'est cette performance qui vient influencer la confiance de l'humain envers l'automation en plus de son état de confiance au début de la période.

Le problème est donc de déterminer la politique de renvois de tâches, considérant la confiance humaine, qui sera suivie par l'automation afin de maximiser les performances de l'équipe.

Nous fournissons maintenant un exemple concret d'application de ce scénario.

Exemple 1. Un opérateur humain a comme tâche d'effectuer la surveillance aérienne d'un territoire. À l'aide d'un radar, il doit classifier les différentes cibles qui survolent le territoire à surveiller. Une cible est considérée comme étant soit hostile ou non hostile (état binaire).

À chaque période, l'opérateur a un nombre prédéterminé de cibles à classifier. Pour effectuer toutes les classifications, il est aidé par un système automatisé. Le système observe toutes les cibles qui doivent être classifiées et réfère celles pour lesquelles il est le plus incertain quant à sa capacité de les classifier correctement. Toutefois, la décision finale quant à la prise en charge des cibles revient à l'opérateur humain, qui peut choisir de prendre plus ou moins de tâches que celles recommandées par l'automation, en fonction de son niveau de confiance.

Pour réaliser chacune des classifications, l'opérateur se réfère à certaines règles apprises par expérience. La performance de l'opérateur dépend de sa charge de travail, puisque pour un nombre de tâches élevé à classifier, il possède moins de temps pour effectuer chacune des décisions conformes aux règles apprises. On considère donc que la probabilité d'effectuer une décision erronée va augmenter à mesure que le temps pour effectuer la décision diminue.

Pour visualiser cet exemple concret, nous supposons à titre d'exemple que les règles apprises par expérience par l'opérateur pourraient être traduites par un arbre décisionnel présenté à la Figure 1.1. Nous présentons également sur la même figure le radar utilisé pour détecter les cibles.

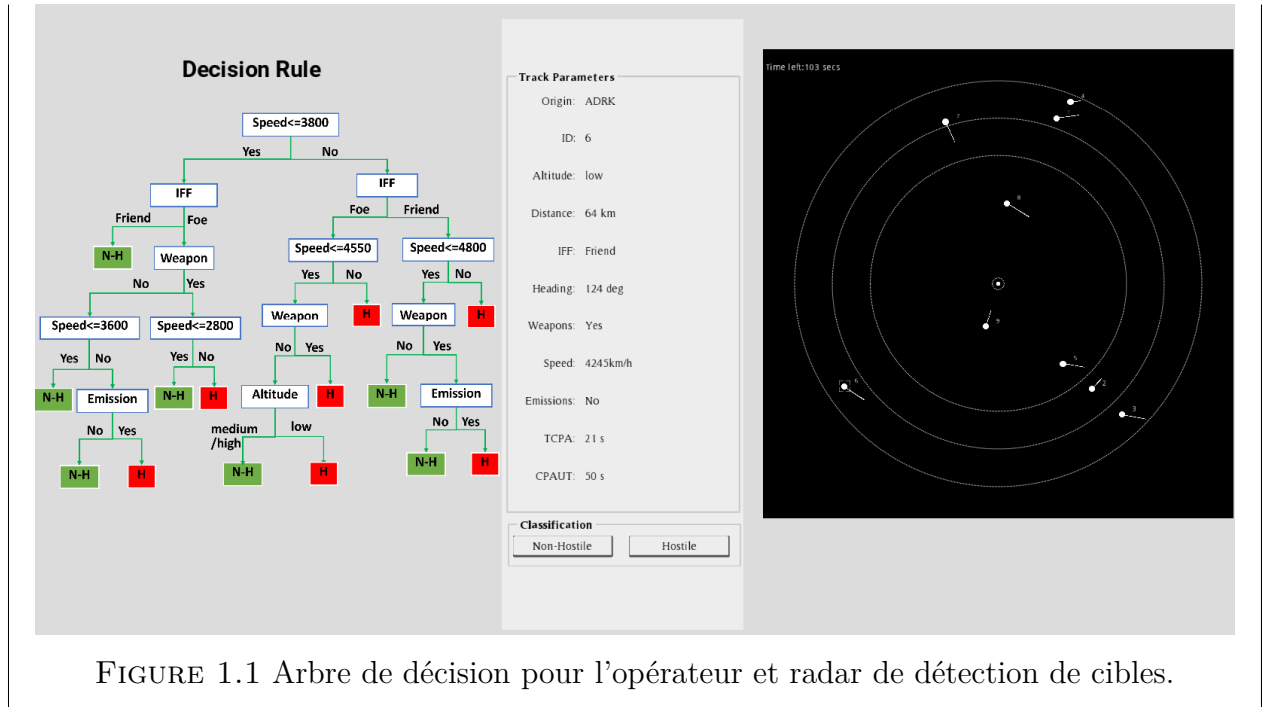


FIGURE 1.1 Arbre de décision pour l'opérateur et radar de détection de cibles.

1.2.2 Justification du scénario

Nous décidons de formuler un scénario axé sur la classification de tâches binaires puisque cela permet de bénéficier d'un cadre de travail suffisamment précis pour développer des résultats pertinents et applicables. De plus, ce scénario représente un premier exemple de collaboration relativement simple qui pourrait être enrichi ultérieurement selon diverses orientations, telles que l'introduction de classifications comprenant plus de deux catégories ou la considération d'un environnement non statique, parmi d'autres possibilités.

Le scénario formulé met en évidence les compétences spécifiques d'un décideur humain qui doit s'appuyer sur ses connaissances et son expertise pour prendre des décisions dans un environnement incertain. Traditionnellement, l'humain est considéré comme étant plus performant que les systèmes automatisés dans ce type de situation [11]. Cependant, avec l'émergence de nouveaux outils d'apprentissage machine, il est de plus en plus envisageable que l'automatisation puisse également exceller dans des problèmes de prise de décisions sous incertitude. Conséquemment, la collaboration entre l'humain et le système automatisé devient d'autant plus pertinente dans le contexte de la classification d'événements.

1.3 Objectifs de recherche

1.3.1 Développer une politique de renvois de tâches

Nous souhaitons développer une politique de renvois de tâches qui détermine de manière optimale les tâches décisionnelles spécifiques qui devraient être assignées à l'humain à chaque période. Cette politique devra considérer différents facteurs humains, tels que la confiance envers l'automation ainsi que les performances humaines en fonction de sa charge de travail.

1.3.2 Évaluer la politique d'allocation de tâches

Nous cherchons à évaluer la pertinence de cette approche en déterminant si la mise en œuvre d'une politique optimale présente un avantage réel par rapport à d'autres politiques sous-optimales plus simples. La comparaison avec d'autres politiques doit être effectuée en termes de performances globales de l'équipe, tout en prenant également en compte la facilité d'implémentation.

1.4 Plan du mémoire

Ce mémoire s'organise de la façon suivante. Le chapitre 2 présente une revue de littérature qui situe ce travail dans son contexte de recherche tout en fournissant des justifications quant à certains choix effectués. Le chapitre 3 présente une formulation complète du problème. Le chapitre 4, quant à lui, développe une politique de renvois de tâches optimale pour un état de confiance humaine parfaitement observable, alors que le chapitre 5 considère une observation imparfaite de la confiance. Nous présentons également dans ces chapitres des simulations numériques qui permettent d'évaluer la pertinence de la politique développée. Finalement, le dernier chapitre conclut le mémoire en présentant les limites et les potentielles améliorations du travail effectué.

CHAPITRE 2 REVUE DE LITTÉRATURE

Dans ce chapitre, nous introduisons et contextualisons les différents concepts qui sont exploités dans la formulation du problème. Plus particulièrement, nous présentons d’abord les concepts de charge de travail et de charge mentale ainsi que leurs impacts sur les performances humaines. Ensuite, la littérature concernant le concept de confiance envers l’automation ainsi que des différentes méthodes utilisées pour la modéliser est présentée. Enfin, nous étudions les différentes méthodes de collaboration entre un système automatisé et un opérateur humain en portant une attention particulière au scénario de renvois de décisions.

2.1 Charge de travail et charge mentale

2.1.1 Définitions

Bien qu’ils soient souvent liés, la charge de travail et la charge mentale sont deux concepts distincts qu’il est pertinent de définir lorsqu’on s’intéresse au facteur humain. Nous définissons dans ce mémoire la charge de travail comme étant simplement la quantité de travail à accomplir dans un temps imparti (*taskload*). Cette charge se réfère autant au travail physique que mental qui doit être effectué pour compléter la tâche. La charge de travail est donc influencée par des facteurs tels que le nombre de tâches, la complexité de chaque tâche, et les ressources physiques disponibles pour effectuer les tâches. Pour sa part, la charge mentale fait plutôt référence à la charge cognitive qui est associée à la réalisation des tâches (*mental workload*). Elle réfère à la pression mentale subie en résultat à une prise de décision, une gestion des tâches, ou une planification de la méthode de travail. C’est donc un construit hypothétique qui représente les ressources mentales nécessaires pour accomplir une tâche avec un certain niveau de performance [12]. On notera que la charge de travail a un impact considérable sur la charge mentale d’un opérateur humain bien qu’il ne s’agisse pas du seul facteur à considérer. En effet, la charge de travail est un déterminant exogène qui ne dépend pas de l’individu, alors que plusieurs déterminants endogènes propres à l’individu, tels que son expérience, ses ressources attentionnelles, son état physique, sont à prendre en compte dans l’évaluation de la charge mentale [13].

2.1.2 Mesure de la charge mentale

Différentes méthodes de mesure de la charge mentale ont été développées afin d’approfondir la compréhension de ce concept. La mesure adéquate de la charge mentale d’un humain revêt sa

pertinence du fait qu'elle est en grande partie responsable des performances de l'humain. En effet, une bonne gestion de la charge mentale d'un opérateur humain permettrait d'améliorer sa productivité, de réduire son stress, et ainsi diminuer ses erreurs [14]. Nous présentons maintenant quelques-unes des méthodes de mesure de la charge mentale les plus utilisées [15] :

1. Échelle d'évaluation subjective : L'opérateur est amené à évaluer sa propre charge mentale perçue par le biais d'un questionnaire (p. ex. *NASA-Task Load index* [12]) qui contient plusieurs échelles (p. ex. échelle de Likert [16]). Ces outils fournissent une estimation subjective de la quantité d'effort mental ressentie lors de l'exécution d'une tâche.
2. Mesures physiologiques : Les mesures physiologiques permettent d'obtenir des données objectives de la charge mentale. Par exemple, l'électroencéphalographie (EEG) peut être utilisée pour mesurer l'activité cérébrale et identifier certains patrons d'ondes cérébrales à différentes charges mentales [17]. La variabilité de la fréquence cardiaque (VFC) et la dilation des pupilles sont d'autres exemples de mesures physiologiques courantes pour mesurer la charge mentale.
3. Performances comportementales : La vitesse de réaction, la précision des réponses et le temps de réponse sont des indicateurs de performances comportementales qui permettent d'évaluer la charge mentale associée à une tâche. En effet, la variation de ces indicateurs de performances peut indiquer le changement du niveau de charge mentale ressenti par l'opérateur [15].

2.1.3 Modélisation de la charge mentale

Nous nous intéressons maintenant aux différents modèles quantitatifs de la charge mentale qui ont été proposés dans la littérature. Une attention plus particulière est accordée aux modèles de charge mentale basés sur une interaction avec un agent automatisé. Il convient de souligner que, dans la pratique, l'attention est principalement portée sur le modèle de performance humaine en relation avec la charge mentale, car l'objectif ultime est de maximiser les performances.

Une des notions à la base de la mesure et la modélisation de la charge se trouve à être la capacité totale de travail, soit les ressources cognitives globales disponibles. On considère que chaque individu a des ressources attentionnelles limitées qu'il peut utiliser pour effectuer une tâche. Plus la demande en ressources associées à une tâche est élevée, plus les ressources allouées à cette tâche seront sollicitées, diminuant ainsi les ressources disponibles. Au-delà d'un certain seuil, si les ressources demandées excèdent les ressources disponibles, on s'attend à une baisse de la performance de l'opérateur.

Dans [18], les auteurs proposent un modèle de charge mentale d'attention multiple qui définit différents systèmes de traitement de l'information utilisés pour différentes tâches cognitives (p. ex. perception visuelle ou auditive, traitement spatial ou verbal, réponse manuelle ou vocale). Chacun de ces systèmes a des ressources limitées, faisant en sorte que la sollicitation par une tâche d'un même système a pour effet d'augmenter la charge mentale et possiblement affecter les performances. Un autre modèle proposé dans [19] repose davantage sur le concept de charge cognitive qui se distingue entre la charge mentale intrinsèque (liée à la complexité inhérente de la tâche), la charge mentale extrinsèque (liée aux ressources cognitives pour accomplir la tâche) et la charge mentale essentielle (*germane cognitive load*) (liée à l'intégration des connaissances). Selon ce modèle, une charge mentale excessive entraînerait une surcharge cognitive pouvant nuire aux performances.

Les deux modèles précédents indiquent que des performances réduites peuvent être observées lorsque l'opérateur fait face à une charge mentale excessive. Parallèlement, il existe également un modèle largement cité dans la littérature qui suggère que les performances de l'opérateur peuvent également être affectées en cas de sous-charge cognitive. En effet, la loi de Yerkes et Dodson met en évidence que lorsque l'opérateur est confronté à une sous-charge, il peut se retrouver dans un état d'ennui, ce qui peut réduire sa capacité à réagir de manière efficace [20]. Plusieurs critiques ont toutefois été formulées à l'égard de ce modèle, notamment concernant sa simplification excessive et son application à des scénarios de tâches plus complexes [21]. Plus particulièrement, dans une optique de collaboration active entre un humain et un agent automatisé (c.-à.-d. que l'humain n'a pas comme unique tâche la surveillance du système automatisé), on considère plus souvent qu'autrement que la performance cognitive de l'humain est maximale lorsqu'il est sujet à une charge de travail minimale [22].

Par ailleurs, plusieurs modèles de charge mentale ont été développés pour analyser la dynamique de la collaboration humain-machine. C'est notamment le cas du travail de [23] qui propose un modèle de charge mentale à trois dimensions (temps imparti pour la tâche, distance perçue de l'objectif, niveau d'effort requis) pour une application de réallocation dynamique de tâches pour maintenir une charge de travail optimale pour l'opérateur. Une approche encore plus simple est présentée dans [24] où les auteurs évaluent la charge mentale humaine par un modèle linéaire robuste.

Finalement, nous portons notre attention sur le concept de schémas mentaux partagés (*Shared Mental Models* (SMM)) qui viennent de pair avec la charge mentale de l'opérateur dans le cas d'un scénario d'équipe collaborative. Les SMMs permettent à chaque agent de comprendre les rôles et objectifs des autres agents dans l'équipe afin de faciliter l'aspect collaboratif au sein de l'équipe et ainsi augmenter les performances globales [25]. Une meilleure compréhension

globale des objectifs de tout un chacun favorise une meilleure communication implicite au sein de l'équipe en plus de rendre plus fluide la coordination entre les agents. Par exemple, dans [26] l'agent automatisé utilise des modèles de Markov cachés pour évaluer la charge cognitive et le schéma mental de l'opérateur afin d'ajuster son comportement en fonction de son partenaire.

2.2 Confiance envers l'automation

La confiance est un facteur humain essentiel à considérer lors de la conception d'une équipe humain-machine, en particulier dans les scénarios où l'opérateur humain a la mainmise sur la machine. En effet, un opérateur qui n'aurait pas confiance envers les capacités d'un système automatisé à réaliser une tâche pourrait décider d'effectuer lui-même les tâches qui auraient dû être assignées à l'automation. Cependant, il a été discuté à la section précédente qu'une telle surcharge de travail peut au final nuire aux performances de l'humain et conséquemment aux performances globales de l'équipe [19]. Inversement, la surévaluation des capacités de la machine pourrait porter l'humain à délaissier certaines tâches pour lesquelles l'automation ne possède pas les compétences nécessaires, menant encore une fois à des performances sous-optimales, voire des accidents [27]. Dans le premier cas de figure, on parle de sous-exploitation de l'automation (*disuse*), alors qu'on utilise le terme surexploitation (*misuse*) pour décrire la deuxième situation [10]. La connaissance de l'état de confiance de l'opérateur humain rend donc possible l'adaptation des stratégies utilisées par le système automatisé afin de s'assurer d'une utilisation adéquate de l'automation par l'opérateur.

2.2.1 Définitions

Bien que la confiance soit un sujet largement étudié en psychologie, on s'intéresse plus précisément dans cette section à la définition de la confiance envers l'automation dans le cadre d'une collaboration humain-machine. La confiance envers l'automation est définie dans [28] comme étant l'attitude positive qu'un opérateur développe envers les systèmes automatisés, basée sur des croyances et des expériences passées, et qui influence la volonté de s'appuyer sur ces systèmes pour l'accomplissement de tâches critiques dans un environnement caractérisé par l'incertitude et la vulnérabilité. On parle également de la croyance positive qu'a l'opérateur envers les capacités, la fiabilité et l'intégrité d'un système automatisé pour accomplir une tâche spécifique de manière sûre, efficace et conforme aux attentes [29]. La confiance envers l'automation est donc un processus psychologique qui permet d'éviter à l'humain une utilisation coûteuse de ressources cognitives lors de prises de décision en collaboration avec une machine complexe. Cependant, on comprendra que ce n'est pas tant la confiance que

la dépendance envers la machine (*reliance*) qu'on cherche à maintenir à un niveau adéquat. Un opérateur qui n'aurait pas confiance envers un système automatisé, mais qui l'utiliserait de manière adéquate ne nuirait pas pour autant aux performances de l'équipe. Toutefois, on a en pratique une forte corrélation entre la confiance de l'humain et sa propension à se fier à l'automation, donc sa dépendance à l'automation [28, 30]. C'est pourquoi ce mémoire fait référence à la confiance envers l'automation en sous-entendant la dépendance de l'humain envers la machine, soit sa propension à respecter les décisions prises par le système automatisé.

2.2.2 Facteurs influençant la confiance envers l'automation

La confiance étant un concept psychologique complexe, il serait difficile de faire mention de tous les facteurs qui l'impactent. La littérature fait toutefois mention de principaux facteurs d'influence. Parmi ceux-ci on retrouve la confiance en soi. En effet, la confiance envers l'automation est souvent liée à la confiance que l'opérateur accorde à ses propres compétences et capacités à interagir avec cette automatisation. Lorsqu'un opérateur a confiance en ses propres capacités à utiliser et à s'adapter à l'automatisation, il est plus enclin à développer une confiance positive envers la machine elle-même [28, 31]. Le niveau de transparence de l'automation est un autre facteur ayant un impact considérable sur la confiance de l'opérateur. Un opérateur ayant accès à des informations claires et compréhensibles sur le fonctionnement du système automatisé est plus enclin à faire confiance aux décisions de la machine [10, 32]. Évidemment, la confiance de l'opérateur envers l'automation est également impactée de manière importante par les compétences, les performances et la fiabilité perçue de l'automation [33]. Un système automatisé performant qui commet peu d'erreurs inspire donc davantage de confiance chez l'humain [28, 34]. Finalement, d'autres facteurs d'influence peuvent également être pris en compte tel que la confiance initiale de l'opérateur [35], la pression temporelle [36], la difficulté perçue de la tâche [37], la charge mentale de travail [10], ainsi que des considérations culturelles propres à chaque individu [38].

2.2.3 Sur-confiance

À présent, nous abordons de manière plus précise la notion de sur-confiance afin de faciliter la discussion quant à l'hypothèse de son omission dans ce mémoire.

Tout d'abord, nous différencions le concept de sur-confiance (*over-trust*) et de complaisance (*complacency*). La sur-confiance se réfère à une situation de confiance excessive envers l'automatisation dans laquelle l'opérateur attribue à l'automatisation des capacités supérieures à ses capacités réelles. Cela peut se produire en raison de biais cognitifs, d'expériences passées

positives ou de croyances préexistantes. La sur-confiance peut donc conduire à une dépendance (*reliance*) excessive envers l'automatisation. La complaisance, quant à elle, est un état de satisfaction ou de confort excessif qui se développe lorsque les utilisateurs deviennent trop confiants envers l'automatisation et cessent de se préoccuper activement de la tâche à accomplir [10]. L'opérateur est donc plus à risque de tomber dans un état de complaisance lorsqu'il doit effectuer des tâches en parallèle à la surveillance de l'automation. Dans les deux cas, on peut s'attendre à une baisse de vigilance, une attention réduite et une diminution des performances de l'humain.

L'auteur de [39] identifie trois types de sur-confiance :

1. L'opérateur a un manque ou une réduction de vigilance face à un mal fonctionnement du système automatisé.
2. L'opérateur s'attend à ce que l'automation effectue une action qui se trouve à l'extérieur des capacités spécifiées par la conception du système.
3. L'opérateur a des attentes face aux capacités de l'automation qui ne correspondent pas aux capacités réelles du système.

Nous remarquons donc que la sur-confiance est essentiellement une conséquence d'une compréhension erronée des capacités réelles du système automatisé.

2.2.4 Modèles quantitatifs de la confiance

Différents modèles quantitatifs de confiance ont été développés dans la littérature. On considère que l'aspect quantitatif d'un modèle confère un avantage significatif par rapport à un modèle reposant sur des mesures qualitatives, principalement en raison de la prédictibilité qu'offrent les modèles quantitatifs. En effet, il existe des méthodes de mesure de la confiance qui sont similaires à celles présentées pour mesurer la charge mentale d'un individu. Les échelles de confiance (p. ex. *Trust Perception Scale-HRI* [40]) et les questionnaires [41, 42] sont des outils qui, avec les mesures physiologiques et comportementales [43, 44], permettent d'avoir une mesure qualitative de l'état de confiance humain. Néanmoins, le caractère invasif et subjectif des outils qui nécessitent une auto-évaluation de l'opérateur durant sa tâche attirent plusieurs critiques à l'égard de tels modèles basés sur des mesures qualitatives qui sont suffisantes pour pouvoir formuler des prédictions quantitatives.

Avant de présenter les différents modèles quantitatifs, il est important de rappeler que la pertinence et la validité d'un modèle sont fortement influencées par le scénario d'application précis dans lequel il est utilisé. En effet, étant donné la multitude de facteurs qui influencent la confiance dont il a été question à la section précédente, il semble impossible de penser qu'un

seul modèle quantitatif serait pertinent dans tous les scénarios de collaboration humain-machine [45]. La présente revue des différents modèles existants ne se veut donc pas une comparaison de la performance des différents modèles, qui ont tous leur pertinence dans des scénarios d’application précis.

Tout d’abord, plusieurs modèles de régression ont été développés afin de prédire l’évolution de la confiance. C’est notamment le cas des auteurs de [46] qui proposent un modèle linéaire pour prédire l’utilisation de l’automatisation qui sera faite par l’opérateur en fonction des performances de l’équipe et de l’état de confiance de l’opérateur à la période précédente. Un autre modèle similaire dans [1] décrit en plus les différents processus psychologiques associés à l’état de confiance de l’humain afin de prédire son utilisation de l’automatisation.

Afin de considérer les caractéristiques dynamiques liées à l’état de confiance humain, d’autres chercheurs ont proposé des modèles déterministes. Parmi ceux-ci, on retrouve dans [34] un modèle ARMA (*AutoRegressive Moving Average*) qui permet de prédire la stratégie d’allocation de tâche de l’opérateur en fonction de sa confiance en soi et sa confiance envers l’opérateur. Dans [47], on modélise le niveau de confiance humain en fonction de l’expérience, de la confiance cumulative et des attentes de l’opérateur.

Ensuite, l’aspect incertain et stochastique inhérent au processus cognitif de la confiance a conduit au développement de modèles probabilistes. Les auteurs de [48] ont développé le modèle OPTIMo (*Online Probabilistic Trust Inference Model*) : un réseau bayésien dynamique qui infère la confiance de l’humain à partir de ses interactions avec la machine. Dans [49], un modèle probabiliste basé sur des arguments qualitatifs est utilisé afin de mieux représenter l’incertitude des prédictions que pourrait avoir un modèle déterministe. En ce qui concerne d’autres approches probabilistes, plusieurs chercheurs ont modélisé le comportement de confiance humaine en utilisant des modèles de Markov, plus précisément des modèles de Markov cachés (*Hidden Markov Model* (HMM)) [50, 51]. Bien que les HMM puissent être utilisés pour inférer l’intention de l’opérateur et pour intégrer l’incertitude liée au comportement humain, par exemple dans [52, 53], ils n’incluent toutefois pas les effets des entrées et des actions des systèmes autonomes qui influencent le comportement humain. Il est possible d’utiliser des modèles basés sur les processus de décision markoviens (*Markov Decision Process* (MDP)) pour prendre en compte l’influence des actions de l’automatisation dans la modélisation du comportement humain [54, 55], bien que de tels modèles ne tiennent pas compte du fait que la confiance humaine est un processus cognitif qui n’est pas directement observable. On peut donc avoir recours à une extension des HMMs et des MDPs, soit les processus de décision markoviens partiellement observables (*Partially Observable Markov Decision Process* (POMDP)). Les POMDPs proposent un cadre de travail qui permet de tenir compte à la

fois de l'influence des actions de l'automation ainsi que du caractère non observable dans la modélisation de la confiance humaine [56, 57].

Finalement, une autre approche consiste à modéliser la confiance par l'apprentissage d'un réseau de neurones artificiels. L'intérêt croissant pour différents outils d'apprentissage machine a permis le développement de plusieurs modèles qui prédisent la confiance de l'humain en identifiant les différents paramètres d'influence grâce à de grandes bases de données. On peut citer par exemple [58] et [59].

2.3 Conception de stratégies de collaboration humain-machine

Nous nous intéressons dans cette section aux diverses stratégies de collaboration développées dans la littérature afin d'améliorer les performances d'équipes humain-machine. Nous ne cherchons pas à présenter de manière exhaustive l'ensemble de méthodes publiées dans la littérature, mais plutôt à cibler les stratégies les plus étudiées qui se basent sur un modèle quantitatif du comportement humain. Encore une fois, il est important de relever qu'une stratégie donnée est pertinente dans le scénario spécifique dans lequel elle a été développée et que certaines stratégies ne peuvent être utilisées dans certaines applications.

2.3.1 Stratégie d'allocation de contrôle

Une première stratégie de collaboration humain-machine est de laisser l'humain prendre la décision de passer en contrôle manuel ou automatique. Dans le cas où l'humain choisit un contrôle automatique, il se retrouve dans une situation de surveillance de l'automation. L'intérêt d'une telle stratégie est d'éviter les situations de sous-exploitation ou de surexploitation de la machine en laissant l'opérateur gérer sa propre charge mentale. Nombre de modèles quantitatifs linéaires de la confiance humaine envers l'automation sont justement basés sur un tel scénario d'allocation de contrôle [1, 28, 46]. Cette stratégie peut être bonifiée en permettant à l'automation d'avertir l'opérateur lorsqu'elle considère qu'il serait avantageux pour l'équipe que l'opérateur change le contrôle du système. Un exemple de stratégie d'allocation de contrôle dans un contexte de téléopération d'un robot mobile est donné dans [60] où les auteurs montrent une performance globale accrue et une charge mentale réduite.

2.3.2 Stratégies de collaboration influençant le comportement de la machine

Les stratégies d'allocation de contrôle décrite à la section précédente peuvent se révéler efficaces pour éviter directement les situations de sous-exploitation ou de surexploitation. Cependant, selon [61], il n'est pas nécessaire de faire varier de manière drastique le niveau

d'automation (manuel ou automatique) afin de concevoir une stratégie de collaboration efficace. Nous présentons deux types de collaboration pour lesquelles l'automation varie son comportement en considérant la confiance de l'opérateur.

Un premier type d'une telle stratégie de collaboration serait le changement de comportement de l'automation via des états dégradés. En fonction de l'état de confiance de l'opérateur, l'automation passerait donc d'un état de fonctionnement normal à un état de fonctionnement dégradé, et vice-versa. Par exemple, dans [62], où une équipe humain-machine doit compléter une tâche de manipulation d'objets, le robot considère la confiance de l'humain avant de décider s'il devrait être pro actif dans la décision de la trajectoire optimale ou s'il serait préférable de laisser davantage de charge décisionnelle à l'humain. Les auteurs de [63] développent un modèle automatisé d'un robot qui va chercher à rétablir la confiance (*trust-seeking robot*) en adoptant un comportement plus prévisible (vitesse réduite et mouvements plus lisses) en cas de sous-confiance de l'opérateur.

Dans les deux exemples cités précédemment, le système automatisé manifestait via un état dégradé un comportement qui ne reflétait pas pleinement ses capacités réelles lorsqu'il y avait une faible confiance de la part de l'opérateur. Un autre type de stratégie permet d'intégrer de manière continue la confiance de l'opérateur dans le comportement de l'automation sans toutefois dégrader les performances de la machine. Par exemple, dans [64], un bras robotique doit déplacer plusieurs objets associés à des risques et des récompenses différents, tout en étant supervisé par un opérateur humain capable d'interrompre la tâche du bras robotique s'il anticipe un échec éventuel. Les auteurs proposent une stratégie qui indique quel objet devrait être déplacé par le robot en fonction de la confiance humaine. Il s'agit donc d'un scénario dans lequel le comportement de l'automation va être influencé par la confiance humaine sans toutefois être dégradé pour autant.

2.3.3 Stratégies de renvois de tâches

Discutons à présent des stratégies qui considèrent également l'impact de la charge de travail en plus du niveau de confiance de l'opérateur. De telles stratégies sont développées dans des cadres de travail où un ensemble de tâches doit être effectué dans une période de temps souvent restreinte. On parle de renvois de décisions (*decision referrals*) dans la mesure où on laisse à l'automation la décision d'effectuer elle-même la tâche ou de la référer à l'humain. Une telle décision se base donc en partie sur un modèle de charge de travail de l'humain, à savoir que les performances de l'opérateur diminueront s'il doit effectuer un nombre trop important de tâches, ainsi que sur un modèle de confiance, dans les cas où l'humain conserve l'autorité sur la machine et qu'il peut décider de ne pas suivre ses décisions. De plus, la déci-

sion de renvois de tâches est aussi influencée de manière importante par les caractéristiques intrinsèques de la tâche à effectuer qui sont d’abord observées par l’automation. Dans [55], les auteurs proposent une stratégie d’allocation de tâches adaptative basée sur un modèle quantitatif de confiance. L’automation doit déterminer la proportion des tâches qui devraient être accomplies par l’humain en fonction de sa confiance et en connaissant son modèle de performance par rapport à sa charge de travail. Les auteurs de [4] vont s’intéresser davantage à la nature spécifique de chaque tâche afin de formuler des seuils d’observation au-delà desquels l’automation devrait référer la tâche à l’opérateur. La stratégie d’allocation de tâches présentée dans [3], qui est d’un intérêt particulier pour le travail présenté dans ce mémoire, permet à l’automation d’observer toutes les tâches avant de décider lesquelles devraient être complétées par l’humain en considérant le modèle de performance humain, mais également sa propre capacité de compléter avec succès les tâches qu’elle décide de conserver. Le présent travail a comme objectif d’ajouter un modèle de confiance simple à la formulation du problème proposée dans [3].

Finalement, des stratégies basées sur l’apprentissage machine ont également été développées dans ce même esprit de renvois de décisions à l’opérateur. On parle dans la littérature d’apprentissage à référer (*learning to defer*). Dans [65], le système automatisé apprend à la fois un modèle de classification de tâches et un modèle de renvois de tâches en se basant sur ses observations des décisions prises par l’humain. Les auteurs de [66] proposent également une démarche visant à ce que l’automation apprenne quelles tâches devraient être effectuées par un opérateur en tentant de maximiser l’équité et la précision dans la prise de décision de sorte à limiter les biais humains. D’autres travaux, tels que ceux cités dans l’article [67], ont également démontré que la maximisation de la performance d’un système automatisé basé sur l’intelligence artificielle au sein d’une équipe humain-machine ne conduit pas toujours à une optimisation des performances globales.

2.4 Conclusion de la revue de littérature

Ce chapitre a joué un rôle crucial en établissant le contexte de ce mémoire et en mettant en évidence les différents éléments de littérature nécessaires pour justifier le choix des modèles utilisés. Plus spécifiquement, il a abordé les concepts de charge mentale, charge de travail et confiance, qui sont trois facteurs essentiels à considérer lors de l’évaluation des performances d’un opérateur au sein d’une équipe humain-machine. Différentes stratégies de collaboration pour une telle équipe ont également été présentées, en accordant une attention particulière à une stratégie de renvois de décisions, qui sera étendue dans le reste du travail. En effet, la solution présentée dans ce mémoire s’inscrit dans cette même démarche puisqu’il s’agit de

proposer une stratégie de renvois de décisions qui se base sur un modèle quantitatif de la charge de travail humaine en plus de considérer un modèle de la confiance envers l'automatisation. Le développement d'une telle stratégie de renvois de décisions permet d'évaluer la pertinence à l'implémenter en comparant ses performances à d'autres stratégies de collaboration humain-machine existantes.

L'originalité de la solution proposée dans ce mémoire réside donc dans la combinaison des éléments suivants :

1. Le contexte spécifique des renvois de décisions au sein d'une équipe humain-machine ;
2. La prise en compte de la charge de travail et de son impact sur les performances humaines ;
3. L'incorporation de la confiance en tant que facteur crucial dans la stratégie de renvois de décisions.

Bien que des travaux existent déjà sur chacun de ces aspects, la contribution novatrice de ce mémoire réside dans leur association au sein d'une même étude, ce qui en fait son intérêt majeur.

CHAPITRE 3 FORMULATION DU PROBLÈME

Dans ce chapitre, nous développons une formalisation du problème ainsi qu'une description des différents modèles numériques utilisés dans les simulations du scénario décrit.

3.1 Description formelle du scénario

Considérons une équipe formée d'un humain et d'un système de support à la décision (SSD). À chaque période discrète t sur un horizon fini T , l'équipe reçoit un ensemble de K tâches binaires à classifier dénoté $\mathcal{K} = \{1, \dots, K\}$. Chaque tâche $k \in \mathcal{K}$ a un état binaire caché $H_{k,t} \in \{\mathcal{H}_0, \mathcal{H}_1\}$ avec une probabilité connue p d'être $\mathcal{H}_1$ (i.e., un événement intéressant). Pour chaque tâche k à la période t , le SSD et l'humain reçoivent chacun une observation aléatoire dénotée $Y_{k,t}^a \in \mathcal{Y}^a$ et $Y_{k,t}^h \in \mathcal{Y}^h$ respectivement, qui dépend du vrai état $H_{k,t}$ de la tâche observée. Les états $H_{k,t}$ pour différents k sont indépendants et, pour un $H_{k,t}$ donné, les observations $Y_{k,t}^a$ et $Y_{k,t}^h$ sont conditionnellement indépendantes. Toutes les tâches sont identiquement distribuées.

On dénote $D_{k,t}$ la décision de classification finale de la tâche $k \in \mathcal{K}$ de l'ensemble à la période t , par tout preneur de décision, humain ou automation. Le coût encouru pour chacune des décisions est défini comme étant

$$C(D_{k,t}, H_{k,t}) = \begin{cases} c_{tp} & \text{si } (D_{k,t}, H_{k,t}) = (\mathcal{H}_1, \mathcal{H}_1), \\ c_{fp} & \text{si } (D_{k,t}, H_{k,t}) = (\mathcal{H}_1, \mathcal{H}_0), \\ c_{tn} & \text{si } (D_{k,t}, H_{k,t}) = (\mathcal{H}_0, \mathcal{H}_0), \\ c_{fn} & \text{si } (D_{k,t}, H_{k,t}) = (\mathcal{H}_0, \mathcal{H}_1). \end{cases} \quad (3.1)$$

Notons que c_{fp} et c_{fn} sont les coûts associés à un faux positif et à un faux négatif respectivement. Nous formulons l'hypothèse naturelle suivante.

Hypothèse 1. *Les coûts satisfont $c_{fn} \geq c_{tp}$ et $c_{fp} \geq c_{tn}$.*

En effet, il est plus que raisonnable d'assumer qu'une erreur de classification (c_{fp} ou c_{fn}) engendre un coût plus élevé pour l'équipe qu'une tâche ayant été correctement classifiée (c_{tp} ou c_{tn}). On pourrait même prendre $c_{tp} = c_{tn} = 0$ afin de simplifier la formulation du problème.

À chaque période t , le SSD considère d'abord ses observations $Y_{1:K,t}^a$ pour l'ensemble de tâches de la période courante t , ainsi que sa perception de l'état de confiance de l'humain τ_t , que

nous assumons comme étant parfaitement observable dans le chapitre 4 et imparfaitement observable dans le chapitre 5. À partir de ces informations, le SSD recommande un nombre de tâches w_t^r que l'humain devrait prendre et que nous avons défini comme étant les tâches *renvoyées*. La classification des tâches restantes dans l'ensemble de tâches serait donc effectuée par le SSD. En plus du nombre de tâches w_t^r recommandé à l'humain, l'automatisme fournit également une liste ordonnée des tâches, spécifiée par une permutation $\pi \in \mathcal{S}_K$ où $\mathcal{S}_K$ est un groupe symétrique sur K éléments. On considère que l'humain effectue ses tâches dans l'ordre de la liste π spécifiée.

Nous supposons que la confiance de l'humain τ_t à chaque période peut être classifiée comme étant soit un état de confiance normal ($\tau_t = \mathbf{o}$) ou un état de sous-confiance ($\tau_t = -$). La justification de ce choix d'état est donnée à la section 3.2.1. La confiance de l'humain évolue de manière stochastique au fil du temps et est influencée à la période $t + 1$ à la fois par l'état de confiance τ_t à la période t et par le coût total de classification réalisé par l'équipe pour les tâches de l'ensemble $\mathcal{K}$ à la période t . On suppose que le modèle stochastique est stationnaire et qu'il peut donc être complètement caractérisé par la fonction

$$f_\tau(\tau, c) := \mathbb{P}[\tau_{t+1} = \mathbf{o} | \tau_t = \tau, C_t = c], \quad (3.2)$$

pour $\tau \in \{-, \mathbf{o}\}$, $c \in \mathbb{R}$. L'état de confiance de l'humain donne une indication quant à sa volonté de suivre la recommandation w_t^r de l'automatisme et de laisser le SSD s'occuper de la classification de plus de tâches. Concrètement, en fonction de son état de confiance, l'humain peut choisir de prendre plus ou moins de tâches, dénotées w_t^s , que le nombre référé w_t^r par l'automatisme. Au final, le nombre de tâches que l'humain a à classifier dans la liste π fournie par l'automatisme est $w_t^h = w_t^r + w_t^s$. L'humain classifie les tâches $\pi_1, \dots, \pi_{w_t^h}$ dans l'ordre spécifié de la liste π . Le SSD s'occupe de la classification des $K - w_t^h$ tâches restantes.

L'objectif est d'identifier la stratégie optimale pour l'automatisme afin de déterminer à chaque période une liste ordonnée de tâches π_t et une charge de travail suggérée w_t^r afin de minimiser les coûts de classification sur un horizon de temps fini T .

3.1.1 Impact de la confiance

Le modèle proposé pour la réponse humaine w_t^s suppose qu'elle est moins «prévisible» en situation de sous-confiance qu'en confiance normale. De façon plus concrète, réécrivons

$$w_t^s = \mathbb{E}[w_t^s | \tau_t, w_t^r] + \tilde{w}_t^s \quad (3.3)$$

qui définit la déviation $\tilde{w}_t^s$ par rapport à la moyenne (conditionnelle à la connaissance de τ_t , de telle sorte que $\mathbb{E}[\tilde{w}_t^s | \tau_t, w_t^r] = 0$). Nous formulons l'hypothèse suivante.

Hypothèse 2. *À tout instant $0 \leq t \leq T$, pour toute charge de travail w_t^r suggérée initialement par le SSD, nous supposons que pour toute fonction convexe φ ,*

$$\mathbb{E}[\varphi(\tilde{w}_t^s) | \tau_t = \mathbf{o}, w_t^r] \leq \mathbb{E}[\varphi(\tilde{w}_t^s) | \tau_t = -, w_t^r],$$

à condition que l'espérance existe.

En d'autres mots, la distribution de $\tilde{w}_{t,\tau}^s$ étant donné $\tau_t = \mathbf{o}$ est *plus petite dans le sens convexe* que la distribution de $\tilde{w}_{t,\tau}^s$ étant donné $\tau_t = -$. L'ordre convexe est une relation d'ordre partiel entre les variables aléatoires [68] qui capture l'intuition selon laquelle une variable aléatoire plus grande dans cet ordre est plus incertaine. En particulier, cela implique la relation

$$\text{Var}[\tilde{w}_t^s | \tau_t = \mathbf{o}, w_t^r] \leq \text{Var}[\tilde{w}_t^s | \tau_t = -, w_t^r]$$

entre les variances, en prenant $\varphi(x) = x^2$ dans la définition.

Nous justifions cette hypothèse simplement par la définition des états de sous-confiance et de confiance normale. C'est-à-dire qu'il est plus probable que l'opérateur en état sous-confiance ne suive pas la recommandation du SSD (d'où une réponse w^s plus incertaine et une variance plus grande) que lorsqu'il se trouve en état de confiance normale.

Notons finalement que w^s peut prendre les valeurs $\{-w^r, \dots, K - w^r\}$. En effet, l'opérateur a au minimum aucune tâche (soit $w^h = w^r - w^r$) et au maximum toutes les tâches ($w^h = w^r + K - w^r$). La distribution de w^s est donc dépendante de w^r .

3.1.2 Performance humaine

Comme le montre [3], pour que le SSD puisse référer les tâches appropriées vers l'opérateur, il doit former une probabilité postérieure sur les décisions de classification de l'humain. Il s'avère qu'il suffit de connaître les modèles des probabilités de classification de vrais positifs et de faux positifs de l'humain, c'est-à-dire,

$$\begin{aligned} P_{tp}(w) &= \mathbb{P}[D_k = \mathcal{H}_1 | \mathcal{H}_k = \mathcal{H}_1, w_t^h = w], \\ P_{fp}(w) &= \mathbb{P}[D_k = \mathcal{H}_1 | \mathcal{H}_k = \mathcal{H}_0, w_t^h = w], \end{aligned}$$

que nous supposons indépendants de la tâche k par hypothèse. Nous supposons que ces probabilités dépendent de la charge de travail, car avoir plus de tâches tend à diminuer les

performances pour chaque tâche tel que discuté à la section 2.1.3. Supposons que les fonctions $w \mapsto P_{tp}(w)$ et $w \mapsto P_{fp}(w)$ peuvent être obtenues à partir d'expériences de calibration préalables avec l'opérateur. Nous formulons l'hypothèse naturelle suivante.

Hypothèse 3. P_{tp} et P_{fp} sont respectivement non-croissant et non-décroissant avec w , c'est-à-dire pour tout $0 \leq w \leq K - 1$,

$$P_{tp}(w + 1) \leq P_{tp}(w), \quad P_{fp}(w + 1) \geq P_{fp}(w).$$

En effet, nous considérons que la probabilité que l'opérateur classe correctement (incorrectement) une certaine tâche va décroître (augmenter) à mesure qu'il devient plus chargé et, qu'au mieux, elle restera la même. Toutefois, il est important de noter que ceci n'est pas toujours observé expérimentalement tout dépendamment du scénario spécifique étudié [3]. Cette hypothèse est utilisée au chapitre suivant lorsque nous cherchons à démontrer l'optimalité d'une politique de renvois de tâches myope.

3.2 Formulation du problème en MDP

Formulons dans un premier temps le problème sous la forme d'un processus de décision markovien (MDP), en considérant que la confiance est un état qui puisse être complètement observé. La résolution de ce problème sous cette hypothèse est donnée au chapitre 4. Un MDP est défini par un quadruplet $\{S, \mathcal{A}, P, R\}$, soit un espace d'état S , un espace d'action $\mathcal{A}$, une fonction de transition $P_a(s, s')$, et une fonction de coût $R_a(s, s')$.

3.2.1 Définition de l'état

L'état du MDP est l'état de confiance de l'opérateur τ_t à l'instant t , qui peut prendre les valeurs $\tau_t = \mathbf{o}$ pour un état de confiance normal, et $\tau_t = -$, pour un état de sous-confiance.

Nous jugeons qu'il est raisonnable d'omettre un état de sur-confiance dans ce scénario en considérant la discussion concernant les raisons qui mènent à un état de sur-confiance présentée à la section 2.2.3. Nous considérons que le comportement d'un opérateur en état de sur-confiance qui sur-évaluerait les capacités du SSD serait de simplement se fier aveuglement à la recommandation w_t^r de l'automatisme qu'il estimerait comme étant exacte. Il y aurait donc encore moins d'incertitude dans la réponse w_t^s de l'humain lorsqu'il se trouverait dans cet état. Le scénario proposé ne laisse donc pas place à une perte de vigilance ou de la complaisance étant donné que l'opérateur a tout de même des tâches spécifiques à effectuer

à chaque période. Ainsi, on peut omettre la considération de cet état afin de simplifier la résolution du problème.

Nous choisissons également un espace d'état réduit qui comprend seulement deux états afin de faciliter le calcul analytique de la politique d'action optimale, bien qu'il est possible d'avoir un état discrétisé représentant un spectre de confiance entre l'état de sous-confiance et l'état de confiance normal.

3.2.2 Définition de l'action

L'action du MDP comprend la liste ordonnée de tâches π_t à suivre par l'opérateur ainsi que la charge de travail suggérée w_t^r , soit les tâches renvoyées à l'humain. Nous définissons l'action du MDP comme étant $a_t = \{\pi_t, w_t^r\}$. Nous fournissons un exemple simple d'une action que pourrait être prise par l'automation.

Exemple 2. Soit un ensemble de $K = 5$ tâches dénoté $\mathcal{K} = \{1, 2, 3, 4, 5\}$. Après avoir observé toutes les tâches, l'automation fournit une liste $\pi = \{\pi_1, \pi_2, \pi_3, \pi_4, \pi_5\}$. Pour notre exemple, cette liste pourrait être $\pi = \{2, 3, 1, 5, 4\}$. Avec cette liste, l'automation fournit également un nombre de tâches qu'il suggère à l'humain, par exemple $w^r = 2$. Après avoir reçu la liste et la recommandation de l'automation, l'humain peut décider de prendre plus ou de moins de tâches, qu'on dénote w^s . Considérons que l'humain décide de prendre une tâche de plus (i.e. $w^s = 1$). Sa charge de travail totale serait donc $w^h = w^r + w^s = 3$ et les tâches qu'il aurait à classifier seraient les trois premières de la liste π , soit $\{2, 3, 1\}$.

Il aurait également été possible de considérer un scénario pour lequel la liste de tâches à classifier par l'humain est obtenue une fois le nombre de tâches supplémentaires w_t^s connu. Toutefois, en pratique, ce scénario est plus difficilement observable puisqu'il peut être complexe pour un opérateur d'estimer au début de la période le nombre exact de tâches qu'il compte effectuer. Le fait de fournir une seule liste en considérant la distribution de w_t^s reflète donc mieux la réalité et justifie le choix d'action du SSD.

3.2.3 Définition de la fonction de coût

La fonction de coût du MDP est décrite par le coût espéré de classification réalisé à chaque période, qu'on cherche à minimiser. Pour la définir, nous utilisons le raisonnement proposé par [3] qui définit le coût de classification de l'automation ainsi que celui de l'opérateur. Afin

d'alléger la notation et puisque nous définissons le coût espéré pour une seule période, nous omettons l'index t dans cette section.

Coût de classification de l'automaton

À partir de ses observations $Y_{1:K}^a$ de l'ensemble des tâches $\mathcal{K}$, l'automaton peut calculer une probabilité postérieure $p_{1:K}^a$ pour chaque tâche d'être un événement intéressant, c'est-à-dire,

$$p_k^a := \mathbb{P}[H_k = \mathcal{H}_1 | Y_k^a], \quad \text{pour } 1 \leq k \leq K. \quad (3.4)$$

Le coût espéré pour l'automaton en prenant la décision de classification D_k est

$$C_a(D_k, p_k^a) = p_k^a C(D_k, \mathcal{H}_1) + (1 - p_k^a) C(D_k, \mathcal{H}_0) \quad (3.5)$$

où $C(D_k, \mathcal{H}_i), i \in \{0, 1\}$ est défini en (3.1).

Donc, la décision optimale D_k de l'automaton pour la classification d'une tâche k est

$$D_k = \begin{cases} \mathcal{H}_0, & \text{si } C_a(\mathcal{H}_0, p_k^a) \leq C_a(\mathcal{H}_1, p_k^a), \\ \mathcal{H}_1, & \text{autrement,} \end{cases}$$

et le coût espéré associé est

$$C_a(p_k^a) = \min\{C_a(\mathcal{H}_0, p_k^a), C_a(\mathcal{H}_1, p_k^a)\}.$$

Nous formulons l'hypothèse suivante concernant la probabilité postérieure p_t^a .

Hypothèse 4. À chaque période $0 \leq t \leq T - 1$, la probabilité de passer à un état de confiance normal étant donné un coût réalisé c_t et un état de confiance τ_t est conditionnellement indépendante de la probabilité postérieure p_t^a , de l'action de l'automaton a_t et de la réponse de l'humain w_t^s , i.e.,

$$\mathbb{P}[\tau_{t+1} = \mathbf{o} \mid \tau_t, p_t^a, c_t, a_t, w_t^s] = \mathbb{P}[\tau_{t+1} = \mathbf{o} \mid \tau_t, c_t].$$

Cette hypothèse capture l'intuition selon laquelle, une fois le coût de classification d'une période connu, la probabilité postérieure des tâches observées par l'automaton ainsi que son action n'apportent aucune d'information supplémentaire quant à l'évolution de la confiance humaine.

Coût de classification de l'humain

L'espérance du coût de laisser à l'opérateur classifier une tâche k , en considérant que sa charge de travail est w^h , est donné par

$$\Gamma_h(p_k^a, w^h) = p_k^a [P_{tp}(w^h)c_{tp} + (1 - P_{tp}(w^h))c_{fn}] + (1 - p_k^a) [P_{fp}(w^h)c_{fp} + (1 - P_{fp}(w^h))c_{tn}]. \quad (3.6)$$

Ce coût espéré peut être calculé par l'automation puisque nous supposons connus les modèles de performance humaine décrits par $P_{tp}(\cdot)$ et $P_{fp}(\cdot)$.

Coût de classification de l'équipe

Définissons maintenant les indices G qui correspondent au gain à envoyer une tâche à l'opérateur plutôt que de laisser cette tâche être classifiée par le SSD, i.e.,

$$G(p_k^a, w^h) := C_a(p_k^a) - \Gamma_h(p_k^a, w^h), \quad 1 \leq k \leq K. \quad (3.7)$$

La Figure 3.1 illustre les coûts espérés associés à la classification d'une tâche par l'humain et par l'automation.

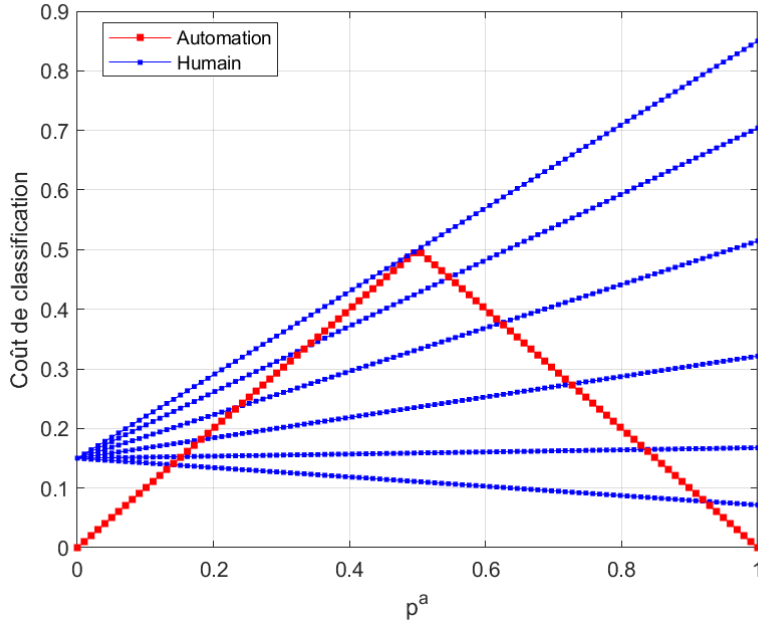


FIGURE 3.1 Coûts de classification avec $c_{fp} = c_{fn} = 1$ et $c_{tn} = c_{tp} = 0$. La courbe en rouge correspond au coût de classification de l'automation $C_a(p_k^a)$ en fonction du postérieur p_k^a . La courbe en bleu montre le coût de classification espéré pour l'humain $\Gamma_h(p_k^a, w^h)$ pour différentes charges de travail w^h . Le gain G correspond à la différence entre la fonction en rouge et celle en bleue.

En supposant que l'action de l'automatisme $a = \{\pi, w^r\}$ est spécifiée, on a que le coût espéré de classification de toutes les tâches pour l'équipe est donné par

$$\begin{aligned} C(\tau, p^a; \pi, w^r) &= \mathbb{E}_{w^s} \left[\sum_{i=1}^{w^r+w^s} \Gamma_h(p_{\pi_i}^a, w^r + w^s) + \sum_{i=w^r+w^s+1}^K C_a(p_{\pi_i}^a) \middle| \tau, w^r, p^a \right] \\ &= \sum_{k=1}^K C_a(p_k^a) - \mathbb{E}_{w^s} \left[\sum_{i=1}^{w^r+w^s} G(p_{\pi_i}^a, w^r + w^s) \middle| \tau, w^r, p^a \right], \end{aligned} \quad (3.8)$$

en définissant $\sum_{i=1}^0 \dots = 0$, et où $p^a = [p_1^a \dots p_K^a]^\top$.

3.2.4 Définition de la fonction de transition

La fonction de transition du MDP décrit la probabilité de passer d'un état τ_t à un état τ_{t+1} en posant l'action a_t . La dynamique du modèle à deux états est décrite par (3.2) où la probabilité de transition dépend du coût réalisé, qui lui même dépend de l'action choisie par le SSD. Le modèle utilisé pour cette fonction sera discuté au chapitre 4 puisqu'une hypothèse doit être posée afin d'assurer la validité de la solution optimale.

3.2.5 Description du scénario en MDP

Nous supposons l'état de confiance humaine comme étant parfaitement observable dans le cas de la formulation du problème en MDP. Le scénario est le même que celui décrit à la section 3.1. À la fin de chaque période t , un coût de classification c_t est réalisé par l'équipe, influençant la probabilité pour l'humain de se trouver dans un état de confiance normale à la prochaine période, tel que décrit par (3.2).

Le problème à résoudre est donc de trouver la politique d'action optimale pour le SSD de sorte à minimiser l'espérance des coûts réalisés par l'équipe. Soit $J_t^*(\tau)$ le coût optimal à la période t pour un état de confiance τ . Nous utilisons la notation $c(\tau, p^a, a) = \mathbb{E}_{w^s}[C(\tau, p^a; \pi, w^r)]$ pour le coût espéré défini en (3.8) lorsque l'opérateur a un état de confiance τ et que l'action du SSD est a . Nous introduisons également la notation

$$H_t^*(\tau) = \mathbb{E}_{Y^a}[J_t^*(\tau, p^a(Y^a))], \quad (3.9)$$

pour prendre l'espérance du coût optimal sur les observations Y^a . Rappelons que la distribution de Y^a ne dépend pas de t ce qui permet d'omettre l'indice dans la notation. L'équation

de Bellman pour ce problème est donnée par

$$\begin{aligned}
J_t^*(\tau, p^a) &= \min_{a=(\pi, w^r)} \{c(\tau, p^a, a_t = a) + \mathbb{P}[\tau_{t+1} = \mathbf{o} | \tau_t = \tau, a_t = a] H_{t+1}^*(\mathbf{o}) \\
&\quad + \mathbb{P}[\tau_{t+1} = - | \tau_t = \tau, a_t = a] H_{t+1}^*(-)\} \\
&= \min_a \{c(\tau, p^a, a_t = a) + \mathbb{P}[\tau_{t+1} = \mathbf{o} | \tau_t = \tau, a_t = a] (H_{t+1}^*(\mathbf{o}) - H_{t+1}^*(-))\} \\
&\quad + H_{t+1}^*(-).
\end{aligned}$$

Notons qu'il a été possible de simplifier l'équation de Bellman étant donné que la confiance prend seulement deux états de sorte que

$$\mathbb{P}[\tau_{t+1} = - | \tau_t = \tau, a_t = a] = 1 - \mathbb{P}[\tau_{t+1} = \mathbf{o} | \tau_t = \tau, a_t = a].$$

3.3 Formulation du problème en POMDP

Formulons à présent le problème sous la forme d'un POMDP dans le cas où la confiance n'est pas parfaitement observable. La résolution de ce problème sous cette hypothèse est donnée au chapitre 5. Un POMDP est défini par un tuple $\{S, A, P, R, \Omega, O\}$ où les éléments $\{S, A, P, R\}$ sont les mêmes que dans le cas du MDP.

Définissons Ω comme étant l'ensemble des éléments observables. Dans ce scénario, les tâches supplémentaires w^s prises par l'opérateur constituent une observation indirecte de sa confiance à partir de laquelle le SSD peut former une croyance quant à l'état de confiance réel de l'humain. Ensuite, nous avons la fonction d'observation O qui associe la probabilité d'observer un w^s étant donné l'état τ . Les définitions explicites de la fonction d'observation et de l'état de croyance seront données au chapitre 5.

3.4 Modèles numériques

Nous présentons à présent des exemples de modèles numériques pour chacune des fonctions et variables décrites dans ce chapitre. Ces exemples de modèles sont ceux utilisés pour les différentes simulations numériques. Il convient de rappeler qu'en pratique, de tels modèles peuvent être obtenus de façon expérimentale et que la validité d'un modèle dépend fortement du scénario dans lequel il est utilisé.

3.4.1 Modèle d'observation de l'automaton

La probabilité postérieure $p_{1:K}^a$ qu'une tâche soit un événement intéressant, que forme l'automaton à partir de ses observations $Y_{1:K}^a$, est définie en (3.4). Nous supposons que $\mathcal{H}_0 = 0$ et $\mathcal{H}_1 = d_0 \in \mathbb{R}$ et définissons l'observation de l'automaton de la tâche k

$$Y_k^a = \mathcal{H}_k + N_k^a, \quad k \in \mathcal{K},$$

où N_k^a est un processus gaussien indépendant de $\mathcal{H}_k$, avec $N_k^a \sim \mathcal{N}(0, \sigma_a^2)$, et où σ_a est la variance associée au processus gaussien de l'automaton. Il s'agit du même modèle donné en exemple dans [3], qu'on nomme modèle Gaussien à Variances Égales, et qui est central dans la théorie de détection de signal [69, Chapitre 2].

La probabilité postérieure $p_{1:K}^a$ s'écrit

$$\begin{aligned} p_{1:K}^a &= \mathbb{P}[H_k = \mathcal{H}_1 | Y_k^a] \\ &= \frac{\mathbb{P}[Y_k^a | \mathcal{H}_1] \mathbb{P}[\mathcal{H}_1]}{\mathbb{P}[Y_k^a | \mathcal{H}_1] \mathbb{P}[\mathcal{H}_1] + \mathbb{P}[Y_k^a | \mathcal{H}_0] \mathbb{P}[\mathcal{H}_0]} \end{aligned} \quad (3.10)$$

où $\mathbb{P}[\mathcal{H}_1] = p$, $\mathbb{P}[\mathcal{H}_0] = 1 - p$, et où $\mathbb{P}[Y_k^a | \mathcal{H}_1]$ est distribué selon une loi normale $\mathcal{N}(d_0, \sigma_a^2)$ et $\mathbb{P}[Y_k^a | \mathcal{H}_0]$ selon une loi normale $\mathcal{N}(0, \sigma_a^2)$.

3.4.2 Modèle de tâches supplémentaires w_t^s

Nous présentons maintenant un exemple de modèle pour les tâches supplémentaires w_t^s prises par l'humain, utilisé dans les différentes simulations numériques.

Nous utilisons la loi de Skellam pour représenter ce modèle. La loi de Skellam est une loi de probabilité discrète obtenue par la différence de deux variables aléatoires qui suivent des lois de Poisson de paramètres λ_1 et λ_2 [70]. La fonction de masse de la loi de Skellam est donnée par

$$f(k; \lambda_1, \lambda_2) = e^{-(\lambda_1 + \lambda_2)} \left(\frac{\lambda_1}{\lambda_2} \right)^{k/2} I_{|k|} \left(2\sqrt{\lambda_1 \lambda_2} \right)$$

où $I_{|k|}$ est la fonction de Bessel modifiée de première espèce.

L'espérance d'une variable aléatoire $X \sim \text{Skellam}(\lambda_1, \lambda_2)$ est simplement $\mathbb{E}[X] = \lambda_1 - \lambda_2$, alors que sa variance est $\text{Var}[X] = \lambda_1 + \lambda_2$.

Le choix d'un tel modèle est tout d'abord motivé par l'allure normale (au sens probabiliste) de sa fonction de masse. En effet, puisque la réponse humaine w_t^s est discrète, on ne peut utiliser directement une loi normale pour la modéliser. De plus, cette loi permet de travailler

facilement avec une distribution qui prend seulement deux paramètres et pour lesquels on connaît directement l'influence sur l'espérance et la variance. Il devient alors facile de choisir $\lambda_1(\tau)$ et $\lambda_2(\tau)$ pour $\tau \in \{-, \mathbf{o}\}$ de sorte à respecter l'Hypothèse 2 et à obtenir des résultats intéressants.

Notons finalement que la distribution de Skellam a un support infini alors que w^s prend des réalisations finies. Pour contrevenir à ce problème, nous sommes les probabilités à l'extérieur du support de w^s aux valeurs limites qui peuvent être prises ($w^s = -w^r$ et $w^s = K - w^r$).

La Figure 3.2 présente la distribution des probabilités de w^s pour différents paramètres λ . Pour chacune des figures tracées, nous posons $w^r = 10$ et $K = 20$ afin d'avoir un support symétrique $w^s \in \{-w^r, \dots, K - w^r\}$. Nous associons l'état de confiance normal $\tau = \mathbf{o}$ à des paramètres $\lambda_1 = \lambda_2 = 1$ fixes de sorte à avoir une moyenne nulle pour cet état et on détériore pour chaque sous-figure la réponse de l'humain en cas de sous-confiance $\tau = -$ en augmentant les valeurs des paramètres λ de sorte à augmenter la variance.

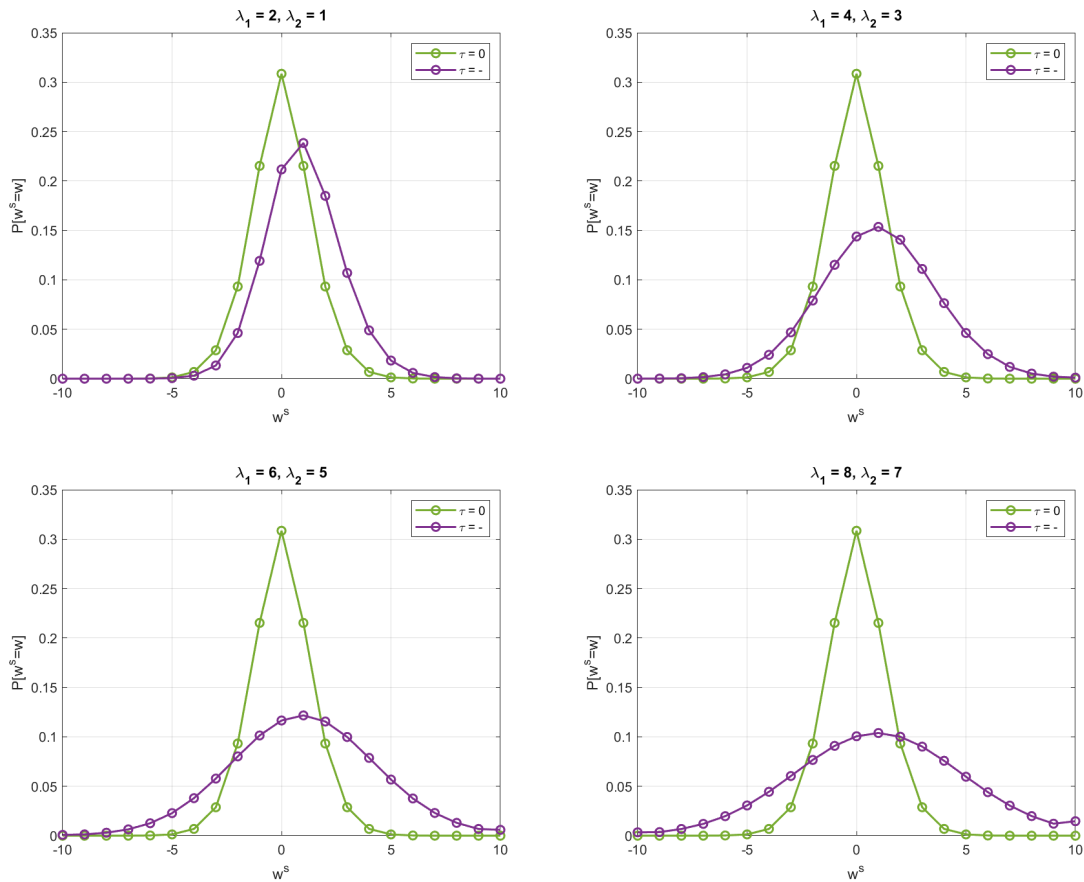


FIGURE 3.2 Distribution des probabilités de w^s avec une moyenne fixe égale à 1.

3.4.3 Modèle de performance humaine

Nous utilisons le modèle présenté dans [55] pour modéliser les courbes de performances humaines en fonction du workload $P_{tp}(w)$ et $P_{fp}(w)$. Ce modèle se base sur les caractéristiques de performance (*Receiver Operating Characteristic* (ROC)) qui sont représentées par un ensemble de courbes qui dépendent de la charge de travail de l'humain [71]. De manière similaire à ce qui est fait pour l'observation de l'automatisme, nous supposons que l'observation de l'opérateur Y_k^h lorsque $\mathcal{H}_k = \mathcal{H}_1$ suit une loi normale $\mathcal{N}(d(w^h/K), \sigma_h^2)$, alors qu'elle suit une loi normale $\mathcal{N}(0, \sigma_h^2)$ pour une tâche $\mathcal{H}_k = \mathcal{H}_0$. Nous introduisons $d(w^h/K)$ qui est un paramètre défini comme étant la sensibilité et qui dépend de la charge de travail $w^h/K \in [0, 1]$ de l'opérateur, et σ_h la variance associée au processus gaussien de l'humain. Nous définissons la sensibilité comme étant

$$d(w^h/K) = d_h(1 - w^h/K)$$

où d_h est un paramètre à identifier par expérience. Ainsi, plus la charge de travail w^h augmente, plus la sensibilité de l'opérateur, soit sa capacité à distinguer $\mathcal{H}_1$ de $\mathcal{H}_0$, diminue.

Soit λ_h le seuil pour lequel l'opérateur prend la décision $D_k = \mathcal{H}_1$ si Y_k^h est à droite du seuil et $D_k = \mathcal{H}_0$ autrement. Ce seuil est défini par l'opérateur lui-même basé sur ses expériences et ses connaissances et est supposé comme étant fixe pour une période donnée. Nous avons

$$\begin{aligned} P_{tp}(w) &= \mathbb{P}[Y_k^h \geq \lambda_h \mid \mathcal{H}_k = \mathcal{H}_1] \\ &= 1 - \mathbb{P}[X + d(w) < \lambda_h] \quad \text{avec } X \sim \mathcal{N}(0, 1) \\ &= 1 - \Phi(\lambda_h - d(w)) \\ &= \Phi(d(w) - \lambda_h) \end{aligned}$$

où $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$ est la fonction de répartition d'une loi normale standard $\mathcal{N}(0, 1)$. Nous avons également que la probabilité de réaliser un faux positif est donnée par

$$P_{fp} = \mathbb{P}[Y_k^h \geq \lambda_h \mid \mathcal{H}_0] = \Phi(-\lambda_h).$$

Puisque λ_h est supposé comme étant fixe, nous avons donc que la probabilité de faux positif P_{fp} se trouve à être également fixée et indépendante de la charge de travail w^h . La courbe ROC résultant de l'utilisation du modèle de GVE est donc donnée par

$$P_{tp}(w) = \Phi(d(w) + \Phi^{-1}(P_{fp})).$$

Le modèle utilisé pour les performances humaines respecte l'Hypothèse 3 énoncée précédem-

ment utilisée dans la preuve d'optimalité de la politique au chapitre suivant. La Figure 3.3 montre l'allure des courbes de performances de l'humain pour cet exemple de modèle.

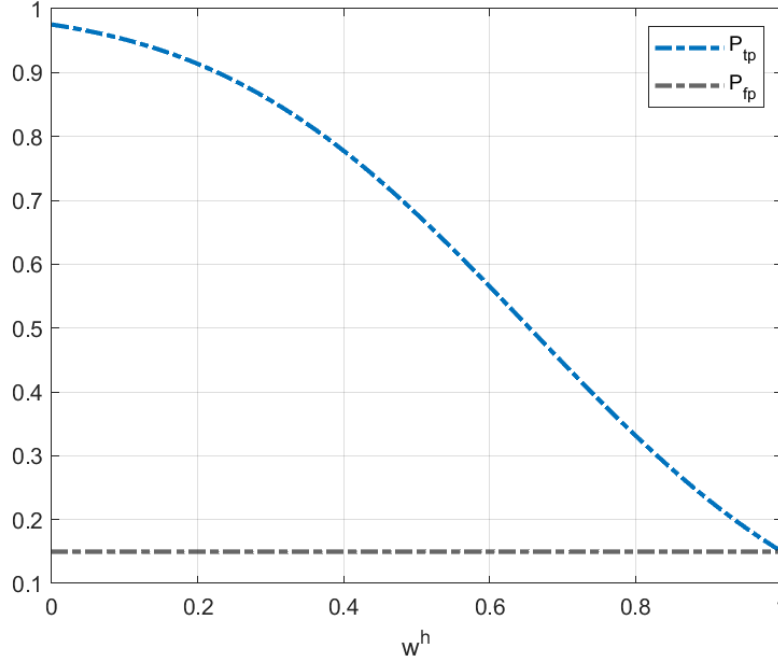


FIGURE 3.3 Courbes de performance de classification de l'humain en fonction de sa charge de travail w^h .

Dans le scénario donné, nous supposons un taux de faux positif P_{fp} fixe qui est implémenté par l'opérateur et qui est connu de l'automation. Lorsque plusieurs tâches doivent être effectuées, on peut respecter cette exigence en suivant des règles de décision basées sur des seuils, qui sont implémentés de manière naturelle par l'humain [55].

CHAPITRE 4 OBSERVATION PARFAITE DE LA CONFIANCE

Dans ce chapitre, nous considérons l'état de confiance de l'opérateur comme étant parfaitement observable à chaque début de période. Conséquemment, le SSD connaît de manière exacte la distribution probabiliste des tâches supplémentaires w_t^s qui sont prises par l'opérateur. Dans un tel contexte, nous démontrons l'optimalité d'une politique myope et fournissons un algorithme permettant de la calculer. Des simulations numériques comparant cette politique à d'autres politiques sous-optimales sont également présentées.

4.1 Calcul de la politique myope

Définissons $\mathbb{P}_u$ l'opérateur de probabilité conditionnel à l'utilisation d'une politique markovienne u , de sorte que

$$\mathbb{P}_u[\tau_{t+1} = \mathbf{o} | \tau_t, p_t^a] := \mathbb{P}[\tau_{t+1} = \mathbf{o} | \tau_t, a_t = u_t(\tau_t, p_t^a)],$$

et notons $\mathbb{E}_u$ l'opérateur de l'espérance correspondant à $\mathbb{P}_u$.

Définition 1. Une politique myope $\mu(\cdot)$ pour le SSD est une politique stationnaire qui choisit à chaque période l'action qui minimise le coût espéré pour cette période, i.e.,

$$\forall p^a \in [0, 1]^K, \forall \tau \in \{\mathbf{o}, -\}, \forall a \in \mathcal{A}, \quad c(\tau, p^a, \mu(\tau, p^a)) \leq c(\tau, p^a, a).$$

Notons que dans la Définition 1, plusieurs actions peuvent exister qui minimisent le coût, auquel cas on choisit de manière arbitraire l'une de ces actions. On va tout de même se référer à cette action comme étant «la» politique myope malgré le fait qu'elle ne soit pas forcément unique.

Dans le reste de cette section, étant donné qu'une seule période de temps t est considérée, nous omettons l'indice t de la notation à des fins de simplicité. Rappelons que l'opérateur se voit fournir une liste ordonnée π des tâches à effectuer et qu'il effectue sa classification dans l'ordre spécifié de cette liste. Il reçoit également une recommandation de charge de travail w^r qu'il peut ne pas suivre de manière exacte selon son état de confiance. Ainsi, la charge de travail recommandée optimale et la permutation de tâches optimale sont la solution du problème d'optimisation suivant

$$\max_{w^r, \pi} \mathbb{E}_{w^s} \left[\sum_{i=1}^{w^r + w^s} G(p_{\pi_i}, w^r + w^s) \middle| \tau, w^r, p^a \right], \quad (4.1)$$

avec la fonction G définie en (3.7).

Démonstration. Une fois π et w^r spécifiés pour la période, le coût espéré pour toutes les décisions est donné par (3.8). Puisque le premier terme dans (3.8) est constant, la minimisation du coût espéré est équivalente à la maximisation du second terme. ■

Ce problème d'optimisation peut être résolu en deux étapes. Premièrement, nous calculons la permutation de tâches optimale pour un choix donné de w^r . Deuxièmement, nous pouvons effectuer une recherche exhaustive sur $w^r \in \{0, 1, \dots, K\}$. Pour la première étape, étant donné w^r , notons que la fonction objectif à maximiser (4.1) s'écrit explicitement

$$\begin{aligned} \mathbb{E}_{w^s} \left[\sum_{i=1}^{w^r+w^s} G(p_{\pi_i}, w^r + w^s) \middle| \tau, w^r, p^a \right] &= \sum_{w=1-w^r}^{K-w^r} \mathbb{P}(w^s = w | \tau, w^r) \sum_{i=1}^{w^r+w} G(p_{\pi_i}^a, w^r + w) \\ &= \sum_{\tilde{w}=1}^K \mathbb{P}(w^s = \tilde{w} - w^r | \tau, w^r) \sum_{i=1}^{\tilde{w}} G(p_{\pi_i}^a, \tilde{w}) \end{aligned} \quad (4.2)$$

après le changement de variable $\tilde{w} = w^r + w^s$ et en utilisant le fait que $\sum_{i=1}^0 \dots = 0$, ou encore plus explicitement

$$\begin{aligned} &\mathbb{P}(w^s = 1 - w^r | \tau, w^r) G(p_{\pi_1}^a, 1) + \mathbb{P}(w^s = 2 - w^r | \tau, w^r) (G(p_{\pi_1}^a, 2) + G(p_{\pi_2}^a, 2)) + \dots \\ &+ \mathbb{P}(w^s = K - w^r | \tau, w^r) (G(p_{\pi_1}^a, K) + \dots + G(p_{\pi_K}^a, K)). \end{aligned}$$

En groupant les termes pour chaque indice spécifique π_i ensemble (i.e. en changeant l'ordre des sommes de (4.2)), il est finalement possible de réécrire l'objectif comme étant

$$\sum_{i=1}^K A_{i, \pi_i}(w^r; p^a, \tau), \quad (4.3)$$

où, pour tout $1 \leq i, j \leq K$, nous définissons

$$A_{i,j}(w^r; \tau, p^a) := \sum_{w=i}^K \mathbb{P}(w^s = w - w^r | \tau, w^r) G(p_j^a, w).$$

Pour une valeur donnée de w^r (et de l'état de confiance τ et des probabilités postérieures p^a connues), nous calculons les K^2 éléments $A_{i,j}(w^r; \tau, p^a)$, et alors le calcul de la permutation π maximisant (4.3) est un problème d'affectation linéaire, qui peut être résolu via la programmation linéaire ou via l'algorithme hongrois [72] ou l'algorithme Jonker-Volgenant [73] en temps $O(K^3)$. Ainsi, avec la recherche exhaustive de la valeur optimale de w^r , le problème d'optimisation (4.1) peut être résolu en temps $O(K^4)$.

4.2 Optimalité de la politique myope

Dans cette section, nous identifions les conditions pour lesquelles la politique myope est optimale. Intuitivement, puisque la politique myope se concentre uniquement sur la période actuelle, tant que l'on considère des scénarios pour lesquels il est généralement préférable d'être dans l'état de confiance normal et où la minimisation des coûts contribue à maintenir l'opérateur dans cet état, la politique myope devrait être optimale. Cette affirmation est formulée de manière plus formelle dans la suite.

Propriétés de la politique myope

Rappelons les hypothèses 1 et 3 énoncées au chapitre 3, que nous supposons vérifiées pour tous les résultats de cette section.

Lemme 1. *Pour tout $p \in [0, 1]$ et tout $1 \leq w \leq K - 1$*

$$G(p, w + 1) \leq G(p, w).$$

Démonstration. Sous les hypothèses 1 et 3, nous montrons immédiatement que le coût espéré de renvoi de tâche à l'humain $\Gamma_h(p, w)$ défini par (3.6) est non-décroissant avec w , donc $G(p, \cdot)$ est non-croissant pour tout p . ■

Lemme 2. *Fixons $w^r \in \{0, \dots, K\}$. Il existe toujours une permutation π qui maximise (4.1) et qui de plus satisfait*

$$G(p_{\pi_{w+1}}, w + 1) \leq G(p_{\pi_w}, w). \quad (4.4)$$

pour tout $1 \leq w \leq K - 1$.

Démonstration. Soit la notation $\tilde{w} = w^r + w^s$. Alors pour une permutation π , on peut réécrire l'objectif en (4.1) comme étant

$$\begin{aligned} R(\pi) = & \mathbb{P}(\tilde{w} = 1 \mid \tau, w^r) G(p_{\pi_1}^a, 1) + \mathbb{P}(\tilde{w} = 2 \mid \tau, w^r) (G(p_{\pi_1}^a, 2) + G(p_{\pi_2}^a, 2)) + \dots \\ & + \mathbb{P}(\tilde{w} = K \mid \tau, w^r) (G(p_{\pi_1}^a, K) + \dots + G(p_{\pi_K}^a, K)). \end{aligned}$$

Supposons que

$$G(p_{\pi_{w_0}}, w_0) < G(p_{\pi_{w_0+1}}, w_0 + 1) \quad (4.5)$$

pour un certain $1 \leq w_0 \leq K - 1$. Définissons une nouvelle permutation $\tilde{\pi}$ en échangeant π_{w_0} et π_{w_0+1} , en d'autres mots $\tilde{\pi}_i = \pi_i$ pour $i \notin \{w_0, w_0 + 1\}$, et $\tilde{\pi}_{w_0} = \pi_{w_0+1}$ et $\tilde{\pi}_{w_0+1} = \pi_{w_0}$. On calcule que la différence de coût est

$$R(\tilde{\pi}) - R(\pi) = \mathbb{P}(\tilde{w} = w_0)(G(p_{\pi_{w_0+1}}, w_0) - G(p_{\pi_{w_0}}, w_0)).$$

Selon le Lemme 1 et l'hypothèse de (4.5), nous avons que

$$G(p_{\pi_{w_0+1}}, w_0) \geq G(p_{\pi_{w_0+1}}, w_0 + 1) > G(p_{\pi_{w_0}}, w_0),$$

donc $R(\tilde{\pi}) \geq R(\pi)$ (strictement si $\mathbb{P}(\tilde{w} = w_0) \neq 0$) et on peut changer π à $\tilde{\pi}$ sans diminuer l'objectif. En d'autres mots, en partant d'une permutation qui ne satisfait pas la propriété (4.4), il est toujours possible de changer cette permutation de sorte à ne pas diminuer l'objectif et aussi satisfaire la condition. Cela prouve le lemme. ■

Rappelons la définition suivante.

Définition 2. Une fonction $f : \mathbb{N} \rightarrow \mathbb{R}$ est dite concave discrète si

$$2f(w) \geq f(w - 1) + f(w + 1),$$

pour tout w dans le support de f . De plus, f est dite convexe discrète si $-f$ est concave discrète.

Une fonction concave discrète peut être étendue à une fonction concave sur $\mathbb{R}$, par exemple, via l'interpolation linéaire [74].

Posons maintenant une hypothèse sur la convexité discrète du modèle de performance humaine.

Hypothèse 5. Nous supposons que $w \mapsto P_{tp}(w)$ est concave discrète et $w \mapsto P_{fp}(w)$ est convexe discrète.

Cette hypothèse est nécessaire pour compléter le reste de la preuve et est satisfaite par l'exemple de modèle de la performance humaine donné en 3.4.3.

Le résultat suivant est immédiat.

Lemme 3. Sous les hypothèses 1 et 5, les fonctions $w \mapsto G(p, w)$ sont concaves discrètes,

pour tout $p \in [0, 1]$.

Démonstration. On montre que sous les hypothèses, $\Gamma_h(p, \cdot)$ est convexe discrète, d'où il s'ensuit que $G(p, \cdot)$ est concave discrète. En effet, réécrivons (3.6) sous la forme suivante en isolant pour P_{tp} et P_{fp} ,

$$\begin{aligned}\Gamma_h(p_k^a, w) &= p_k^a[P_{tp}(w)c_{tp} + (1 - P_{tp}(w))c_{fn}] + (1 - p_k^a)[P_{fp}(w)c_{fp} + (1 - P_{fp}(w))c_{tn}] \\ &= P_{tp}(w)p_k^a(c_{tp} - c_{fn}) + P_{fp}(w)(1 - p_k^a)(c_{fp} - c_{tn}) + p_k^a c_{fn} + (1 - p_k^a)c_{tn}\end{aligned}$$

L'hypothèse 1 donne $p_k^a(c_{tp} - c_{fn}) \leq 0$ et $(1 - p_k^a)(c_{fp} - c_{tn}) \geq 0$ alors que l'hypothèse 5 assume $P_{tp}(w)$ concave discrète et $P_{fp}(w)$ convexe discrète. La somme de deux fonctions convexes est convexe et puisque $C_a(p_k^a)$ est indépendant de w , alors $G(p, \cdot)$ est concave pour tout $p \in [0, 1]$. ■

Lemme 4. *Supposons que les hypothèses 1, 3 et 5 soient validées. Soit w^r une charge de travail suggérée et soit π une permutation maximisant (4.1) pour cette valeur de w^r et satisfaisant la condition (4.4) du Lemme 2. Alors la fonction*

$$w \mapsto \bar{G}_{w^r}^\pi(w, p^a) = \sum_{i=1}^{w^r+w} G(p_{\pi_i}, w^r + w)$$

apparaissant dans (4.1) est concave discrète pour tout postérieur $p^a \in [0, 1]^K$, i.e.,

$$2\bar{G}_{w^r}^\pi(w, p^a) \geq \bar{G}_{w^r}^\pi(w - 1, p^a) + \bar{G}_{w^r}^\pi(w + 1, p^a), \quad (4.6)$$

pour tout w tel que $-w^r + 1 \leq w \leq K - w^r - 1$.

Démonstration. Soit $\tilde{w} = w^r + w$. Nous pouvons écrire

$$\begin{aligned}2\bar{G}_{w^r}^\pi(w, p^a) &= 2G(p_{\pi_{\tilde{w}}}, \tilde{w}) + \sum_{i=1}^{\tilde{w}-1} 2G(p_{\pi_i}, \tilde{w}) \\ &\geq 2G(p_{\pi_{\tilde{w}}}, \tilde{w}) + \sum_{i=1}^{\tilde{w}-1} G(p_{\pi_i}, \tilde{w} - 1) + G(p_{\pi_i}, \tilde{w} + 1) \\ &= \bar{G}_{w^r}^\pi(w - 1, p^a) + \bar{G}_{w^r}^\pi(w + 1, p^a) + 2G(p_{\pi_{\tilde{w}}}, \tilde{w}) \\ &\quad - G(p_{\pi_{\tilde{w}}}, \tilde{w} + 1) - G(p_{\pi_{\tilde{w}+1}}, \tilde{w} + 1),\end{aligned}$$

où l'inégalité découle du Lemma 3. De plus,

$$G(p_{\pi_{\tilde{w}}}, \tilde{w}) \geq G(p_{\pi_{\tilde{w}}}, \tilde{w} + 1)$$

à partir du Lemme 1 et

$$G(p_{\pi_{\tilde{w}}}, \tilde{w}) \geq G(p_{\pi_{\tilde{w}+1}}, \tilde{w} + 1)$$

à partir de la condition (4.4) du Lemme 2 satisfaite par la permutation optimale π . Ceci montre (4.6). ■

Proposition 1. *Sous les hypothèses 1, 2, 3 et 5, pour tout postérieur $p^a \in [0, 1]^K$, le coût espéré de la politique myope à chaque période est plus faible dans l'état de confiance normale que dans l'état de sous-confiance, i.e.,*

$$c(\mathbf{o}, p^a, \mu(\mathbf{o})) \leq c(-, p^a, \mu(-)). \quad (4.7)$$

Démonstration. Considérons la décomposition (3.3), et dénotons $\nu^s(\tau, w^r) := \mathbb{E}[w^s | \tau, w^r]$. Un simple changement de variable dans (4.1) permet de réécrire le problème sous la forme

$$\max_{\bar{w}^r, \pi} \mathbb{E}_{\tilde{w}^s} \left[\bar{G}_{\bar{w}^r}^{\pi}(\tilde{w}^s, p^a) \middle| \tau, \bar{w}^r, p^a \right], \quad (4.8)$$

où $\tilde{w}^s$ est la variable aléatoire de moyenne nulle définie dans (3.3). En particulier, si w^r et π maximisent (4.1) alors $\bar{w}^r = w^r + \nu^s(\tau, w^r)$ et le même π maximise (4.8), et les valeurs optimales sont les mêmes. En d'autres mots, la valeur optimale du problème d'optimisation ne dépend pas de la moyenne de w^s , puisque cette moyenne est connue et peut être compensée par un simple décalage dans le choix w^r .

Soit $\bar{w}_-^{r,*}$, π_-^* le couple maximisant (4.8) pour $\tau = -$, tel que π_-^* satisfait la condition du Lemme 2. Alors $\bar{G}_{\bar{w}_-^{r,*}}^{\pi_-^*}$ est concave discrète par le Lemme 4, et d'après l'hypothèse 2, on a

$$\begin{aligned} \mathbb{E}_{\tilde{w}^s} \left[\bar{G}_{\bar{w}_-^{r,*}}^{\pi_-^*}(\tilde{w}^s, p^a) \middle| \tau = -, \bar{w}_-^{r,*}, p^a \right] &\leq \mathbb{E}_{\tilde{w}^s} \left[\bar{G}_{\bar{w}_-^{r,*}}^{\pi_-^*}(\tilde{w}^s, p^a) \middle| \tau = \mathbf{o}, \bar{w}_-^{r,*}, p^a \right] \\ &\leq \max_{\bar{w}^r, \pi} \mathbb{E}_{\tilde{w}^s} \left[\bar{G}_{\bar{w}^r}^{\pi}(\tilde{w}^s, p^a) \middle| \tau = \mathbf{o}, \bar{w}^r, p^a \right]. \end{aligned}$$

Puisque la première expression est égale à $\sum_{k=1}^K C_a(p_k^a) - c(-, p^a, \mu(-))$ et la dernière expression est égale à $\sum_{k=1}^K C_a(p_k^a) - c(\mathbf{o}, p^a, \mu(\mathbf{o}))$, cela démontre (4.7). ■

Nous posons l'hypothèse suivante pour le reste de la section.

Hypothèse 6. À chaque période $0 \leq t \leq T - 1$, la probabilité de passer à un état de confiance normal diminue linéairement à mesure que le coût augmente, i.e., pour chaque $\tau \in \{\mathbf{o}, -\}$, il existe une constante positive $\kappa_\tau, \lambda_\tau$ tel que

$$f_\tau(\tau, c) = \kappa_\tau - \lambda_\tau c, \quad \forall c, \quad (4.9)$$

où la fonction f_τ est définie dans (3.2).

Cette hypothèse est nécessaire afin de permettre la comparaison directe de la distribution des probabilités de transition à un état de confiance normal entre la politique optimale et toute autre action de l'automatisme. Notons toutefois qu'il s'agit ici de l'hypothèse qui affaiblit le plus la pertinence de ce travail. En pratique, il faudrait calibrer ce modèle par des expériences avec l'opérateur et il est peu probable que l'évolution de la confiance suive un modèle linéaire. Notons toutefois que certains travaux ont utilisé de tels modèles [24, 46].

Nous dérivons maintenant quelques propriétés utiles pour la politique myope.

Proposition 2. Sous les hypothèses 4 et 6, à chaque période $0 \leq t \leq T - 1$, l'action d'une politique myope garantit une probabilité plus élevée d'être dans un état de confiance normale à la période $t + 1$, par rapport à toute autre action, i.e.,

$$\mathbb{P}_\mu[\tau_{t+1} = \mathbf{o} | \tau_t, p^a] \geq \mathbb{P}[\tau_{t+1} = \mathbf{o} | \tau_t, p^a, a_t = a], \quad \forall a. \quad (4.10)$$

Démonstration. À chaque période t , selon la loi de probabilité totale et en utilisant l'hypothèse 4,

$$\begin{aligned} \mathbb{P}_\mu[\tau_{t+1} = \mathbf{o} | \tau_t, p^a] &= \mathbb{E}[\mathbb{P}[\tau_{t+1} = \mathbf{o} | \tau_t, p^a, \mu(\tau, p^a), c_t] | \tau_t], \\ &= \mathbb{E}_\mu[\mathbb{P}[\tau_{t+1} = \mathbf{o} | \tau_t, c_t] | \tau_t], \\ &= \mathbb{E}_\mu[f_\tau(\tau_t, c_t) | \tau_t]. \end{aligned}$$

Ainsi, en utilisant l'hypothèse 6, nous obtenons pour toute action a :

$$\begin{aligned} \mathbb{P}_\mu[\tau_{t+1} = \mathbf{o} | \tau_t, p^a] &= \kappa_{\tau_t} - \lambda_{\tau_t} \mathbb{E}_\mu[C_t | \tau_t, p^a] \\ &\geq \kappa_{\tau_t} - \lambda_{\tau_t} c(\tau_t, p^a, a), \\ &= \mathbb{E}[f_\tau(\tau_t, c_t) | \tau_t, p^a, a_t = a] \\ &= \mathbb{P}[\tau_{t+1} = \mathbf{o} | \tau_t, p^a, a_t = a], \end{aligned}$$

où l'inégalité découle de la définition de la politique myope. Cela démontre l'équation (4.10). ■

Hypothèse 7. *Nous supposons que pour tout coût c ,*

$$f_\tau(\mathbf{o}, c) \geq f_\tau(-, c),$$

c'est-à-dire qu'il y a une plus grande chance d'être dans l'état de confiance normale à l'instant $t + 1$ si l'on est déjà dans cet état à l'instant t .

Des travaux soutiennent cette hypothèse en démontrant que l'état actuel de confiance de l'opérateur joue un rôle important dans l'évolution de sa confiance [28, 31, 35].

Proposition 3. *Sous les hypothèses 6 et 7, lors de l'application de la politique myope, la probabilité d'être dans un état de confiance normale à la période suivante est plus grande si l'on est déjà dans cet état, c'est-à-dire :*

$$\forall p^a \in [0, 1]^K, \quad \mathbb{P}_\mu[\tau_{t+1} = \mathbf{o} | \tau_t = \mathbf{o}, p^a] \geq \mathbb{P}_\mu[\tau_{t+1} = \mathbf{o} | \tau_t = -, p^a]. \quad (4.11)$$

Démonstration. Par la loi de probabilité totale :

$$\begin{aligned} \mathbb{P}_\mu[\tau_{t+1} = \mathbf{o} | \tau_t = \mathbf{o}, p^a] &= \mathbb{E}_\mu[\mathbb{P}[\tau_{t+1} = \mathbf{o} | \tau_t = \mathbf{o}, p^a, c_t, \mu(\mathbf{o}, p^a)] | \tau_t = \mathbf{o}, p^a] \\ &= \mathbb{E}_\mu[f_\tau(\mathbf{o}, c_t) | \tau_t = \mathbf{o}, p^a] \\ &\geq \mathbb{E}_\mu[f_\tau(-, c_t) | \tau_t = \mathbf{o}, p^a] \end{aligned} \quad (4.12)$$

$$\begin{aligned} &= \kappa_- - \lambda_- \mathbb{E}_\mu[C_t | \tau_t = \mathbf{o}, p^a] \\ &= \kappa_- - \lambda_- c(\mathbf{o}, p^a, \mu(\mathbf{o}, p^a)) \\ &\geq \kappa_- - \lambda_- c(-, p^a, \mu(-, p^a)) \\ &= \mathbb{P}_\mu[\tau_{t+1} = \mathbf{o} | \tau_t = -, p^a], \end{aligned} \quad (4.13)$$

où l'inégalité (4.12) découle de l'hypothèse 7 et (4.13) découle de la Proposition 1. ■

Optimalité de la politique myope

On rappelle la notation $J_t^*(\tau, p^a)$ qui désigne le coût optimal à l'instant t lorsque l'état de confiance actuel est τ et les postérieurs sont p^a . Nous obtenons le résultat suivant.

Théorème 1. *Sous les hypothèses 1 à 7, la politique myope est optimale. De plus, pour tout $0 \leq t \leq T$, on a*

$$J_t^*(\mathbf{o}, p^a) \leq J_t^*(-, p^a), \quad \forall p^a \in [0, 1]^K. \quad (4.14)$$

Démonstration. À la période T , la politique myope est optimale par définition. L'inégalité (4.14) correspond alors exactement à (4.7), c'est-à-dire au résultat de la Proposition 1. Ensuite, nous procédons par induction. Considérons $0 \leq t \leq T-1$ et supposons que la politique myope est optimale à partir de l'instant $t+1$, et que (4.14) est satisfait à $t+1$. Rappelons la notation $H_t^*(\tau) = \mathbb{E}_{Y^a}[J_t^*(\tau, p^a(Y^a))]$ définie en (3.9). L'équation de Bellman s'écrit

$$\begin{aligned} J_t^*(\tau, p^a) = & \min_a \{ c(\tau, p^a, a) + \mathbb{P}[\tau_{t+1} = \mathbf{o} | \tau_t = \tau, p^a, a] H_{t+1}^*(\mathbf{o}) \\ & + \mathbb{P}[\tau_{t+1} = - | \tau_t = \tau, p^a, a] H_{t+1}^*(-) \}, \end{aligned}$$

qu'il est possible de réécrire comme étant

$$J_t^*(\tau, p^a) = H_{t+1}^*(-) + \min_a \{ c(\tau, p^a, a) + \mathbb{P}[\tau_{t+1} = \mathbf{o} | \tau_t = \tau, a] (H_{t+1}^*(\mathbf{o}) - H_{t+1}^*(-)) \}. \quad (4.15)$$

Dans la minimisation par rapport à a , l'action de la politique myope minimise le premier terme $c(\tau, p^a, a)$ par la Définition 1. Ensuite, puisque $(J_{t+1}^*(\mathbf{o}, p^a) - J_{t+1}^*(-, p^a)) \leq 0 \forall p^a$ par l'hypothèse de récurrence, alors $H_{t+1}^*(\mathbf{o}) - H_{t+1}^*(-) \leq 0$, et puisque l'action de la politique myope maximise la probabilité de transition $\mathbb{P}[\tau_{t+1} = \mathbf{o} | \tau_t = \tau, p^a, a]$ selon la Proposition 2, elle minimise également le dernier terme dans (4.15). Par conséquent, cette action myope est optimale à t .

Il reste à montrer que (4.14) est vérifiée à l'instant t . Soit $p^a \in [0, 1]^K$. Maintenant qu'on sait que la politique myope est optimale à partir de l'instant t , nous avons

$$\begin{aligned} J_t^*(\mathbf{o}, p^a) &= J_t^\mu(\mathbf{o}, p^a) \\ &= H_{t+1}^*(-) + c(\mathbf{o}, p^a, \mu(\mathbf{o})) + \mathbb{P}_\mu[\tau_{t+1} = \mathbf{o} | \tau_t = \mathbf{o}, p^a] (H_{t+1}^*(\mathbf{o}) - H_{t+1}^*(-)) \\ &\leq H_{t+1}^*(-) + c(-, p^a, \mu(-)) + \mathbb{P}_\mu[\tau_{t+1} = \mathbf{o} | \tau_t = -, p^a] (H_{t+1}^*(\mathbf{o}) - H_{t+1}^*(-)) \quad (4.16) \\ &= J_t^\mu(-, p^a) = J_t^*(-, p^a), \end{aligned}$$

où l'inégalité (4.16) découle de (4.7), du fait que $(J_{t+1}^*(\mathbf{o}) - J_{t+1}^*(-)) \leq 0$ par l'hypothèse d'induction, et de (4.11). Donc (4.14) est vrai à t et le théorème est démontré. $\blacksquare$

4.3 Politique myope : version alternative

Il a été discuté à la section 3.2.2 qu'un scénario alternatif aurait été de considérer le cas où la liste de tâches π_t envoyée à l'opérateur à l'instant t serait calculée seulement après avoir pris connaissance du nombre de tâches supplémentaires w_t^s choisi par l'opérateur. Ainsi, pour une charge de travail donnée, on obtient une permutation $\pi^w = (\pi_1^w, \dots, \pi_K^w)$ de $\mathcal{K}$ en ordonnant de manière décroissante les indices G définis en (3.7) de sorte à avoir

$$G(p_{\pi_1^w}^a, w) \geq \dots \geq G(p_{\pi_K^w}^a, w).$$

Il a été montré dans [3] qu'une fois la charge de travail connue (en considérant les tâches supplémentaires choisies par l'humain), le coût total pour la période en prenant les w premières tâches de la liste π^w spécifiée est donnée par $\sum_{k=1}^K C_a(p_k^a) - \bar{G}(w)$, où

$$\bar{G}(w, p^a) := \sum_{i=1}^w G(p_{\pi_i^w}^a, w). \quad (4.17)$$

Dans le scénario présenté dans le cadre ce mémoire, pour un choix de w^r et un état de confiance τ_t , le coût espéré devient $\sum_{k=1}^K C_a(p_k^a) - \tilde{G}(w^r, \tau_t, p^a)$, où

$$\tilde{G}(w^r, \tau_t, p^a) := \mathbb{E}[\bar{G}(w^r + w^s, p^a) | \tau_t, p^a].$$

Donc, après avoir observé l'état de confiance τ_t , l'automation qui suit une politique myope doit choisir un w^r qui maximise $\tilde{G}(w^r, \tau_t, p^a)$. Ceci peut être fait, par exemple, de manière naïve en effectuant une recherche exhaustive sur toutes les valeurs de $w^r \in \{0, \dots, K\}$. Une fois le choix de w^r fixé et la réponse w^s de l'humain connue, la liste π^w peut être calculée en ordonnant de manière décroissante les indices de G pour le $w^h = w^r + w^s$ de la période.

Tel qu'énoncé dans la formulation du problème, nous préférons toutefois considérer le scénario pour lequel l'automation doit fournir une liste de tâches π sans connaître de manière exacte w^s afin de mieux refléter la réalité. En effet, il est complexe pour un opérateur d'estimer le nombre exact de tâches supplémentaires qu'il compte effectuer durant la période. Ainsi, puisque l'ordre des indices de la permutation π fourni par l'automation change en fonction de la charge de travail de l'humain, il devient problématique d'évaluer l'action optimale à prendre pour l'automation si la réponse de l'humain peut changer au cours de la période. On voit donc que prouver l'optimalité de la politique myope développée n'est pas aussi évident a priori dans ce mode d'interaction.

4.4 Simulations numériques

Nous rappelons le scénario du problème étudié décrit à la section 3.1. Nous simulons à présent la politique myope développée à la section 4.1 et la comparons à d'autres politiques sous-optimales, dans le but d'évaluer la pertinence de son implémentation. Pour ce faire, nous définissons d'abord les deux politiques auxquelles la politique myope sera comparée. Sachant que la politique myope est optimale, nous cherchons à la comparer à d'autres politiques qui minimisent également les coûts pour une seule période. Remarquons que le principal avantage de la politique myope réside dans le fait qu'elle évite la nécessité de résoudre l'équation de Bellman, ce qui permet par conséquent de réduire le temps de calcul. À titre indicatif, on parle d'un temps de résolution d'environ 2 secondes lorsque $K = 20$ pour la politique myope qui est d'ordre $O(K^4)$.

L'implémentation des différentes politiques est faite sous MATLAB en utilisant le package LAPJAV (*Jonker-Volgenant Algorithm for Linear Assignment Problem*) pour la résolution du problème d'assignation de la politique myope [75].

4.4.1 Définitions des politiques de comparaison

Politique naïve

Nous définissons une première politique qui servira de comparaison, appelée politique naïve (*Naive Approach Policy* (NA)), car elle ne prend tout simplement pas en compte l'état de confiance de l'opérateur dans son calcul. Cette politique est basée sur celle développée dans [3], à l'exception que la réponse w^s de l'humain influence la charge de travail finale. Ainsi, une recherche exhaustive est effectuée sur toutes les possibilités de w^r , en fournissant la liste correspondante π^{w^r} de l'ordre dans lequel l'opérateur effectue les tâches. Le calcul de la politique naïve est réalisé à l'aide de l'Algorithme 1 présenté ci-dessous qui retourne la charge de travail optimale pour l'humain w^* ainsi que l'ordre π dans lequel les tâches devraient être effectuées.

La charge de travail réelle de l'opérateur devient donc $w^h = w^* + w^s$, ce qui diffère de la charge optimale w^* calculée dans tous les cas où $w^s \neq 0$. Étant donné qu'elle ne prend pas en compte la confiance de l'opérateur, on s'attend à ce que cette politique présente des performances plus faibles, car elle dépend fortement de l'influence de w^s sur la charge de travail finale w^h . Toutefois, dans le cas où l'espérance de w^s est près de zéro et que sa variance est faible, la politique naïve devrait avoir une performance similaire à la politique optimale. Les avantages principaux de cette politique sont qu'elle ne nécessite pas la connaissance des différents modèles de la confiance humaine τ et de la réponse w^s en plus d'être d'ordre $O(K)$.

Algorithme 1 : Politique naïve

```

pour  $w \in \{0, 1, 2, \dots, K\}$  faire
    Calculer  $[G(p_k^a, w/K)]_{k \in K}$ ;
    Indicer  $[\pi_1, \dots, \pi_K] \leftarrow \text{Trier} \downarrow ([G(p_k^a, w/K)]_{k \in K})$ ;
    Référencer les  $w$  premières tâches à l'humain :
     $w^r = w$ ;
    Calculer le gain total pour une charge  $w^r$  :
     $\bar{G}(w^r) := \sum_{k=\pi_1}^{\pi_{w^r}} G(p_k^a, w^r)$ ;
fin
 $w^* \leftarrow \arg \max \bar{G}(w)$ 
retourner  $w^*$ ;  $[\pi_1^{w^*}, \pi_2^{w^*}, \dots, \pi_K^{w^*}]$ 

```

Politique de décalage de la moyenne

Nous avons défini en (3.3) le modèle pour la réponse humaine qui considère la moyenne de w^s ainsi qu'une déviation par rapport à la moyenne $\tilde{w}^s$. La politique de décalage de la moyenne (*Mean Shifting Policy* (MS)) est une amélioration de la politique naïve en ce sens qu'elle vient considérer l'espérance du nombre de tâches supplémentaires pris par l'opérateur selon son état de confiance. Après avoir effectué le même algorithme de recherche exhaustive que la politique naïve, la politique MS choisira l'action

$$\tilde{w}^r = w^{r*} - \mathbb{E}[w^s \mid \tau],$$

où w^{r*} est la charge optimale calculée par l'Algorithme 1. La liste de tâches π^w est associée à la charge de travail $\tilde{w}^r$, soit la charge espérée pour l'opérateur, également calculée par l'Algorithme 1. Ainsi, cette politique prend en compte la confiance de l'opérateur tout en étant plus rapide que la politique optimale, qui nécessite un temps de calcul de l'ordre de $O(K^4)$, où K correspond au nombre de tâches à classer. Un autre avantage de cette politique est qu'elle nécessite la connaissance d'un modèle moins complet de la réponse humaine w^s . En effet, on considère seulement la moyenne de la réponse w^s en fonction de l'état de confiance au lieu de devoir considérer l'ensemble de la distribution.

4.4.2 Comparaison des politiques

La comparaison des différentes politiques se fait sur la base des coûts réalisés par l'équipe pour une seule période. Compte tenu de la nature myope de la politique optimale, il devient peu pertinent de comparer l'évolution temporelle des différentes politiques puisqu'au final, elles cherchent toutes à minimiser les coûts réalisés à chaque période.

Nous simulons chacune des politiques pour 50 périodes indépendantes de sorte à minimiser l'effet de stochasticité lié au scénario. Nous considérons l'influence des paramètres λ_1 et λ_2 sur les performances des politiques. Rappelons que ces paramètres sont liés à la distribution de w^s tel que décrit par le modèle présenté en 3.4.2. Dans un premier temps, nous observons l'influence de la variance sur les performances des politiques en maintenant l'espérance w^s fixe. Ensuite, nous simulons les performances des politiques pour différentes valeurs de moyenne de la réponse humaine.

Afin de mieux nous représenter le scénario étudié, nous pouvons associer les simulations pour lesquelles les paramètres λ_1 et λ_2 sont faibles à un état de confiance normal étant donné que l'espérance et la variance de w^s dans ce cas seront moindres et la réponse de l'humain w^s plus prévisible. Au contraire, les cas pour lesquels ces mêmes paramètres augmentent peuvent être associés à un état de sous-confiance. Il est possible de voir directement que notre choix pour le modèle de w^s permet d'étendre la formulation du problème à un état de confiance qui prend plus que deux états. En effet, il suffit de définir un spectre entre l'état de sous-confiance et de confiance normal pour lequel on verrait une évolution des paramètres λ_1 et λ_2 .

Valeurs des paramètres

Les différentes valeurs numériques des modèles utilisés sont présentées dans le tableau 4.1 ci-dessous. La signification de ces paramètres est décrite à la section 3.4. Sauf mention contraire, ces valeurs seront utilisées pour l'ensemble des simulations.

Modèle d'observation de l'automaton	$\sigma_a = 1, d_0 = 2$
Modèle de performance humaine	$\sigma_h = 0.8, d_h = 3, P_{fp} = 0.15$
Paramètres du scénario	$K = 20, p = 0.55, c_{tp} = c_{tn} = 0, c_{fp} = c_{fn} = 1$

TABLEAU 4.1 Valeurs numériques des différents modèles

Justification des paramètres

Nous choisissons les paramètres de performances de sorte que l'humain est plus performant que l'automaton lorsqu'il est moins chargé et moins performant lorsqu'il est surchargé. Nous définissons également les coûts c_{tn} et c_{tp} comme étant nuls afin de simplifier l'analyse du problème et de respecter l'hypothèse 1 concernant les coûts de classification. Le nombre de tâches K a été choisi de sorte à avoir suffisamment de tâches pour éviter qu'une seule erreur soit critique dans l'évaluation des performances (p. ex. une erreur sur cinq tâches équivaldrait à un taux d'échec élevé de 20%), tout en maintenant un nombre de tâches qui correspond vraisemblablement au scénario de l'Exemple 1.

Impact de la variance de w^s sur les performances des politiques

Nous étudions d'abord l'influence de la variance dans la réponse w^s de l'humain sur les performances des différentes politiques. Nous choisissons les paramètres λ_1 et λ_2 de sorte que la moyenne de w^s (c.a.d. $\mathbb{E}[w^s] = \lambda_1 - \lambda_2$) reste constante. Dans les cas où la moyenne anticipée de w^s se rapproche des extrêmes du support de w^s , la moyenne réelle n'est plus exactement $\lambda_1 - \lambda_2$. C'est notamment le cas pour la distribution de la sous-figure en bas à droite de la Figure 3.2 où la moyenne réelle est légèrement plus élevée que $\lambda_1 - \lambda_2$. Dans ces situations, la politique MS effectue alors son décalage par rapport au mode de la distribution, soit la valeur de w^s la plus probable d'être prise par l'opérateur.

La Figure 4.1 présente les résultats de simulation de 50 périodes de travail des trois politiques implémentées pour différents paramètres λ . Notons que les paramètres sont choisis de sorte à conserver la même moyenne de tâche supplémentaire w^s , soit $\mathbb{E}[w^s] = 2$. Nous observons que pour des variances plus faibles, la politique myope et la politique MS ont des distributions somme toute similaires. La politique NA, quant à elle, a des coûts de classification qui sont en général plus élevés et une distribution plus allongée signifiant une plus grande variance dans les coûts réalisés à chaque période.

Afin de pouvoir mieux comparer les performances des politiques, nous présentons sur la Figure 4.2 le coût moyen réalisé à chaque période pour les différentes politiques implémentées. Il est possible de voir clairement que la politique NA a des résultats de performance inférieurs aux deux autres politiques autant par rapport à sa variance qu'à ses écarts types, et ce, pour l'ensemble des variances simulées. Néanmoins, les performances des trois politiques sont détériorées par une augmentation de la variance puisqu'on observe un coût moyen plus élevé pour une variance plus grande. Notons la similitude de performance entre la politique optimale et la politique MS qui sera discutée à la section 4.4.3. Finalement, pour une grande variance dans la réponse humaine w^s , on observe que les politiques sous-optimales NA et MS ont des performances très similaires. Ceci s'explique dû au fait qu'elles ne considèrent justement pas cette variance dans leur calcul.

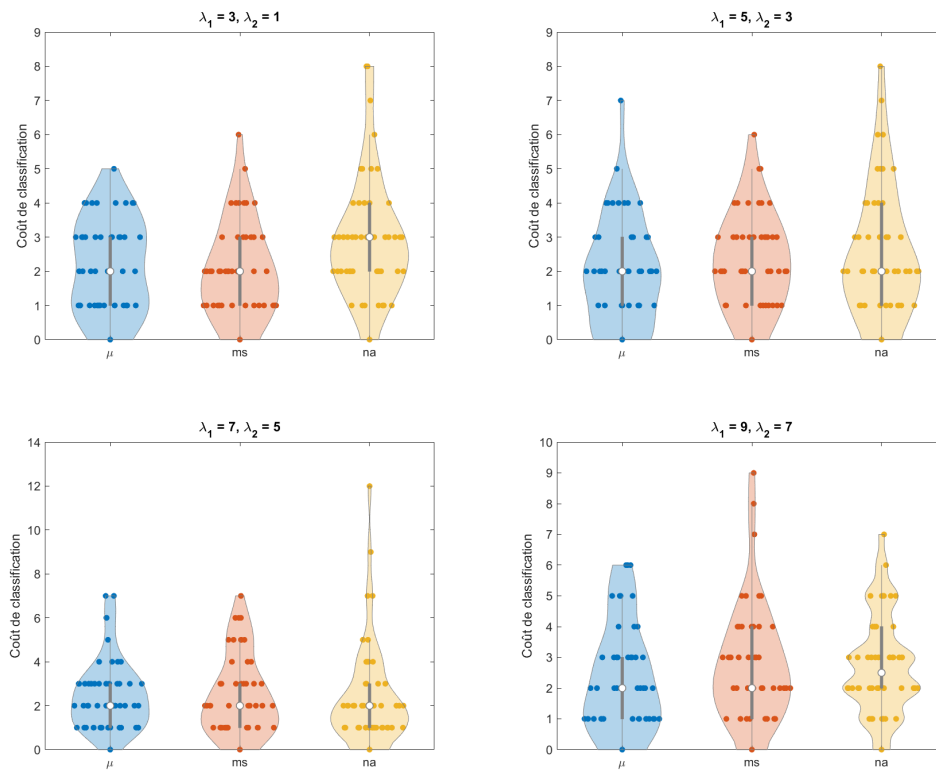


FIGURE 4.1 Coûts de classification pour 50 périodes avec une moyenne de w^s fixe. Les résultats sont présentés sous la forme de diagramme en violon. Le point blanc correspond à la médiane, la ligne grise correspond à la plage interquartile et la densité de la distribution des points est représentée par la forme en violon du diagramme. Les points de couleurs identifient les coûts de classification, soit le nombre d'erreurs, sur $K = 20$ tâches, pour les 50 périodes simulées.

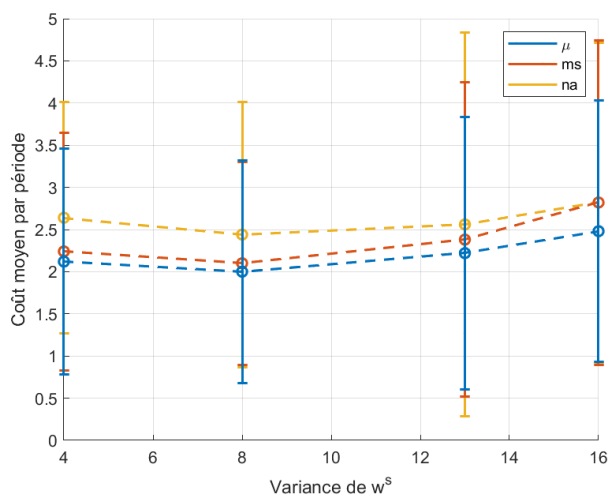


FIGURE 4.2 Coût moyen par période avec une variance croissante. Les cercles représentent la moyenne de coûts réalisés et les barres illustrent l'écart type.

Nous présentons finalement sur la Figure 4.3 l'évolution des coûts cumulatifs des différentes politiques implémentées pour les 50 périodes de simulations, et ce, pour les mêmes paramètres que ceux utilisés à la Figure 4.1.

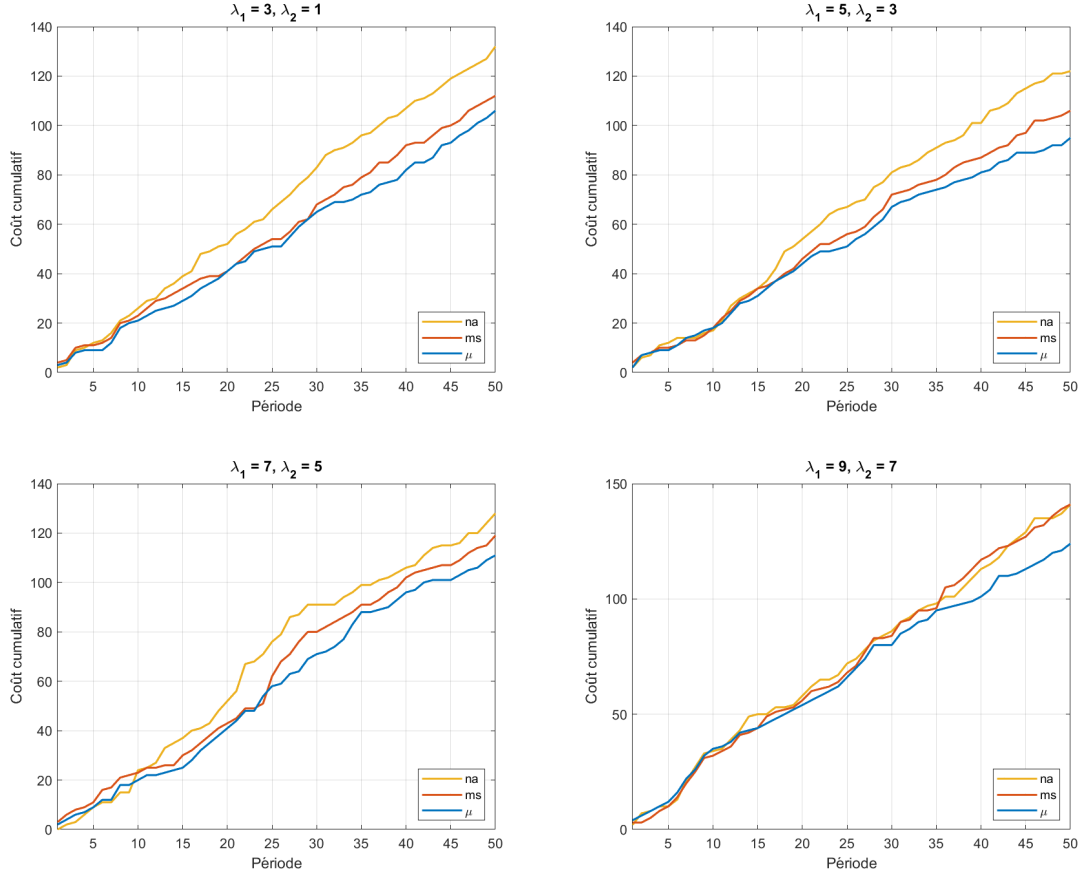


FIGURE 4.3 Coût cumulé pour 50 périodes avec une moyenne fixe.

Nous pouvons donc conclure que pour une moyenne de w^s fixe, la politique myope présente de meilleures performances que les deux autres politiques auxquelles elle est comparée bien que la politique MS puisse présenter des performances similaires pour des variances plus faibles. Nous observons ensuite l'influence d'une augmentation de la moyenne de w^s sur les performances des différentes politiques. Les distributions pour différents paramètres de λ sont présentés dans la Figure 4.4.

Nous générons les mêmes figures que celles présentées à la section précédente à la différence que nous faisons maintenant augmenter la moyenne de w^s .

La Figure 4.5 illustre les diagrammes de violons pour les différents paramètres λ . Nous observons encore une fois une grande similitude dans l'allure des distributions des politiques myope et MS alors que la politique NA présente des coûts de classification plus élevés.

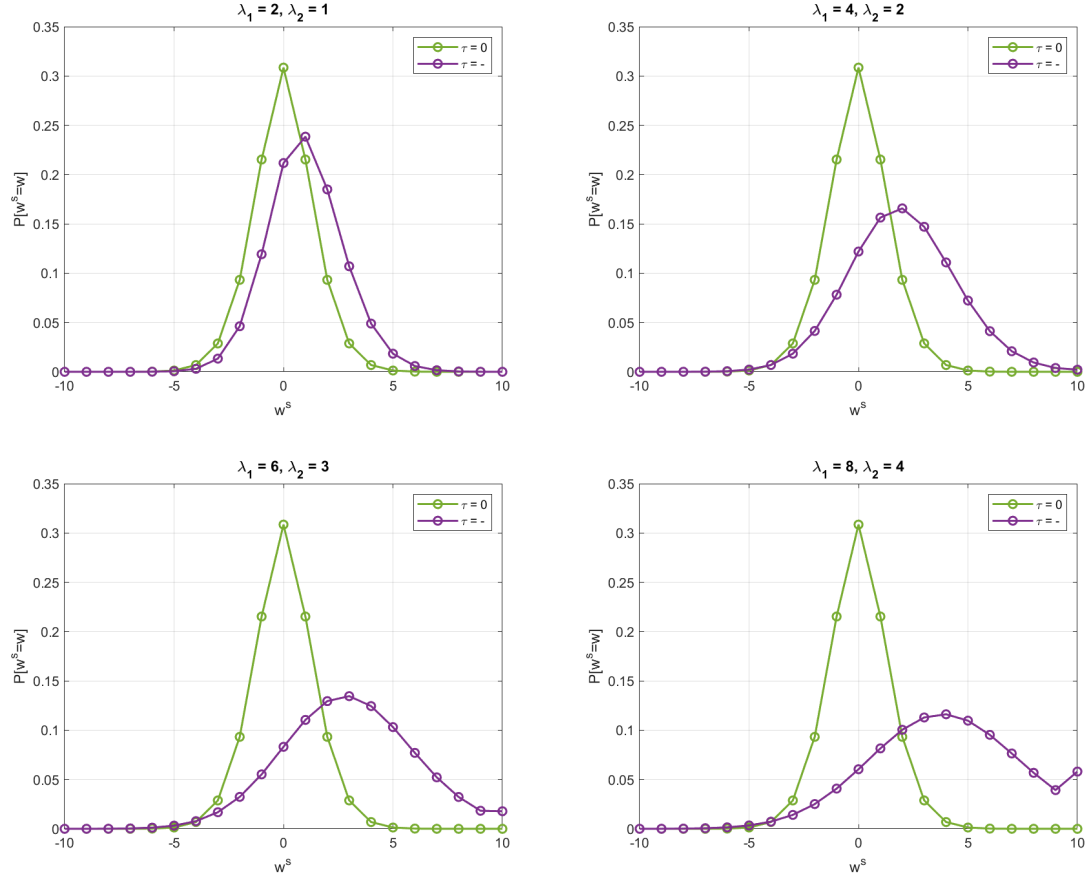


FIGURE 4.4 Distribution des probabilités de w^s avec différentes moyennes.

Nous pouvons confirmer la proximité des performances entre la politique myope et la politique MS avec la Figure 4.6 sur laquelle on observe des coûts moyens très rapprochés pour chaque moyenne de w^s simulées. C'est sans surprise que la performance de la politique NA décroît à mesure que la moyenne de w^s augmente puisque cette politique ne considère tout simplement pas cette moyenne.

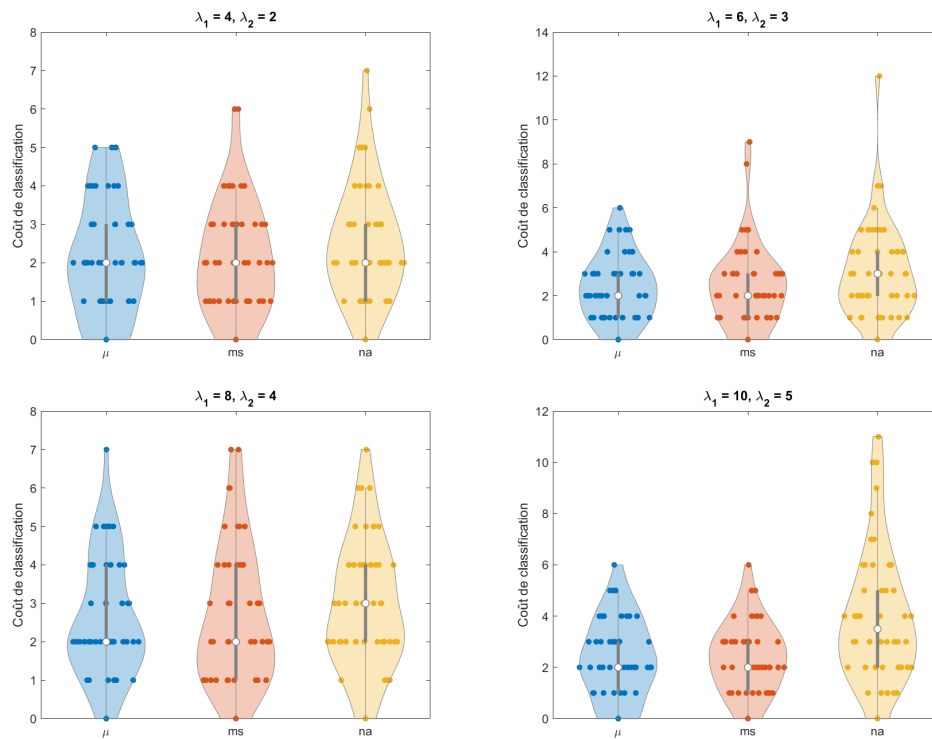


FIGURE 4.5 Coût de classifications pour 50 périodes avec différentes moyennes.

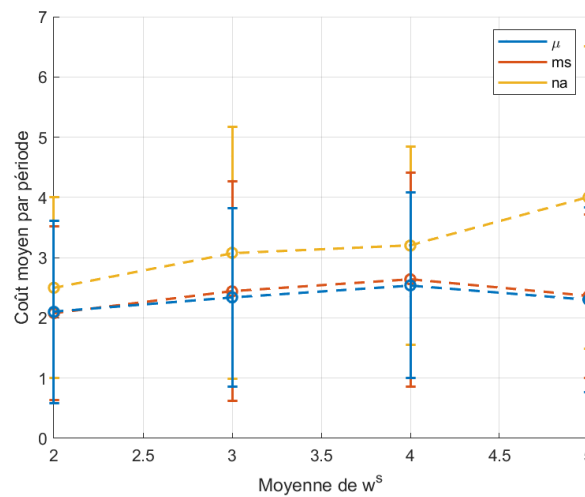


FIGURE 4.6 Coût moyen par période avec une moyenne croissante.

Nous présentons finalement sur la Figure 4.7 les coûts cumulatifs des trois politiques sur les 50 périodes. Nous observons encore une fois la grande similitude de performance entre la politique MS et la politique myope malgré un léger avantage de la politique optimale.

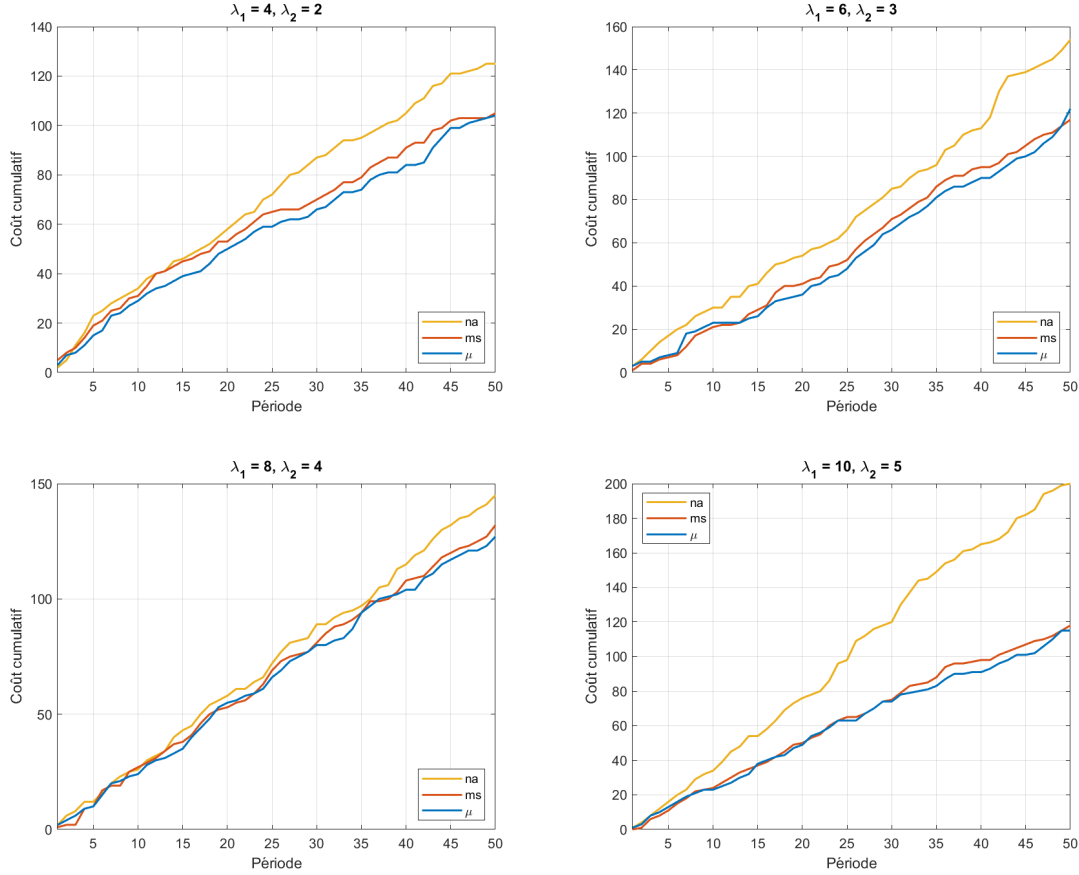


FIGURE 4.7 Coût cumulatif pour 50 périodes avec une moyenne croissante.

4.4.3 Conditions de quasi-optimalité pour la politique MS

Nous remarquons à partir des résultats de simulations présentés précédemment que la politique MS présente des performances similaires à la politique optimale pour l'ensemble des valeurs de λ_1 et λ_2 simulées. Nous nous intéressons à présent aux conditions sous lesquelles la politique MS fournit la même action optimale que la politique myope. En effet, s'il est possible d'identifier des cas pour lesquels la politique MS fournit le même résultat et les mêmes performances que la politique myope, on pourra alors décider de choisir cette politique pour ces cas précis puisque son implémentation est plus simple et plus rapide.

Répartition des postérieurs de l'automation

Dans les cas où la liste de permutation optimale des tâches π ne change pas en fonction de la charge de travail, l'action de l'automation est simplifiée à trouver la charge de travail w^r qu'il est optimal de référer à l'opérateur, puisque la même liste optimale π sera toujours envoyée.

Dans ces cas, nous remarquons à partir des simulations numériques que les w^r obtenus pour

la politique MS sont les mêmes que ceux obtenus pour la politique optimale. Nous présentons un exemple d'un tel cas pour lequel l'ordre des tâches est indépendant de la charge de travail.

Exemple 3. Soit $c_{tp} = c_{tn} = 0, c_{fn} > 0, c_{fp} > 0$. On peut réécrire (3.6) explicitement comme étant

$$\begin{aligned}\Gamma_h(p_k^a, w) &= p_k^a(1 - P_{tp}(w))c_{fn} + (1 - p_k^a)P_{fp}(w)c_{fp} \\ &= p_k^a[(1 - P_{tp}(w))c_{fn} - P_{fp}(w)c_{fp}] + P_{fp}(w)c_{fp} \\ &= \alpha p_k^a + \beta\end{aligned}$$

où $\alpha = (1 - P_{tp}(w))c_{fn} - P_{fp}(w)c_{fp}$ et $\beta = P_{fp}(w)c_{fp}$ sont indépendants de p_k^a . Nous avons donc que $\Gamma_h(p_k^a, w)$ est linéaire en p_k^a . Ce résultat est visible directement sur la Figure 3.1.

Ensuite, nous pouvons réécrire (3.5) comme étant

$$\begin{aligned}C_a(\mathcal{H}_0, p_k^a) &= c_{fn} p_k^a, \\ C_a(\mathcal{H}_1, p_k^a) &= -c_{fp} p_k^a.\end{aligned}$$

Ainsi, la minimisation du coût espéré de l'automation donné en (3.2.3) résulte en deux droites, linéaires en p_k^a qui se croisent à un point maximal $\tilde{p}^a$, i.e.,

$$\begin{aligned}\tilde{p}^a &= \arg \max_{p_k^a} (\min\{C(\mathcal{H}_0, p_k^a), C(\mathcal{H}_1, p_k^a)\}) \\ &= \frac{c_{fp}}{c_{fn} + c_{fp}}.\end{aligned}$$

Nous avons que Γ_h est linéaire en p_k^a , que C_a est linéaire en p_k^a de part et d'autre de $\tilde{p}^a$, faisant en sorte que $G = C_a - \Gamma_h$ est également linéaire de part et d'autre de $\tilde{p}^a$.

Conséquemment, en choisissant les postérieurs $p_{1:K}^a$ de sorte qu'ils se trouvent tous d'un même côté de $\tilde{p}^a$ (tous $\leq \tilde{p}^a$ ou tous $\geq \tilde{p}^a$), nous aurons toujours le même ordre de la liste π lors de l'ordonnancement décroissant des indices G , et ce, peu importe la charge de travail w .

Même lorsque cette condition n'est pas respectée, nous observons dans les simulations que la génération aléatoire de postérieurs produit tout de même des listes qui varient très peu en fonction de la charge de travail. Dans une majorité des cas, nous voyons l'ordre de seulement deux tâches être permuté lorsque la charge de travail augmente d'une tâche supplémentaire. Ainsi, même dans ces cas où il y a effectivement un changement de la liste en fonction de

la charge de travail, la politique MS fournit tout de même des résultats très similaires, voire identiques à ceux de la politique myope développée. Rappelons que la politique MS est plus simple à implémenter et plus rapide dans son exécution que la politique myope qui prend un temps d'ordre $O(K^4)$. Toutefois, en pratique, c'est le fait qu'elle utilise moins d'information sur le modèle humain qui la rend davantage intéressante. Mentionnons également le fait que la politique optimale sert à fournir une borne de performance qui permet ensuite d'évaluer la sous-optimalité de diverses heuristiques.

Il est donc question de se demander si des situations existent pour lesquelles il est effectivement considérablement avantageux d'implémenter la politique myope optimale. Pour répondre à cette question, développons un exemple simple pour lequel la politique MS est sous-optimale.

Exemple 4. *Considérons $K = 6$ tâches réparties de manière égale en deux postérieurs θ_a et θ_b (i.e. 3 tâches en θ_a , 3 tâches en θ_b) de sorte à avoir, pour deux valeurs w_1, w_2 de charge de travail*

$$G(\theta_a, w_1) < G(\theta_b, w_1)$$

$$G(\theta_a, w_2) > G(\theta_b, w_2),$$

c'est-à-dire que pour w_1 , il est préférable de mettre les tâches à θ_a au début de la liste alors qu'on préférera les tâches à θ_b pour w_2 . Nous choisissons $w_1 = 2$ et $w_2 = 3$ pour les simulations de sorte à ce que la liste change complètement entre deux charges de travail consécutives. La Figure 4.8 illustre cette situation.

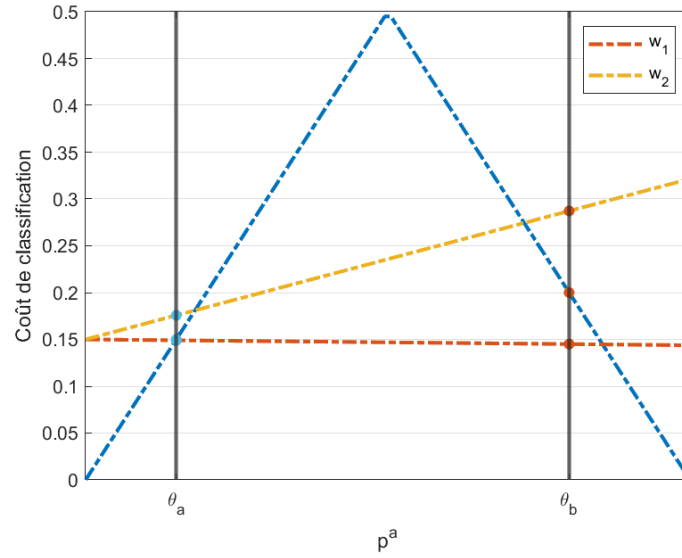


FIGURE 4.8 Gains G à deux postérieurs pour différentes charges de travail w_1 et w_2 . Le gain G correspond à la différence entre la ligne bleue et les lignes jaune et orange. Pour w_1 , on remarque que ce gain est nul à θ_a et positif à θ_b , alors que pour w_2 , le gain est moins négatif en θ_a qu'en θ_b .

Nous comparons les listes fournies par la politique MS et la politique myope en considérant $\mathbb{E}[w^s] = 0$ pour simplifier l'analyse (la politique MS devient donc simplement la politique NA). Notons que π_i correspond à la tâche ordonnée à la i^e position dans la liste alors que a et b correspondent respectivement aux tâches en θ_a et en θ_b .

Politique myope	π_1	π_2	π_3	π_4	π_5	π_6
$w^r \in \{0, 1, 2\}$	b	b	a	a	a	b
$w^r \in \{3, 4\}$	b	a	a	a	b	b
$w^r \in \{5, 6\}$	a	a	a	b	b	b
Politique MS	π_1	π_2	π_3	π_4	π_5	π_6
$w^r \in \{0, 1, 2\}$	b	b	b	a	a	a
$w^r \in \{3, 4, 5, 6\}$	a	a	a	b	b	b

Nous observons que la politique MS ne tient pas compte de la variance de la distribution de w^s . En effet, elle classe simplement les 3 tâches b en premier pour les charges de travail de 2 et moins, et les 3 tâches a en premier pour les charges de travail de 3 et plus. Toutefois, la politique myope semble intuitivement supérieure puisqu'elle considère le fait que bien que $\mathbb{E}[w^s] = 0$, il y a une probabilité non nulle que w^s prenne une valeur différente de zéro, ce qui devrait influencer la liste. Par exemple, pour un $w^r = 2$, si $w^s > 0$ il est souhaitable que la 3^e tâche de la liste soit une tâche en θ_a puisque $G(\theta_a, 3) > G(\theta_b, 3)$.

Le même raisonnement s'applique aussi pour $w^r = 3$ puisque si $w^s < 0$, il est préférable d'avoir une tâche en θ_b dans la liste.

Nous simulons 10 000 périodes avec les deux politiques en appliquant les coûts $c_{fp} = c_{fn} = 1$ et $c_{tp} = c_{tn} = 0$. Nous obtenons les coûts moyens $\mathbb{E}[C^\mu] = 0.164$ et $\mathbb{E}[C^{ms}] = 0.179$ avec des écarts-types de 0.425 et 0.457 pour la politique myope et pour la politique MS, respectivement. En normalisant la différence de la somme des coûts pour l'ensemble des simulations nous obtenons

$$\frac{\sum C^{ms} - \sum C^\mu}{\sum C^\mu} \times 100 = 8.83\%,$$

avec une différence d'écart-type de 1.55%. Puisque nous avons une moyenne de coût plus faible avec un écart-type plus petit pour la politique myope, nous pouvons croire que cette politique est meilleure que la politique MS. Nous avons donc montré par un exemple un scénario pour lequel on trouve un avantage significatif à implémenter la politique myope optimale au détriment de la politique MS plus simple.

4.5 Conclusion

Nous avons montré dans ce chapitre l'optimalité d'une politique myope ainsi que des simulations démontrant une certaine supériorité quant aux performances de cette politique par rapport à d'autres politiques sous-optimales. Toutefois, la politique myope a été montrée comme étant optimale dans le cas où l'état de confiance de l'opérateur est supposé parfaitement connu. En pratique, une telle hypothèse ne peut être totalement validée. C'est pourquoi nous nous intéressons au chapitre suivant au scénario pour lequel l'état de confiance est seulement observable de manière imparfaite.

CHAPITRE 5 OBSERVATION IMPARFAITE DE LA CONFIANCE

5.1 Formulation du problème

En pratique, il est difficile, voire impossible, de connaître de manière parfaite la confiance de l'humain envers l'automatisation. Nous avons toutefois accès à des observations indirectes de l'état de confiance. Nous considérons que les tâches supplémentaires w^s prises par l'humain reflètent son état de confiance à chaque période. Dans ce chapitre, nous proposons une approche exploratoire à partir d'exemples numériques afin de donner une intuition quant à la possibilité du maintien de l'optimalité d'une politique myope dans le cas où la confiance est seulement partiellement observable.

5.1.1 Modèle dynamique de la confiance

À la différence du problème où l'état était parfaitement observable, la dynamique du POMDP repose à présent sur la probabilité que l'humain se trouve en situation de confiance normale en considérant : cette même probabilité à la période précédente, le coût réalisé à la période précédente, le nombre de tâches supplémentaire prises à la période précédente, ainsi que les postérieurs et les actions de l'automatisation. Définissons notre nouvel état comme étant

$$s_t := \mathbb{P}[\tau_t = \mathbf{o} \mid p_{0:t}^a, a_{0:t-1}, w_{0:t-1}^s, c_{0:t-1}], \quad (5.1)$$

et définissons $I_t = \{p_{0:t}^a, a_{0:t-1}, w_{0:t-1}^s, c_{0:t-1}\}$, l'état d'information à la période t juste avant de prendre la décision a_t . Notons le décalage d'indice lié à $p_{0:t}^a$ qui est dû au fait que l'automatisation observe les postérieurs p_t^a avant de poser son action a_t , alors que la réponse de l'humain w_t^s et le coût réalisé c_t surviennent après l'action de l'automatisation pour cette période.

On cherche d'abord à établir la dynamique de cet état, soit la fonction $s_{t+1} = f(s_t, p_{t+1}^a, a_t, w_t^s, c_t)$. De manière similaire au cas de la confiance parfaitement observable, nous considérons connu le modèle de distribution de probabilité $f_{W^s}(w \mid \tau, a) = \mathbb{P}[w_t^s = w \mid \tau_t, a_t]$ des tâches supplémentaires w^s prises en fonction de la confiance τ de l'humain et de l'action $a = \{w^r, \pi\}$ de l'automatisation. Nous rappelons la notation $f_\tau(\tau, c) = \mathbb{P}[\tau_{t+1} = \mathbf{o} \mid \tau, c]$ utilisée en (3.2) pour décrire la probabilité de transition de la confiance de l'humain. On introduit également la notation $f_c(c_t \mid w_t^s, a_t, p_t^a) = \mathbb{P}[c_t \mid w_t^s, a_t, p_t^a]$ qui décrit la probabilité de réaliser un coût c_t étant donné la réponse de l'humain, l'action de l'automatisation, et les postérieurs. La dynamique de

l'état s_{t+1} est donnée par

$$\begin{aligned}
s_{t+1} &:= \mathbb{P}[\tau_{t+1} = \mathbf{o} \mid I_t, p_{t+1}^a, a_t, w_t^s, c_t] \\
&= \mathbb{P}[\tau_{t+1} = \mathbf{o} \mid I_t, a_t, w_t^s, c_t, \tau_t = \mathbf{o}] \mathbb{P}[\tau_t = \mathbf{o} \mid I_t, a_t, w_t^s, c_t] \\
&+ \mathbb{P}[\tau_{t+1} = \mathbf{o} \mid I_t, a_t, w_t^s, c_t, \tau_t = -] \mathbb{P}[\tau_t = - \mid I_t, a_t, w_t^s, c_t] \\
&= f_\tau(\mathbf{o}, c_t) \mathbb{P}[\tau_t = \mathbf{o} \mid I_t, a_t, w_t^s, c_t] + f_\tau(-, c_t) (1 - \mathbb{P}[\tau_t = \mathbf{o} \mid I_t, a_t, w_t^s, c_t]), \tag{5.2}
\end{aligned}$$

où nous utilisons l'hypothèse 4, qui propose qu'une fois le coût c_t réalisé pour une période t , la réponse de l'humain w^s , l'action de l'automatisme a ainsi que les postérieurs p^a n'apportent aucune information supplémentaire quant à l'évolution de la confiance humaine.

Réécrivons le facteur de droite du premier terme de (5.2) en utilisant la loi de Bayes de sorte à avoir

$$\begin{aligned}
\mathbb{P}[\tau_t = \mathbf{o} \mid I_t, a_t, w_t^s, c_t] &= \frac{\mathbb{P}[w_t^s, c_t \mid \tau_t = \mathbf{o}, I_t, a_t] \mathbb{P}[\tau_t = \mathbf{o} \mid I_t, a_t]}{\sum_{i \in \{\mathbf{o}, -\}} \mathbb{P}[w_t^s, c_t \mid \tau_t = i, I_t, a_t] \mathbb{P}[\tau_t = i \mid I_t, a_t]} \\
&= \frac{\mathbb{P}[c_t \mid w_t^s, a_t, p_t^a] \mathbb{P}[w_t^s \mid \tau_t = \mathbf{o}, a_t] s_t}{\chi} \\
&= \frac{f_c(c_t \mid w_t^s, a_t, p_t^a) f_{W^s}(w_t \mid \mathbf{o}, a_t) s_t}{\chi} \tag{5.3}
\end{aligned}$$

où $\chi = \mathbb{P}[c_t \mid w_t^s, a_t, p_t^a] \mathbb{P}[w_t^s \mid \tau_t = \mathbf{o}, a_t] s_t + \mathbb{P}[c_t \mid w_t^s, a_t, p_t^a] \mathbb{P}[w_t^s \mid \tau_t = -, a_t] (1 - s_t)$ peut être vu comme un facteur de normalisation qui assure que $s_t \in [0, 1]$.

En substituant (5.3) dans (5.2), on obtient

$$\begin{aligned}
s_{t+1} &= f_\tau(\mathbf{o}, c_t) \cdot \frac{f_c(c_t \mid w_t^s, a_t, p_t^a) f_{W^s}(w_t \mid \mathbf{o}, a_t) s_t}{f_c(c_t \mid w_t^s, a_t, p_t^a) f_{W^s}(w_t \mid \mathbf{o}, a_t) s_t + f_c(c_t \mid w_t^s, a_t, p_t^a) f_{W^s}(w_t \mid -, a_t) (1 - s_t)} \\
&+ f_\tau(-, c_t) \left(1 - \frac{f_c(c_t \mid w_t^s, a_t, p_t^a) f_{W^s}(w_t \mid \mathbf{o}, a_t) s_t}{f_c(c_t \mid w_t^s, a_t, p_t^a) f_{W^s}(w_t \mid \mathbf{o}, a_t) s_t + f_c(c_t \mid w_t^s, a_t, p_t^a) f_{W^s}(w_t \mid -, a_t) (1 - s_t)} \right) \tag{5.4}
\end{aligned}$$

Notons qu'il peut être difficile de décrire $f_c(c_t \mid w_t^s, a_t, p_t^a)$ de façon analytique. Dans le cadre de ce chapitre qui se veut une exploration numérique du problème, il est possible d'obtenir une distribution de coût pour des paramètres w_t^s , a_t , et p_t^a connus en simulant plusieurs fois une réalisation stochastique du coût pour une période. Pour les mêmes postérieurs p^a , nous appliquons la politique MS décrite en 4.4.1 qui nécessite seulement la connaissance de l'espérance de w^s . Nous fixons la valeur de w^s pour qu'elle soit égale à son espérance pour chaque simulation et observons le coût réalisé. Le fait que la valeur de w^s soit fixée à son espérance pour chaque simulation fait en sorte que la politique MS est optimale pour

une seule période. En effet, cela nous ramène au problème développé dans [3] pour lequel l’optimalité de l’Algorithme 1 a été montré. La Figure 5.1 illustre un exemple de distribution des coûts pour certains paramètres fixés.

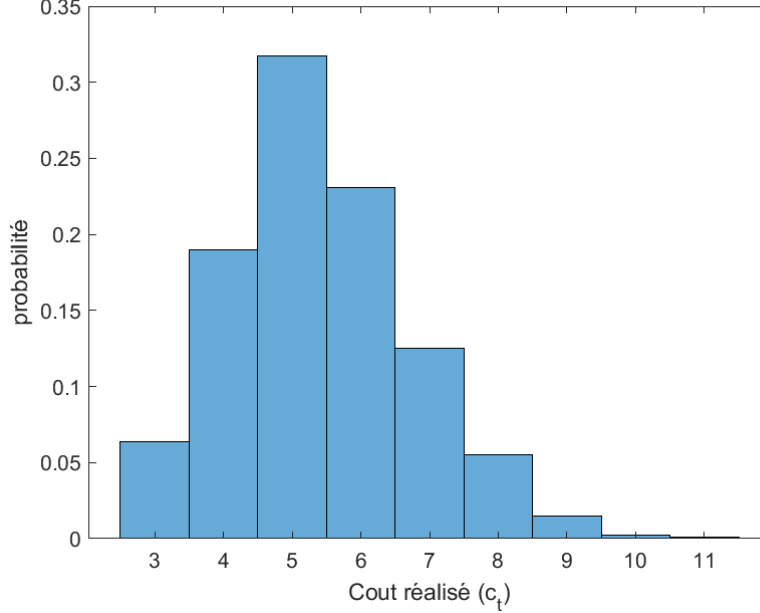


FIGURE 5.1 Distribution du coût réalisé $f_c(c \mid w^s, a, p^a)$ pour 1000 périodes en utilisant la politique MS décrite en 4.4.1 comme action a , avec les mêmes postérieurs p^a et pour un $w^s = 5$ fixe.

En utilisant la distribution de c_t présentée à la figure précédente, il est possible de tracer l’allure de la dynamique de s_{t+1} donnée en (5.4) en fonction du coût et de s_t en fixant les autres paramètres aux mêmes valeurs que celles utilisées à la Figure 5.1. Notons que le modèle utilisé pour $f_\tau(\tau, c)$ respecte l’hypothèse de linéarité utilisée en (4.9). La Figure 5.2 nous porte à croire qu’avec les modèles numériques utilisés, un coût plus élevé résulte en une probabilité plus faible de se trouver dans un état de confiance normale. Des résultats de simulations avec d’autres modèles de $f_\tau(\tau, c)$ nous permettent également de penser que l’allure de s_{t+1} sera non-croissante tant que le modèle de transition $f_\tau(\tau, c)$ est également non-croissant en fonction du coût. En d’autres mots, si on modélise la confiance de l’humain comme étant non-croissante à mesure que le coût réalisé augmente, nous pouvons penser que la dynamique de s_{t+1} sera également non-croissante en fonction du coût réalisé.

5.2 Politique de renvois de tâches

Nous avons montré au chapitre 4 que la politique de renvois de tâche optimale dans le cas où la confiance est parfaitement observable est une politique myope. On cherche à déterminer

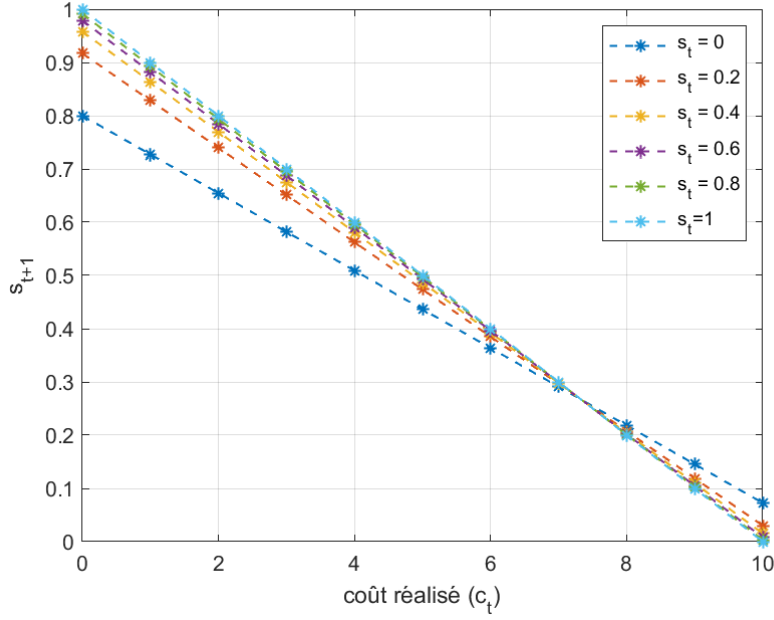


FIGURE 5.2 s_{t+1} en fonction du coût réalisé c_t en utilisant la politique MS décrite en 4.4.1 comme action a , avec les mêmes postérieurs p^a et pour un $w^s = 5$ fixe.

si une telle politique s'avère également optimale dans le cas où la confiance est seulement observable de manière imparfaite.

Rappelons la notation $\bar{G}^\pi(w^r, \tau)$ utilisée pour décrire la fonction objectif (4.2) à maximiser par la politique myope dans le cas du MDP pour une permutation π donnée :

$$\bar{G}^\pi(w^r, \tau) = \mathbb{E}_{w^s} \left[\sum_{i=1}^{w^r + w^s} G(p_{\pi_i}, w^r + w^s) \middle| \tau, w^r, p^a \right]. \quad (5.5)$$

Le problème d'optimisation myope pour le cas du POMDP s'écrit donc

$$\max_{w^r, \pi} \left\{ s_t \bar{G}^\pi(w^r, \tau = \mathbf{o}) + (1 - s_t) \bar{G}^\pi(w^r, \tau = -) \right\} \quad (5.6)$$

où s_t correspond à la probabilité que l'opérateur se trouve dans un état de confiance normal tel que décrit par (5.1).

5.2.1 Impact de la variance sur le POMDP

De manière similaire à ce qui a été fait dans le cas du MDP, nous étudions d'abord l'influence des changements de paramètres λ du modèle de la réponse humaine w^s en maintenant une moyenne fixe et en faisant augmenter la variance. Nous utilisons les mêmes distributions que

celles présentées sur la Figure 3.2.

Analyse du gain espéré

Nous analysons l'espérance du gain $\bar{G}^\pi(w^r, \tau)$ en fonction de l'action de l'automatisme w^r . Nous considérons pour le reste de la discussion que la permutation π est invariante en fonction de la charge de travail de l'humain de sorte que la permutation optimale π^* reste la même pour toute charge de travail w^h de l'opérateur. Cette hypothèse découle de la discussion faite en 4.4.3 où il a été montré que π est indépendant de la charge de travail sous certaines conditions.

Le gain espéré en fonction du nombre de tâches référées w^r et de l'état de confiance τ décrit par (5.5) dépend des paramètres du modèle w^s tel qu'illustré dans la Figure 5.3 ci-dessous. La moyenne et la variance associées aux paramètres λ_1, λ_2 sont décrites à la section 3.4.2.

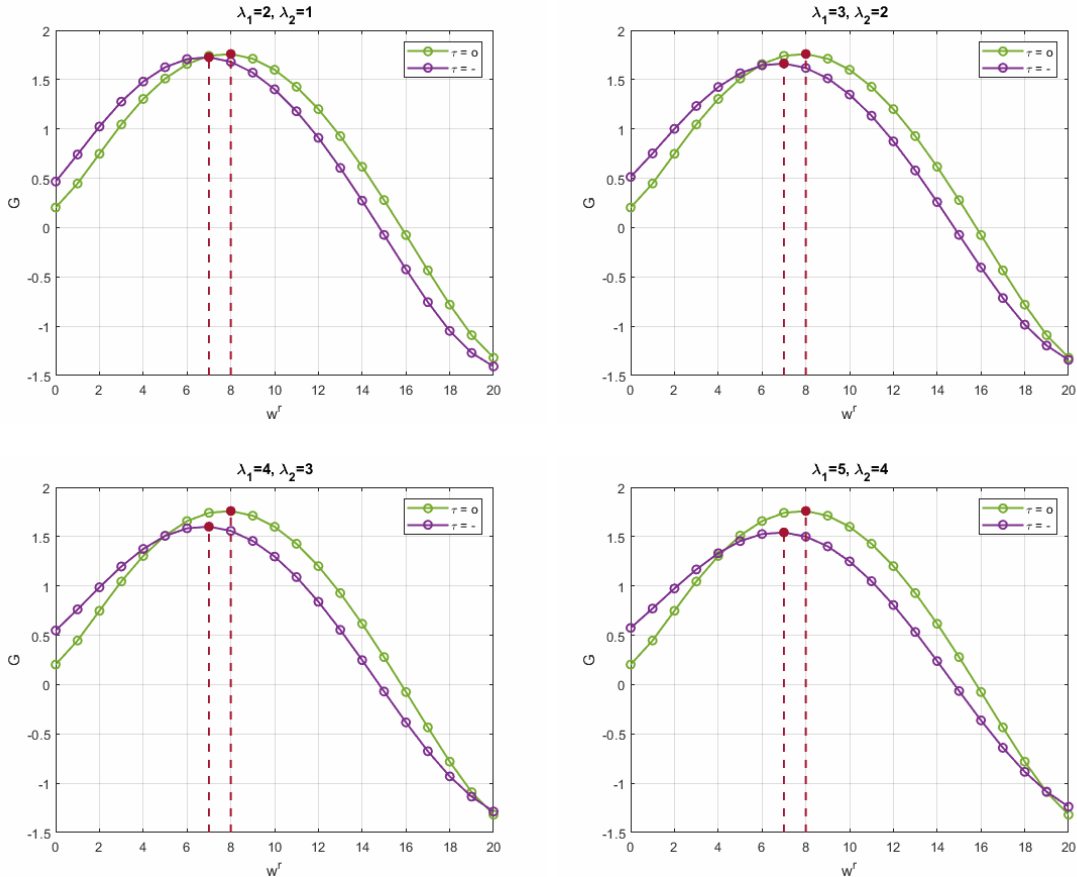


FIGURE 5.3 Gain espéré pour différentes valeurs du modèle de w^s avec une moyenne fixe. Les maximums des différentes courbes sont identifiés en rouge.

Nous observons que pour une moyenne fixe, le maximum du gain espéré demeure au même w^r . Toutefois, plus la variance de w^s augmente, plus la courbe du gain espéré en état de

sous-confiance diminue.

Analyse du problème d'optimisation

Nous introduisons à présent la notation $\hat{G}(s_t)$ pour décrire la valeur maximale du problème d'optimisation (5.6). En considérant une liste optimale fixe π^* pour toute charge de travail w^h , le problème se simplifie à une recherche exhaustive sur tous les w^r possibles en considérant les différentes courbes de $\bar{G}^\pi(w^r, \tau)$. On écrit donc

$$\hat{G}(s_t) = \max_{w^r} \left\{ s_t \bar{G}^\pi(w^r, \tau = \mathbf{o}) + (1 - s_t) \bar{G}^\pi(w^r, \tau = -) \right\}. \quad (5.7)$$

Les Figures 5.4 ci-dessous présentent la valeur maximale $\hat{G}(s_t)$ pour différentes valeurs de s_t en utilisant les modèles w^s avec une moyenne fixe de la Figure 3.2. Nous discrétisons s_t en 100 points sur l'intervalle $[0, 1]$ puisqu'il s'agit d'un état de croyance continu.

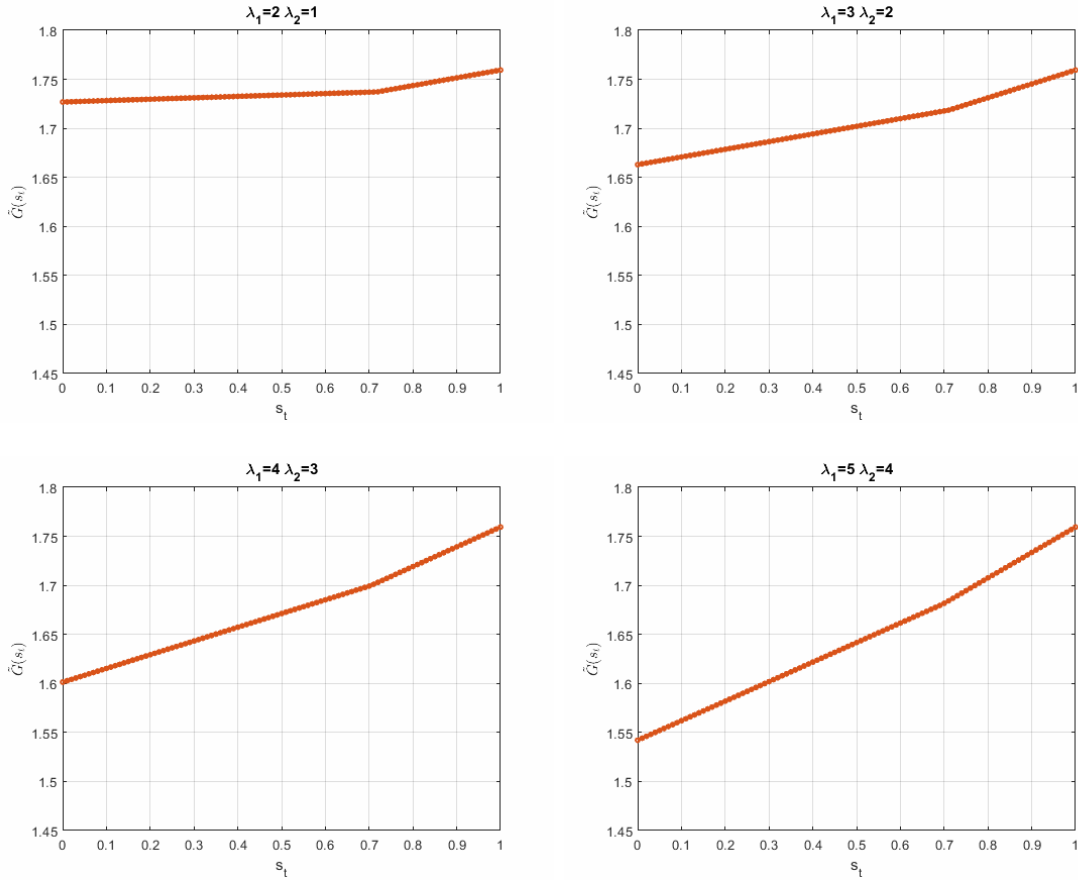


FIGURE 5.4 $\hat{G}(s_t)$ pour différentes valeurs de paramètres du modèle w^s avec une moyenne fixe.

Nous observons pour chacune des sous-figures une fonction non-décroissante en s_t . Il est

possible de remarquer que plus la variance augmente (i.e. plus les paramètres λ sont élevés), plus la différence de gain $\hat{G}(s_t)$ est importante entre l'état de sous-confiance $\tau = -$ et l'état de confiance normal $\tau = \mathbf{o}$. Cela correspond en effet à ce qui a été observé dans les courbes de $\bar{G}(w^r, \tau)$ de la Figure 5.3.

Optimalité de la politique myope pour le POMDP

Il aurait été possible pour le cas du POMDP de résoudre le problème en utilisant un solveur de POMDP afin d'obtenir une politique optimale. Toutefois, une telle méthode utilise des outils numériques complexes qui rendent difficile la compréhension de la politique obtenue. Nous préférons donc utiliser les simulations numériques présentées à la section précédente pour raisonner de manière logique quant à l'optimalité ou non d'une politique myope pour le POMDP.

Rappelons la Définition 1 qui définit la politique myope comme étant une politique qui cherche seulement à minimiser le coût à la période présente. L'état du POMDP s_t correspond à la probabilité que l'opérateur soit dans un état de confiance normal. La Figure 5.2 montre que, en utilisant certains modèles numériques, plus le coût réalisé pour une période est faible, plus la probabilité s_{t+1} de se trouver dans un état de confiance normale à la prochaine période augmente. De plus, il a été montré sur les Figures 5.3 que plus s_t augmente, plus le gain espéré $\hat{G}(s_t)$ est élevé. Conséquemment, il est possible de constater qu'une politique myope qui aurait comme objectif de minimiser le coût réalisé pour une période permettrait d'augmenter la probabilité s_{t+1} d'être en état de confiance, ce qui à son tour augmente le gain espéré à la prochaine période. Nous pouvons émettre la conjecture qu'une telle politique est, dans notre exemple qui utilise des modèles numériques particuliers, encore optimale dans le cas du POMDP sous la condition que la moyenne du modèle w^s reste la même. Nous insistons sur le fait qu'il s'agit ici d'une exploration numérique du problème puisqu'il s'avère complexe de montrer analytiquement l'optimalité de la politique myope dans le cas général du POMDP.

La Figure 5.5 présente cette politique myope, soit la charge de travail w^{r*} qu'il est optimal de référer à l'humain selon l'état s_t .

5.2.2 Impact de la moyenne sur le POMDP

Nous étudions ensuite l'influence d'un changement de la moyenne dans le modèle de la réponse humaine w^s . Nous utilisons les distributions de probabilité présentées à la Figure 4.4 pour le reste de cette section.

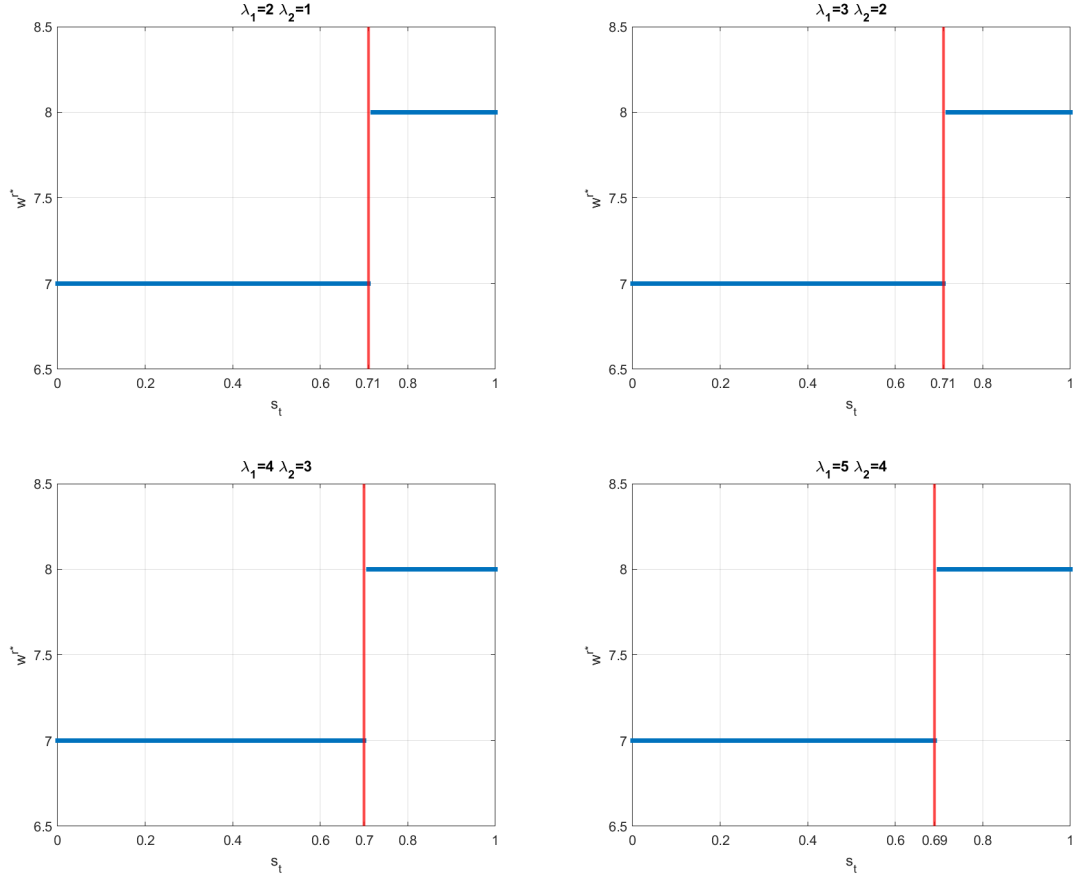


FIGURE 5.5 Politique optimale pour un modèle w^s à moyenne fixe. Nous observons un changement minime, bien que présent, du seuil pour lequel on préfère $w^r = 7$ à $w^r = 8$ pour les différentes variances simulées.

Analyse du gain espéré

Nous avons pour la situation précédente une moyenne fixe faisant en sorte que le w^r qui maximisait $\bar{G}(w^r, \tau)$ demeurerait inchangée malgré une plus grande variance dans la distribution de w^s . La Figure 5.6 montrent les différentes courbes de gains espérés pour différents paramètres λ .

Encore une fois, l'ensemble de la courbe de gain associée à l'état de sous-confiance diminue à mesure que la variance du modèle de w^s augmente. De plus, nous remarquons que la charge de travail référée w^r diminue à mesure que la moyenne de tâches supplémentaires prises par l'opérateur augmente. Conséquemment, nous obtenons une différence de plus en plus importante entre le w^r qui produit le gain maximal dans l'état de sous-confiance et l'état de confiance normal.

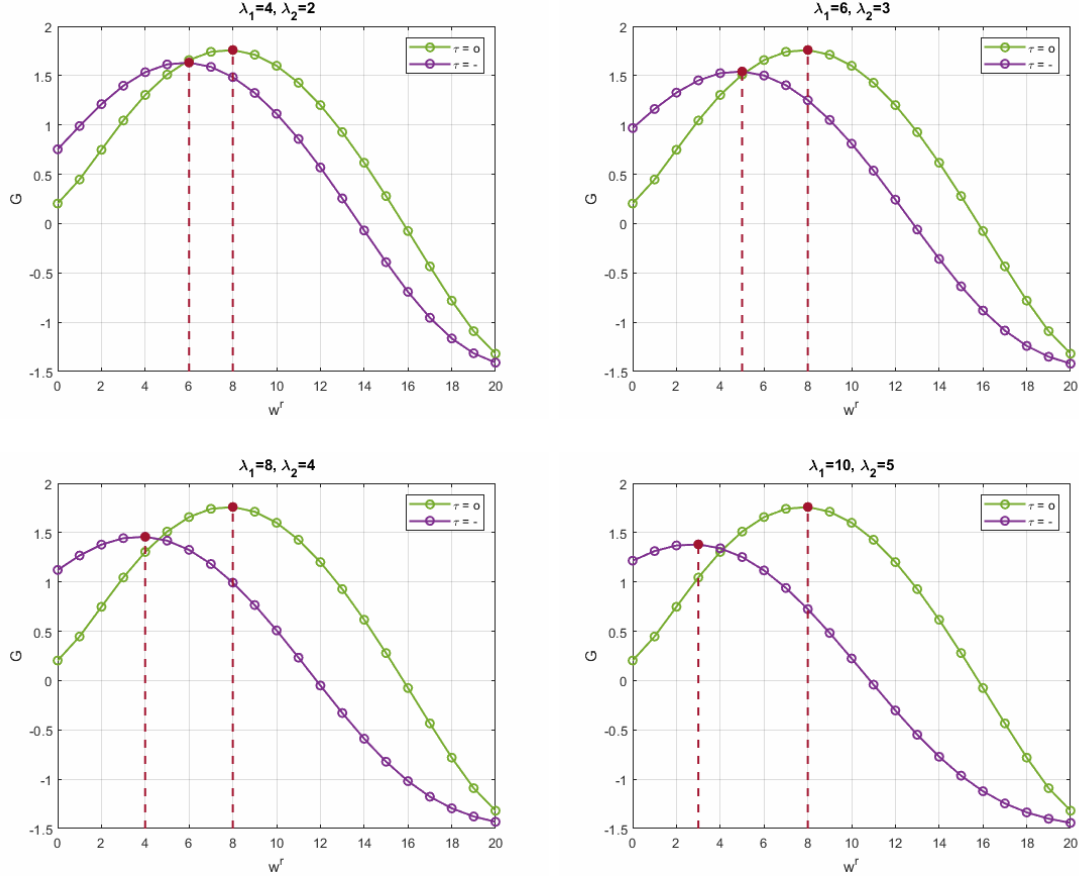


FIGURE 5.6 Gain espéré $\bar{G}(w^r, \tau)$ pour différentes moyennes du modèle de w^s .

Analyse du problème d'optimisation

L'écart grandissant entre les maximums des deux courbes de $\bar{G}(w^r, \tau)$ a un impact direct sur l'allure des courbes de $\hat{G}(s_t)$. Les courbes du gain espéré en fonction de s_t sont présentées sur la Figure 5.7 ci-dessous.

Nous observons que l'allure de $\hat{G}(s_t)$ n'est plus non-décroissante pour des moyennes plus élevées de w^s . En effet, le gain à $s_t = 0$ est supérieur aux gains pour toutes les valeurs de $s_t \in]0, 0.3]$ sur la sous-figure en bas à droite de la Figure 5.7.

Le raisonnement qui avait été effectué à la section précédente pour justifier une conjecture d'optimalité d'une politique myope n'est donc plus directement utilisable dans ce cas. En effet, considérons un cas pour lequel on aurait $s_t = 0$, soit la certitude que l'opérateur est en situation de sous-confiance. Il pourrait être possible que la politique myope, bien qu'elle réduise le coût à la période présente, fasse en sorte qu'à la période suivante l'état devienne $s_{t+1} \in]0, 0.3]$, soit un état pour lequel le gain espéré est inférieur au gain lorsqu'on se trouve à $s_t = 0$. Il n'est donc pas aussi clair de conjecturer qu'il est encore optimal de tout de

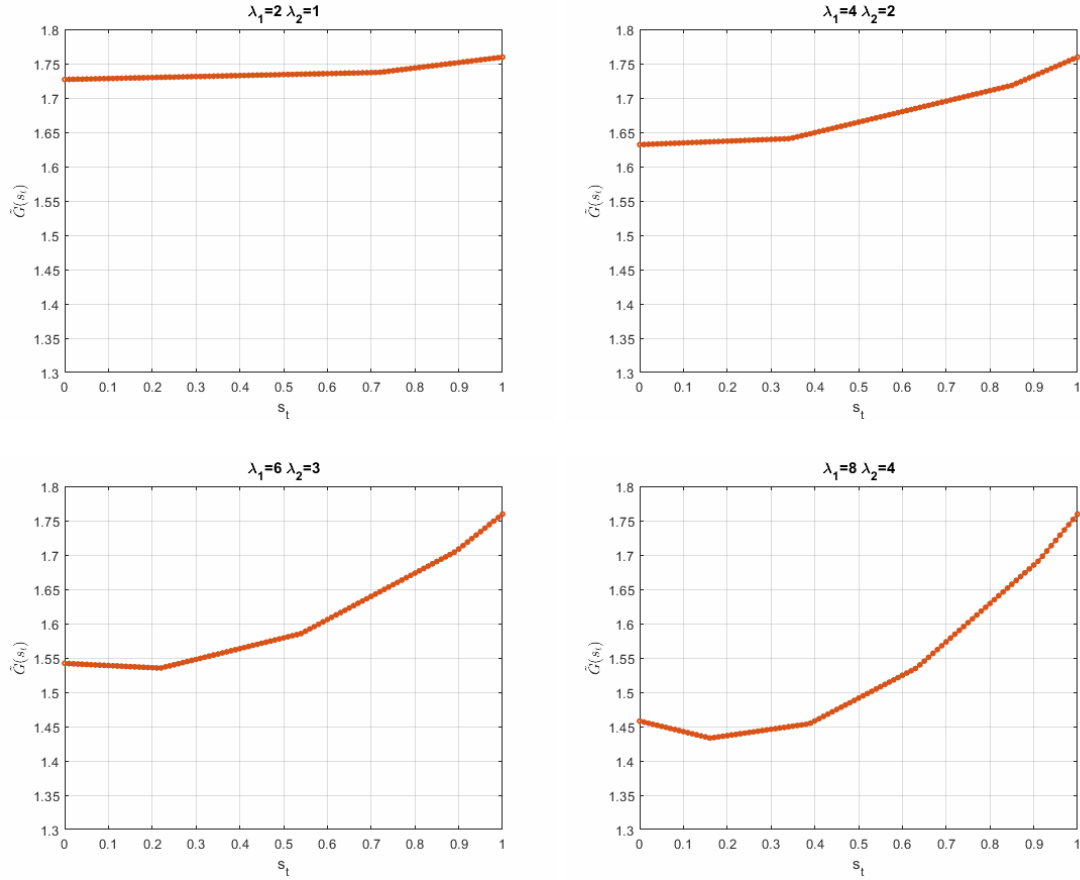


FIGURE 5.7 $\tilde{G}(s_t)$ pour différentes moyennes du modèle w^s .

même minimiser le coût à chaque période. Nous pourrions imaginer une politique qui, dans une telle situation, préfère choisir une action qui ne produit pas un coût minimal de sorte à maintenir l'opérateur dans l'état $s_t = 0$, où l'incertitude sur l'état de confiance est donc faible, et assurer un gain supérieur à la période suivante.

Analyse de la transition s_{t+1}

Nous analysons la fonction de transition de l'état s_t en fonction de w^r afin de déterminer si la situation décrite à la section précédente pose réellement problème et invalide la conjecture d'optimalité d'une politique myope proposée pour le cas du POMDP. La Figure 5.8 ci-dessous présente les courbes de s_{t+1} en fonction de s_t et de w^r où nous fixons w^s à la moyenne attendue en fonction des paramètres λ_1 et λ_2 . Nous utilisons également la distribution des coûts de la Figure 5.1 pour moyenner s_{t+1} selon la probabilité qu'un certain coût c_t se réalise. Ainsi, pour chaque w^r donné et en fixant la réponse humaine w_t^s , il est possible de calculer l'espérance de s_{t+1} selon la distribution de c_t en utilisant la dynamique décrite en (5.4).

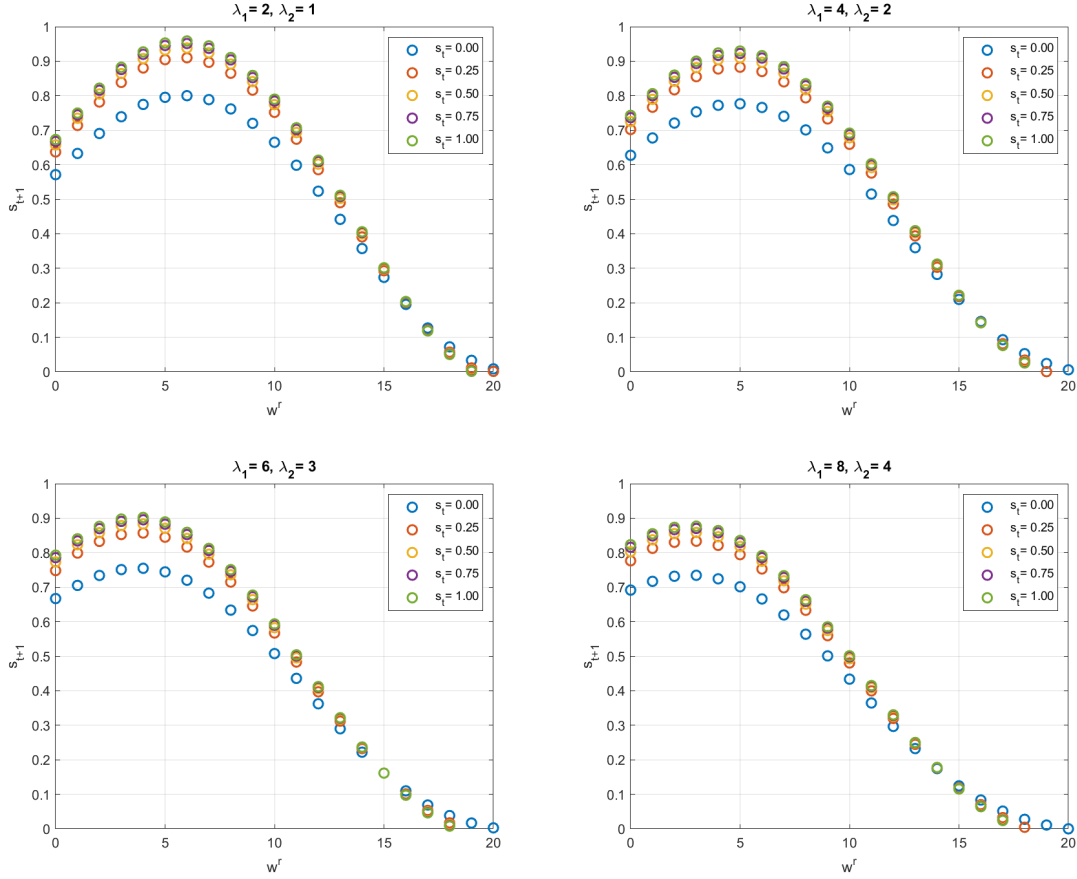


FIGURE 5.8 Probabilité de transition de s_{t+1} pour différentes moyennes de w^s .

Nous observons que la situation problématique décrite précédemment, soit le fait qu'augmenter s_{t+1} pourrait possiblement réduire le gain espéré à la prochaine période, ne se produit pas avec les modèles numériques choisis. En effet, nous remarquons à partir des courbes de s_{t+1} de la Figure 5.8 qu'il est possible de choisir un w^r de sorte à s'assurer que s_{t+1} ne se trouve pas dans l'intervalle $]0, 0.3]$ problématique dont nous avons fait mention sur la Figure 5.7.

Ainsi, nous pouvons utiliser le même raisonnement que pour le cas du POMDP avec une moyenne de w^s fixe pour penser qu'une politique myope pourrait également être optimale dans le cas du POMDP, et ce, même pour un modèle où les moyennes de w^s , dans le cas où l'opérateur est en confiance normale et en sous-confiance, sont éloignées.

Politique myope pour le POMDP

Nous présentons à la Figure 5.9 la politique d'action myope optimale pour différentes moyennes de w^s en rappelant que nous considérons la liste optimale π^* comme étant invariante de sorte à simplifier l'action de l'automatisme à $a = w^r$. Nous remarquons que la politique s'exprime sous forme d'intervalles pour lesquels une action est optimale. En raisonnant de manière

géométrique, nous pouvions nous attendre à un tel résultat puisque la politique myope qui résout le problème défini en (5.6) cherche le maximum des courbes $\bar{G}(w^r, \tau)$ qui sont linéaires en s_t . Ainsi, dans le cas où la liste des tâches ne change pas en fonction de la charge de travail, il suffit de connaître ces intervalles afin de connaître l'action optimale à poser.

5.3 Conclusion

Dans ce chapitre, nous avons effectué une exploration numérique du cas pour lequel la confiance n'est pas parfaitement observable par l'automatisation. Nous avons émis une conjecture quant au maintien de l'optimalité d'une politique myope à partir d'exemples numériques des différents modèles définis dans le problème. Les résultats numériques obtenus nous permettent de penser qu'une politique myope pourrait également être optimale dans le cas du POMDP sous certaines conditions.

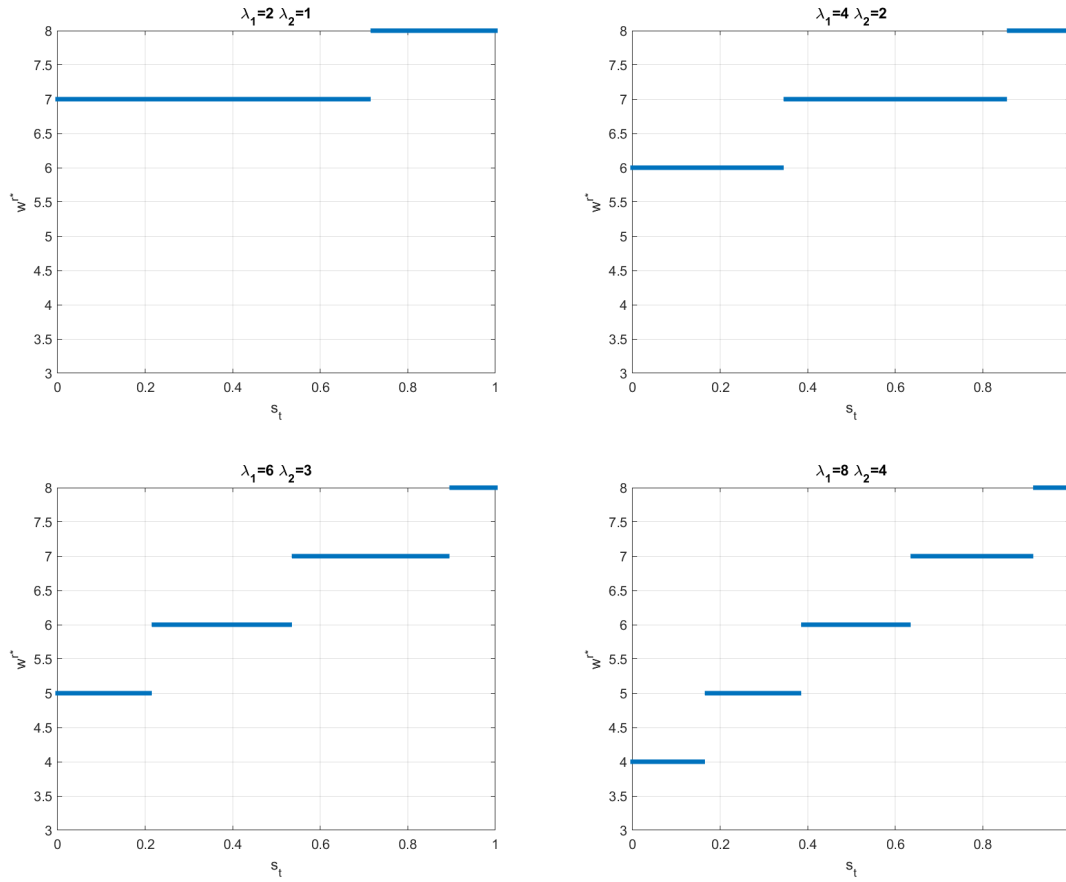


FIGURE 5.9 Politique myope pour différentes moyennes de w^s . Nous observons que l'action optimale lorsque s_t se trouve à ses extrêmes, soit près de $s_t = 0$ et près de $s_t = 1$ correspond aux valeurs de w^r qui maximisent $\bar{G}(w^r, \tau)$ sur la Figure 5.6.

CHAPITRE 6 CONCLUSION

Le dernier chapitre de ce mémoire se concentre sur la synthèse du travail réalisé jusqu'à présent. Les limitations de l'approche avancée sont abordées, et des pistes de réflexion sont proposées pour orienter d'éventuelles améliorations futures.

6.1 Synthèse des travaux

L'objectif de ce travail de recherche était d'améliorer les performances d'une équipe humain-machine en développant une politique de renvois de tâches se basant sur différentes caractéristiques du comportement humain.

Pour ce faire, nous avons d'abord défini un scénario type d'application d'une interaction humain-machine servant de support pour le développement de la politique de renvois de tâches. Le scénario choisi comprend une classification binaire de tâches qui peuvent être réalisées par l'humain ou la machine. Après avoir observé l'ensemble des tâches, l'automation en fournit une liste et renvoie un nombre de tâches qui devraient être effectuées par l'humain à qui revient la décision finale quant aux tâches qu'il décide de compléter. L'automation base son action sur deux facteurs humains dont nous jugeons l'influence sur les performances de l'équipe considérable : la confiance de l'humain envers l'automation et l'impact de la charge de travail sur la performance humaine.

Nous choisissons un modèle à seulement deux états pour prédire la confiance de l'opérateur envers la machine (i.e. sous-confiance et confiance normale) de sorte à simplifier l'analyse du problème. Pour ce qui est de la relation entre la performance de classification de l'humain et sa charge de travail, nous la modélisons par des courbes ROC qui s'inspirent de modèles de la théorie de la détection du signal. Aussi, une distribution de Skellam est utilisée pour modéliser la propension de l'humain à prendre plus ou moins de tâches que celles référées par l'automation.

La nature dynamique de l'état de confiance de l'humain rend naturelle la formulation du problème sous la forme d'un Processus Décisionnel Markovien (MDP) puisque cet outil permet de considérer l'aspect stochastique inhérent à la formulation du problème. La formulation du problème en MDP facilite également son extension à un POMDP pour lequel l'état de confiance n'est qu'imparfaitement observable.

La simplicité de l'état de confiance a permis la démonstration analytique d'une preuve de l'optimalité d'une politique myope, soit une politique qui vise simplement à minimiser les

coûts à chaque période sans considérer l'impact d'une telle action sur l'état futur de l'opérateur. Des simulations numériques ont permis de comparer cette politique myope à d'autres politiques sous-optimales plus simples par rapport auxquelles un avantage considérable quant aux performances de classification a été observé. Cet avantage provient essentiellement du fait que la politique myope développée anticipe à la fois l'espérance, mais également la variance de la réponse humaine dans le calcul de sa stratégie de renvois de tâches. Dans le cas du POMDP, des simulations numériques ont permis de conjecturer quant au maintien de l'optimalité d'une politique myope, même dans le cas où la confiance n'est qu'indirectement observable. Il a toutefois été discuté que cette optimalité dépend de manière importante des modèles numériques implémentés.

Une analyse de l'influence de la variance et de l'espérance du modèle de la réponse humaine a permis de conclure sous quelles conditions on retire un avantage réel à implémenter la politique optimale par rapport aux autres politiques plus simples. Un scénario pour lequel l'état de sous-confiance n'entraîne pas une réponse très différente de l'état de confiance normale pourrait motiver l'utilisation d'une telle politique plus facile à implémenter.

6.2 Limitations de la solution proposée

6.2.1 Modèles quantitatifs utilisés

Une des limitations majeures de ce mémoire relève de la définition de l'état de confiance de l'opérateur. En effet, en considérant seulement deux états de confiance, la formulation du problème ne représente évidemment pas de façon fidèle la complexité du facteur humain qu'est la confiance. La pertinence du travail effectué dans ce mémoire gagnerait en crédibilité si un modèle plus précis de la confiance était utilisé.

La fonction de transition de la confiance qui est considérée par hypothèse comme étant linéaire est également une limitation qui vient affaiblir la portée du théorème développé. Une solution qui ne requerrait pas la linéarité de ce modèle serait davantage pertinente quant à sa portée d'application à des scénarios réels.

6.2.2 Facteurs humains négligés

Bien qu'on ait considéré la confiance et la charge de travail comme étant des facteurs qui influencent les performances humaines, plusieurs autres facteurs n'ont pas été considérés dans le développement de la solution.

Transparence de l'interface de la machine

Plusieurs travaux de la littérature portent sur l'importance de la compréhensibilité de la machine dans son influence sur la confiance humaine [8, 56, 76]. Or, dans le scénario développé, l'opérateur ne reçoit aucune information quant aux informations utilisées par l'automatisation pour prendre ses décisions. Les algorithmes de classification construits par apprentissage machine permettent normalement d'avoir accès à un certain niveau de confiance quant aux classifications qu'elle réalise. La politique de renvois de tâches devrait donc prendre en considération ces indices de confiance comme ayant un impact sur le comportement humain.

Charge mentale

La revue de littérature effectuée distingue la charge mentale de la charge de travail. Néanmoins, ce mémoire considère essentiellement l'impact de la charge de travail sur les performances humaines. En pratique, deux individus pourraient subir une charge mentale associée à des performances différentes pour une même charge de travail. Il pourrait donc être nécessaire d'introduire cette variable dans les modèles utilisés.

Confiance en soi

La confiance en soi est un autre facteur humain dont l'impact sur la confiance envers l'automatisation est documenté dans la littérature [1, 34, 60]. Son impact a toutefois été négligé dans l'ensemble de ce travail.

Fatigue

Dans le cadre du scénario où un humain effectue à chaque période un ensemble de classifications, il semble naturel de considérer l'effet de la fatigue sur les performances de l'humain à long terme. Un opérateur constamment surchargé est à risque de s'épuiser plus rapidement qu'un autre qui aurait une charge de travail adéquate. Il existe dans la littérature des modèles quantitatifs de fatigue et de récupération [77] qu'il serait pertinent de considérer dans la formulation du problème.

6.3 Améliorations futures

6.3.1 Valider les modèles de façon expérimentale

Une des limitations principales de la méthodologie adoptée relève du manque de validation expérimentale des différents modèles utilisés. En effet, il serait plus que pertinent de développer des expériences permettant de calibrer les différents modèles afin de s'assurer de leur validité. Le scénario du POMDP ainsi que la politique associée pourront donc être vérifiés et il sera possible de conclure de manière plus claire quant à la pertinence de l'implémentation d'une politique myope.

6.3.2 Inclure davantage de facteurs humains

La section précédente a relevé différents facteurs humains qui n'avaient pas été considérés dans le cadre de ce travail. Ces différents facteurs ayant déjà été étudiés, il serait pertinent de les inclure dans la modélisation du comportement humain. Par exemple, en ajoutant la charge mentale comme une variable influençant les performances de l'humain, l'automation devrait agir de sorte à ne pas sous-utiliser l'humain et ainsi faire baisser son niveau d'attention ou faire naître de la complaisance.

6.3.3 Comparer la solution à d'autres approches

Les modèles quantitatifs de confiance sont peu nombreux dans la littérature, ce qui témoigne de leur complexité à être définis. Cependant, une fois calibrés, ces modèles offrent l'avantage de permettre la conception d'une stratégie adaptative qui peut être directement implémentée.

D'un autre côté, les méthodes d'apprentissage par renforcement peuvent également développer une stratégie de partage de tâches qui s'adapte à l'état de l'opérateur, sans nécessiter un modèle prédictif du comportement humain. Bien que ces approches se basent uniquement sur les données, une quantité importante d'information est souvent requise pour que les algorithmes convergent vers une solution satisfaisante.

Il serait intéressant de comparer l'approche basée sur un modèle développée dans ce mémoire à celle basée sur une approche «boîte noire» afin de déterminer si elles conduisent à des politiques similaires. L'intégration de données physiologiques dans les algorithmes d'apprentissage pourrait permettre à la machine de mieux comprendre l'état cognitif de l'opérateur, même avant que celui-ci ne se manifeste par un changement de comportement.

RÉFÉRENCES

- [1] J. Gao et J. Lee, “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*, vol. 36, n^o. 5, p. 943–959, 2006.
- [2] Y. Wang, “Regret-based automated decision-making aids for domain search tasks using human-agent collaborative teams,” *IEEE Transactions on Control Systems Technology*, vol. 24, n^o. 5, p. 1680–1695, 2016.
- [3] K. Kaza, J. Le Ny et A. Mahajan, “Decision referrals in human-automation teams,” dans *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE, 2021, p. 2842–2847.
- [4] B. Hyun, P. Kabamba et A. Girard, “Optimal classification by mixed-initiative nested thresholding,” *IEEE Transactions on Cybernetics*, vol. 45, n^o. 1, p. 29–39, 2015.
- [5] W. B. Rouse, “Adaptive aiding for human/computer control,” *Human Factors*, vol. 30, n^o. 4, p. 431–443, 1988. [En ligne]. Disponible : <https://doi.org/10.1177/001872088803000405>
- [6] C. D. Wickens, J. G. Hollands, S. Banbury et R. Parasuraman, *Engineering psychology and human performance*. Psychology Press, 2015.
- [7] K. M. Rabby, M. Khan, A. Karimoddini et S. X. Jiang, “An effective model for human cognitive performance within a human-robot collaboration framework,” dans *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2019, p. 3872–3877.
- [8] S.-Y. Chien, M. Lewis, K. Sycara, A. Kumru et J.-S. Liu, “Influence of culture, transparency, trust, and degree of automation on automation use,” *IEEE Transactions on Human-Machine Systems*, vol. 50, n^o. 3, p. 205–214, 2020.
- [9] W.-L. Hu, K. Akash, T. Reid et N. Jain, “Computational modeling of the dynamics of human trust during human-machine interactions,” *IEEE Transactions on Human-Machine Systems*, vol. 49, n^o. 6, p. 485–497, 2019.
- [10] R. Parasuraman et V. Riley, “Humans and automation : Use, misuse, disuse, abuse,” *Human factors*, vol. 39, n^o. 2, p. 230–253, 1997.
- [11] M. M. Cummings, “Man versus machine or man + machine ?” *IEEE Intelligent Systems*, vol. 29, n^o. 5, p. 62–69, 2014.
- [12] S. G. Hart, “Nasa-task load index (NASA-TLX); 20 years later,” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 50, n^o. 9, p. 904–908, 2006. [En ligne]. Disponible : <https://doi.org/10.1177/154193120605000909>

- [13] M. A. Vidulich et P. S. Tsang, “Mental workload and situation awareness,” dans *Handbook of Human Factors and Ergonomics*, 4^e éd., G. Salvendy, édit. New York : Wiley, 2012, p. 243–273.
- [14] A. DiDomenico et M. A. Nussbaum, “Effects of different physical workload parameters on mental workload and performance,” *International Journal of Industrial Ergonomics*, vol. 41, n^o. 3, p. 255–260, 2011.
- [15] M. S. Young, K. A. Brookhuis, C. D. Wickens et P. A. Hancock, “State of science : mental workload in ergonomics,” *Ergonomics*, vol. 58, n^o. 1, p. 1–17, 2015, pMID : 25442818. [En ligne]. Disponible : <https://doi.org/10.1080/00140139.2014.956151>
- [16] J. T. Croasmun et L. Ostrom, “Using likert-type scales in the social sciences.” *Journal of adult education*, vol. 40, n^o. 1, p. 19–22, 2011.
- [17] R. L. Charles et J. Nixon, “Measuring mental workload using physiological measures : A systematic review,” *Applied Ergonomics*, vol. 74, p. 221–232, 2019. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0003687018303430>
- [18] C. D. Wickens, “Multiple resources and mental workload,” *Human factors*, vol. 50, n^o. 3, p. 449–455, 2008.
- [19] J. Sweller, “Cognitive load during problem solving : Effects on learning,” *Cognitive Science*, vol. 12, n^o. 2, p. 257–285, 1988. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/0364021388900237>
- [20] R. M. Yerkes, J. D. Dodson *et al.*, “The relation of strength of stimulus to rapidity of habit-formation,” 1908.
- [21] M. Corbett, “From law to folklore : work stress and the yerkes-dodson law,” *Journal of Managerial Psychology*, vol. 30, n^o. 6, p. 741–752, 2015.
- [22] K. M. Rabby, M. Khan, A. Karimoddini et S. X. Jiang, “An effective model for human cognitive performance within a human-robot collaboration framework,” dans *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2019, p. 3872–3877.
- [23] P. Hancock et M. Chignell, “Mental workload dynamics in adaptive interface design,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 18, n^o. 4, p. 647–658, 1988.
- [24] W. Rouse, S. Edwards et J. Hammer, “Modeling the dynamics of mental workload and human performance in complex systems,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, n^o. 6, p. 1662–1671, 1993.
- [25] S. Nikolaidis et J. Shah, “Human-robot teaming using shared mental models,” *ACM/IEEE HRI*, 2012.

- [26] X. Fan et J. Yen, “Modeling cognitive loads for evolving shared mental models in human–agent collaboration,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 41, n^o. 2, p. 354–367, 2010.
- [27] P. Robinette, W. Li, R. Allen, A. M. Howard et A. R. Wagner, “Overtrust of robots in emergency evacuation scenarios,” dans *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2016, p. 101–108.
- [28] J. D. Lee et K. A. See, “Trust in automation : Designing for appropriate reliance,” *Human factors*, vol. 46, n^o. 1, p. 50–80, 2004.
- [29] R. Parasuraman, T. Sheridan et C. Wickens, “A model for types and levels of human interaction with automation,” *IEEE Transactions on Systems, Man, and Cybernetics - Part A : Systems and Humans*, vol. 30, n^o. 3, p. 286–297, 2000.
- [30] J. Braithwaite et T. Makkai, “Trust and compliance,” *Policing and Society : An International Journal*, vol. 4, n^o. 1, p. 1–12, 1994.
- [31] H. Saeidi et Y. Wang, “Incorporating trust and self-confidence analysis in the guidance and control of (semi)autonomous mobile robotic systems,” *IEEE Robotics and Automation Letters*, vol. 4, n^o. 2, p. 239–246, 2019.
- [32] K. Akash, G. McMahon, T. Reid et N. Jain, “Human trust-based feedback control : Dynamically varying automation transparency to optimize human-machine interactions,” *IEEE Control Systems Magazine*, vol. 40, n^o. 6, p. 98–116, 2020.
- [33] S. Knocton, A. Hunter, W. Connors, L. Dithurbide et H. F. Neyedli, “The effect of informing participants of the response bias of an automated target recognition system on trust and reliance behavior,” *Human Factors*, vol. 65, n^o. 2, p. 189–199, 2023, pMID : 34078167. [En ligne]. Disponible : <https://doi.org/10.1177/00187208211021711>
- [34] J. D. Lee et N. Moray, “Trust, self-confidence, and operators’ adaptation to automation,” *International Journal of Human-computer Studies*, vol. 40, n^o. 1, p. 153–184, 1994.
- [35] W.-L. Hu, K. Akash, T. Reid et N. Jain, “Computational modeling of the dynamics of human trust during human–machine interactions,” *IEEE Transactions on Human-Machine Systems*, vol. 49, n^o. 6, p. 485–497, 2019.
- [36] S. Rice, D. Keller, G. Hunt et D. Trafimow, “Automation dependency under time pressure,” dans *2009 International Symposium on Aviation Psychology*, 2009, p. 611.
- [37] J. Schwark, I. Dolgov, W. Graves et D. Hor, “The influence of perceived task difficulty and importance on automation use,” dans *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 54, n^o. 19. SAGE Publications Sage CA : Los Angeles, CA, 2010, p. 1503–1507.

- [38] S.-Y. Chien, M. Lewis, K. Sycara, A. Kumru et J.-S. Liu, “Influence of culture, transparency, trust, and degree of automation on automation use,” *IEEE Transactions on Human-Machine Systems*, vol. 50, n^o. 3, p. 205–214, 2019.
- [39] M. Itoh, “A model of trust in automation : Why humans over-trust ?” dans *SICE Annual Conference 2011*, 2011, p. 198–201.
- [40] K. E. Schaefer, “Measuring trust in human robot interactions : Development of the “trust perception scale-hri”,” dans *Robust intelligence and trust in autonomous systems*. Springer, 2016, p. 191–218.
- [41] E. Park et X. Jiang, “Measuring trust of human operators in new generation rescue robots,” *Proceedings of the JFPS International Symposium on Fluid Power*, vol. 2008, 01 2008.
- [42] J.-Y. Jian, A. Bisantz et C. Drury, “Foundations for an empirically determined scale of trust in automated systems,” *International Journal of Cognitive Ergonomics*, vol. 4, p. 53–71, 03 2000.
- [43] P. Kaniarasu, A. Steinfeld, M. Desai et H. Yanco, “Potential measures for detecting trust changes,” dans *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, 2012, p. 241–242.
- [44] A. Freedy, E. DeVisser, G. Weltman et N. Coeyman, “Measurement of trust in human-robot collaboration,” dans *2007 International Symposium on Collaborative Technologies and Systems*, 2007, p. 106–114.
- [45] R. Parasuraman, “Designing automation for human use : empirical studies and quantitative models,” *Ergonomics*, vol. 43, n^o. 7, p. 931–951, 2000, pMID : 10929828. [En ligne]. Disponible : <https://doi.org/10.1080/001401300409125>
- [46] J. Lee et N. Moray, “Trust, control strategies and allocation of function in human-machine systems,” *Ergonomics*, vol. 35, n^o. 10, p. 1243–1270, 1992.
- [47] W.-L. Hu, K. Akash, T. Reid et N. Jain, “Computational modeling of the dynamics of human trust during human-machine interactions,” *IEEE Transactions on Human-Machine Systems*, vol. 49, n^o. 6, p. 485–497, 2019.
- [48] A. Xu et G. Dudek, “Optimo : Online probabilistic trust inference model for asymmetric human-robot collaborations,” dans *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2015, p. 221–228.
- [49] M. S. Cohen, R. Parasuraman, D. Serfaty et R. Andes, “Trust in decision aids : A model and a training strategy,” *Arlington, VA : Cognitive Technologies*, 1997.

- [50] M. E. G. Moe, M. Tavakolifard et S. J. Knapskog, “Learning trust in dynamic multiagent environments using HMMs,” dans *Proceedings of the 13th Nordic Workshop on Secure IT Systems (NordSec 2008)*, 2008.
- [51] Z. Malik, I. Akbar et A. Bouguettaya, “Web services reputation assessment using a hidden markov model,” dans *European Conference on a Service-Based Internet*. Springer, 2009, p. 576–591.
- [52] Z. Wang, A. Peer et M. Buss, “An HMM approach to realistic haptic human-robot interaction,” dans *World Haptics 2009 - Third Joint EuroHaptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 2009, p. 374–379.
- [53] X. Liu et A. Datta, “Modeling context aware dynamic trust using hidden markov model,” dans *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 26, n°. 1, 2012, p. 1938–1944.
- [54] L. Feng, C. Wiltsche, L. Humphrey et U. Topcu, “Synthesis of human-in-the-loop control protocols for autonomous systems,” *IEEE Transactions on Automation Science and Engineering*, vol. 13, n°. 2, p. 450–462, 2016.
- [55] C. Dubois et J. Le Ny, “Adaptive task allocation in human-machine teams with trust and workload cognitive models,” dans *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2020, p. 3241–3246.
- [56] K. Akash, G. McMahon, T. Reid et N. Jain, “Human trust-based feedback control : Dynamically varying automation transparency to optimize human-machine interactions,” *IEEE Control Systems Magazine*, vol. 40, n°. 6, p. 98–116, 2020.
- [57] M. Chen, S. Nikolaidis, H. Soh, D. Hsu et S. Srinivasa, “Planning with trust for human-robot collaboration,” dans *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction*, 2018, p. 307–315.
- [58] U. Jayasinghe, G. M. Lee, T.-W. Um et Q. Shi, “Machine learning based trust computational model for iot services,” *IEEE Transactions on Sustainable Computing*, vol. 4, n°. 1, p. 39–52, 2019.
- [59] M. Yin, J. Wortman Vaughan et H. Wallach, “Understanding the effect of accuracy on trust in machine learning models,” dans *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’19. New York, NY, USA : Association for Computing Machinery, 2019, p. 1–12. [En ligne]. Disponible : <https://doi.org/10.1145/3290605.3300509>

- [60] H. Saeidi et Y. Wang, “Incorporating trust and self-confidence analysis in the guidance and control of (semi)autonomous mobile robotic systems,” *IEEE Robotics and Automation Letters*, vol. 4, n^o. 2, p. 239–246, 2019.
- [61] H. Muslim et M. Itoh, “A theoretical framework for designing human-centered automotive automation systems,” *Cognition, Technology & Work*, vol. 21, p. 685–697, 2019.
- [62] B. Sadrfaridpour, “Trust-based control of robotic manipulators in collaborative assembly in manufacturing,” Thèse de doctorat, Clemson University, 2018.
- [63] A. Xu et G. Dudek, “Maintaining efficient collaboration with trust-seeking robots,” dans *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016, p. 3312–3319.
- [64] M. Chen, S. Nikolaidis, H. Soh, D. Hsu et S. Srinivasa, “Trust-aware decision making for human-robot collaboration : Model learning and planning,” *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 9, n^o. 2, p. 1–23, 2020.
- [65] V. Keswani, M. Lease et K. Kenthapadi, “Designing closed human-in-the-loop deferral pipelines,” dans *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, n^o. 1, 2022, p. 1–9.
- [66] D. Madras, T. Pitassi et R. Zemel, “Predict responsibly : improving fairness and accuracy by learning to defer,” *Advances in Neural Information Processing Systems*, vol. 31, 2018.
- [67] G. Bansal, B. Nushi, E. Kamar, E. Horvitz et D. S. Weld, “Is the most accurate ai the best teammate ? optimizing ai for teamwork,” dans *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, n^o. 13, 2021, p. 11 405–11 414.
- [68] M. Shaked et J. G. Shanthikumar, *Stochastic orders*. Springer, 2007.
- [69] T. D. Wickens, *Elementary signal detection theory*. Oxford university press, 2001.
- [70] J. Bulla, C. Chesneau et M. Kachour, “On the bivariate skellam distribution,” *Communications in Statistics-Theory and Methods*, vol. 44, n^o. 21, p. 4552–4567, 2015.
- [71] T. Fawcett, “Introduction to receiver operator curves,” *Pattern Recognit. Lett*, vol. 27, p. 861–874, 2006.
- [72] H. W. Kuhn, “The hungarian method for the assignment problem,” *Naval Research Logistics (NRL)*, vol. 52, 1955. [En ligne]. Disponible : <https://api.semanticscholar.org/CorpusID:9426884>
- [73] R. Jonker et T. Volgenant, “A shortest augmenting path algorithm for dense and sparse linear assignment problems,” *Computing*, vol. 38, p. 325–340, 1987. [En ligne]. Disponible : <https://api.semanticscholar.org/CorpusID:7806079>

- [74] K. Murota, “Discrete convex analysis,” *Mathematical Programming*, vol. 83, p. 313–371, 1998.
- [75] Y. Cao, “LAPJV - Jonker-Volgenant Algorithm for Linear Assignment Problem V3.0,” <https://www.mathworks.com/matlabcentral/fileexchange/26836-lapjv-jonker-volgenant-algorithm-for-linear-assignment-problem-v3-0>, 2013.
- [76] S. Knocton, A. Hunter, W. Connors, L. Dithurbide et H. F. Neyedli, “The effect of informing participants of the response bias of an automated target recognition system on trust and reliance behavior,” *Human Factors*, vol. 65, n°. 2, p. 189–199, 2023, pMID : 34078167. [En ligne]. Disponible : <https://doi.org/10.1177/00187208211021711>
- [77] L. Ma, D. Chablat, F. Bennis, W. Zhang et F. Guillaume, “A new muscle fatigue and recovery model and its ergonomics application in human simulation,” *Virtual and Physical Prototyping*, vol. 5, n°. 3, p. 123–137, 2010. [En ligne]. Disponible : <https://doi.org/10.1080/17452759.2010.504056>